

A recently evolved diflavin-containing monomeric nitrate reductase is responsible for highly efficient bacterial nitrate assimilation

Wei Tan^{1,2#}, Tian-Hua Liao^{3#}, Jin Wang⁴, Yu Ye¹, Yu-Chen Wei¹, Hao-Kui Zhou³, Youli Xiao^{5*‡a},
Xiao-Yang Zhi^{6*}, Zhi-Hui Shao^{5*}, Liang-Dong Lyu^{2*}, Guo-Ping Zhao^{1,2,3,5,7,8*‡b}

From the ¹Department of Microbiology and Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, Hong Kong, China; ²Key Laboratory of Medical Molecular Virology (MOE/NHC/CAMS), School of Basic Medical Sciences and Department of Microbiology and Microbial Engineering, School of Life Sciences, Fudan University, Shanghai, China; ³Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China; ⁴College of Life Sciences, Shanghai Normal University, Shanghai, China; ⁵CAS Key Laboratory of Synthetic Biology, CAS Center for Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, China; ⁶Yunnan Institute of Microbiology, School of Life Sciences, Yunnan University, Kunming, China; ⁷Bio-Med Big Data Center, Key Laboratory of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institute of Nutrition and Health, Shanghai Institutes for Biological Sciences, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China; ⁸Shanghai-MOST Key Laboratory for Health and Disease Genomics, Chinese National Human Genome Center, Shanghai, China

Running title: *Mycobacterial monomeric assimilatory nitrate reductase NasN*

#These authors contributed equally to the work.

*Correspondence may be addressed to Y. X., or X.-Y. Z. (xyzhi@ynu.edu.cn), or Z.-H. S. (zhshao@sibs.ac.cn), or L.-D. L. (liangdong.lv@gmail.com), or G.-P. Z. (gpzhao@sibs.ac.cn).

‡aThis paper is dedicated to the memory of our very close colleague and friend, Prof. Youli Xiao, who unfortunately passed away on January 30, 2020. Therefore, his email address is not provided.

‡bThis paper is also dedicated to the memory of Prof. Ronald Lamont Somerville of Purdue University, beloved mentor and teacher of Guo-Ping Zhao, who passed away on January 24, 2020.

Keywords: nitrate assimilation, assimilatory nitrate reductase (NAS), diflavin reductase, NasN, *Mycobacterium smegmatis*, mycobacteria, genomic deletion, inorganic nitrogen, electron transfer

ABSTRACT

Nitrate is one of the major inorganic nitrogen sources for microbes. Many bacterial and archaeal lineages have the capacity to express assimilatory nitrate reductase (NAS), which catalyzes the rate-limiting reduction of nitrate to nitrite. Although a nitrate assimilatory pathway in mycobacteria has been proposed and validated physiologically and genetically, the putative NAS enzyme has yet to be identified. Here, we report the characterization of a novel NAS encoded by *Mycobacterium smegmatis* (*Msm*) *Msmeg_4206*, designated NasN, which differs from the canonical NASs in its structure, electron transfer mechanism, enzymatic properties, and phylogenetic distribution.

Using sequence analysis and biochemical characterization, we found that NasN is an NADPH-dependent, diflavin-containing monomeric enzyme composed of a canonical molybdopterin cofactor-binding catalytic domain and an FMN-FAD/NAD-binding, electron-receiving/transferring domain, making it unique among all previously reported hetero-oligomeric NASs. Genetic studies revealed that NasN is essential for aerobic *Msm* growth on nitrate as the sole nitrogen source and that the global transcriptional regulator GlnR regulates *nasN* expression. Moreover, unlike the NADH-dependent heterodimeric NAS enzyme, NasN efficiently supports bacterial growth under nitrate-limiting conditions, likely due to its

significantly greater catalytic activity and oxygen tolerance. Results from a phylogenetic analysis suggested that the *nasN* gene is more recently evolved than those encoding other NASs and that its distribution is limited mainly to the Actinobacteria and Proteobacteria. We observed that among mycobacterial species, most fast-growing environmental mycobacteria carry *nasN*, but that it is largely lacking in slow-growing pathogenic mycobacteria because of multiple independent genomic deletion events along their evolution.

INTRODUCTION

In addition to being an important nitrogen source for plants, nitrate (NO_3^-) is universally used as an inorganic nitrogen source for microorganisms, particularly for soil bacteria such as actinomycetes, as well as marine bacteria such as cyanobacteria, due to fierce competition for limited nutrients in the natural environment (1). In prokaryotes, nitrate is reduced to ammonia (NH_3) through two sequential steps, with the first committed step catalyzed by nitrate reductase (NaR). The assimilatory NaR identified in bacteria, abbreviated hereafter as NAS, is cytoplasmic and reduces nitrate to nitrite (NO_2^-) (2). Then, nitrite is further reduced by assimilatory nitrite reductase (NiR) to ammonia, which is subsequently incorporated into the “organic nitrogen pool”, generally glutamate and/or glutamine, *via* either glutamate dehydrogenase (GDH) or glutamine synthetase-glutamate synthase (GS-GOGAT) routes (3).

The holoenzymes of all prokaryotic NASs characterized to date are highly evolutionarily conserved with respect to their catalytic subunit, which contains a canonical nitrate reduction catalytic domain (CA domain), consisting of three binding subdomains bound with a [4Fe-4S] cluster, a molybdopterin and a *bis*-molybdopterin guanine dinucleotide (MGD), a form of molybdopterin cofactor (MoCo) (Fig. 1). On the other hand, these NAS enzymes are significantly different with respect to the intermolecular transfer of electrons derived from their respective electron donors to the catalytic sites. Therefore, the canonical

prokaryotic NASs are categorized based on the physiological electron donors utilized (2,4,5), *i.e.*, the flavodoxin- or ferredoxin-dependent NAS (generally NarB), and the NADH-dependent NAS that requires an additional electron transfer subunit.

The flavodoxin-dependent heterodimeric NarB (Fv-NarB), represented by the well-studied enzyme from the terrestrial diazotroph *Azotobacter vinelandii* (6), directly acquires electrons from flavodoxin *via* its C-terminal [2Fe-2S] cluster (Fig. 1). In contrast, the ferredoxin-dependent heterodimeric NarB (Fd-NarB) acquires electrons from the photosynthetically reduced ferredoxin-containing a [2Fe-2S] cluster, as identified in the autotrophic cyanobacterium *Synechococcus elongatus* (7,8). Thus, the electron donor (flavodoxin or ferredoxin) functions as a subunit of NarB holoenzymes. Among the NADH-dependent NASs, some are heterodimers, well represented by the NasAC enzyme identified from the rare actinomycete *Amycolatopsis mediterranei* (Ame) (9), composed of a CA domain-containing catalytic subunit (NasA) and a FAD/NAD- and [2Fe-2S] cluster-containing electron transfer subunit (NasC) (Fig. 1). NADH-dependent NASs have also been identified from *Klebsiella pneumoniae* (10) and *Bacillus* species (11,12). Others are heterotrimeric complexes, well represented by *Paracoccus denitrificans* NasBGC, composed of a NasA-like catalytic subunit, a FAD/NAD-containing nitrite reductase subunit, and a [2Fe-2S] cluster-containing electron transfer subunit (13). It is worth noting that the [2Fe-2S] cluster plays an essential role in intermolecular electron transfer for nitrate reduction catalyzed by all of these canonical NASs (8,13). In addition, both heterodimeric and heterotrimeric NADH-dependent NASs carry a FAD/NAD domain with NADH and NADPH-binding sites. However, no prokaryotic NAS capable of utilizing NADPH has ever been reported (2,9,12,13).

Nitrate assimilation has been studied in mycobacteria for more than 50 years (14), including both pathogenic and saprophytic mycobacterial species like *Mycobacterium tuberculosis* (*Mtb*) and *Mycolicibacterium*

smegmatis (*Msm*, formerly *Mycobacterium smegmatis* (15)). A recent report suggests that the membrane-bound respiratory NaR (NarGHJI) of *Mtb* functions in both nitrate respiration and nitrate assimilation (16), establishing a unique strategy for nitrate assimilation in the slow-growing pathogenic mycobacteria. In contrast with the suggested function of *Mtb* NarGHJI, the proposed function of *Msm* NarGHJI in nitrate respiration has yet to be experimentally proven (17) and its function in nitrate assimilation has been explicitly disproved (18,19). In addition, the *narB* homolog in the *Msm* genome has been identified as a nitrate assimilation-unrelated gene (18,19). Therefore, the NAS enzyme responsible for the essential nitrate assimilation function in *Msm* remains unclear.

In this study, we have identified that *Msm* NAS is encoded by the *Msmeg_4206* gene and is essential for *Msm* to assimilate nitrate under aerobic conditions. It is characterized as an NADPH-dependent monomeric NAS containing a canonical CA domain and an FMN-FAD/NAD electron receiving/transferring domain (DF domain) identified in the diflavin reductase family, and is designated NasN. NasN is structurally unique among all of the currently known NASs. Functionally it has significantly greater NAD(P)H-dependent enzyme activity and is more oxygen-tolerant than the NADH-dependent heterodimeric NAS enzyme. Based upon the extremely narrow distribution of *nasN*-containing species across three domains of life, mainly in some environmental bacteria, as well as the uniform phylogenetic clustering of nucleotide sequences homologous to *nasN*, it is reasonable to infer that this enzyme evolved more recently than other NASs. Unlike their fast-growing environmental ancestors, most of the slow-growing pathogenic mycobacteria have lost *nasN* via a few independent genomic deletion events.

RESULTS

The *Msmeg_4206* gene of *Msm* encodes a functional NAS essential for bacterial nitrate assimilation

Previous genetic studies showed that neither the *narGHJI*- nor *narB*-encoded

putative NaR enzyme was responsible for *Msm* nitrate assimilation and suggested that the proposed NAS ought to be another MoCo-containing enzyme (18,19). Actually, the *D806_RS20790* gene encoding a MoCo-containing NaR was annotated in the genome of *Msm* strain MKD8 (20). But for the genome of the reference strain mc²155 sequenced by TIGR (NCBI Genbank accession number NC_008596.1), this protein-coding sequence (*Msmeg_4206*), was annotated as a 4055-bp pseudogene due to a frameshift mutation (deletion of nucleotide G₁₈₅₀ in the gene). In this study, we examined whether *Msmeg_4206* functions in nitrate assimilation. In fact, DNA sequencing did not show the nucleotide G₁₈₅₀-deletion mutation in *Msmeg_4206* in the genome of our mc²155 strain, nor in genomes of any other *Msm* strains (<https://www.ncbi.nlm.nih.gov/genome/genomes/1026>). Also, importantly, the *Msm Msmeg_4206* null mutant strain (mc²155ΔnasN) was clearly defective in nitrate assimilation for growth and was readily complemented by the prototype *Msmeg_4206* but the G₁₈₅₀-deleted mutant was not (Fig. 2A-B). These results suggest that *Msmeg_4206* is involved in nitrate assimilation in *Msm*.

We further measured the nitrogen utilization of the *Msmeg_4206*-null mutant and the wild-type strain in nitrogen defined minimal media to probe the role of *Msmeg_4206* during nitrate assimilation. These two strains grew almost at the same rate with comparable consumption rates against either nitrite or ammonia as the sole nitrogen source, respectively (Fig. S1C-F). However, the disruption of *Msmeg_4206* caused a totally impaired growth of the resulting strain in minimal medium with nitrate as the sole nitrogen source (Fig. S1A), and the concentration of nitrate in the culture medium for *Msm Msmeg_4206* null mutant was almost unchanged over the culturing period (Fig. S1B). This growth defect and inability to utilize nitrate of the *Msmeg_4206*-null mutant in the nitrate-containing minimal medium was rescued when a copy of the prototype *Msmeg_4206* was introduced into the mutant (Fig. S1A-B), suggesting that the disruption of *Msmeg_4206* only affects the utilization of nitrate during

nitrate assimilation. Moreover, the supplement of glutamine could restore the growth of the *Msmeg_4206*-null mutant in the nitrate-containing minimal medium (**Fig. S1G**) although it still lost the ability to assimilate nitrate (**Fig. S1H**), further indicating that the disruption of *Msmeg_4206* only affects the reduction of nitrate during nitrate assimilation. Unlike *Msmeg_4206*, the disruption of *narG* did not affect the growth of the resulting strain in medium with nitrate as the sole nitrogen source (**Fig. 2A**), consistent with previous findings that *Msm* NarGHJI does not function in nitrate assimilation (18,19). Collectively, *Msmeg_4206* is genetically shown to encode the sole functional NAS enzyme catalyzing the first step of nitrate assimilation *in vivo* for *Msm* growth under aerobic conditions, and thus is designated *nasN*.

GlnR directly controls the transcription of nasN in Msm

GlnR is a global transcription regulator for genes related to nitrogen metabolism in high GC Gram-positive actinobacteria (21,22), essential for mycobacterial nitrate/nitrite assimilation (16,23,24). Previous transcriptomic studies have indicated that the expression of *nasN* in *Msm glnR*-null mutant strain is largely turned down, seemingly positively regulated by *GlnR* (23,24). We found that *Msm GlnR* is involved in nitrate assimilation, as shown by the growth defect of *Msm glnR*-null mutant ($mc^2155\Delta glnR$) in the nitrate-containing minimal medium (**Fig. 3A**). To test whether *GlnR* directly regulates *nasN*, we performed an electrophoretic mobility shift assay (EMSA) with a *nasN* promoter region and recombinant *Msm GlnR*. Increasing concentrations of purified *GlnR* protein resulted in shifting bands (**Fig. 3B**), indicating that *GlnR* can directly and specifically bind the *nasN* promoter region *in vitro*. In addition, a 48-nucleotide (nt) *GlnR*-protected DNA sequence in the promoter region was precisely determined by DNase I footprinting assay (**Fig. 3C**). Two conserved *GlnR* boxes (a1-b1 and a2-b2) were identified in the *GlnR*-protected region (**Fig. 3D**), consistent with the *GlnR*-binding consensus sequences characterized in actinomycetes (21). Together, these data

suggest that *nasN* belongs to the *GlnR* regulon for nitrogen metabolism.

NasN confers a growth advantage to Msm under nitrate-limiting conditions compared to that of the canonical NADH-dependent NAS

To investigate whether *NasN* is physiologically different from the canonical *NASs*, such as *Ame NasAC*, an NADH-dependent heterodimeric *NAS* (see **Fig. 1**), we compared the growth phenotype of the complemented strains expressing His-tagged *Msm NasN* or *Ame NasA* using $mc^2155\Delta nasN$ as the host strain. The complemented strains were constructed by specifically employing a constitutive *hsp60* promoter to rule out the interference caused by *GlnR*-mediated regulation. When incubated in minimal medium in the presence of 10 mM nitrate as the sole nitrogen source, both *Ame nasAC* and *Msm nasN* were able to fully restore the growth defect of the *nasN*-null mutant to the wild-type level (**Fig. 2B**). *Ame nasA* alone could not support the growth of the *nasN*-null mutant on nitrate as the sole nitrogen source (**Fig. 2B**), which is consistent with a previous result showing that the *Ame nasA* fails to complement for the nitrate assimilation defect phenotype of *S. coelicolor nasA*-null mutant (9). These results suggest that *NasN* is a functional homolog of *NasAC*.

We further measured the growth of the complemented strains in minimal liquid medium containing high (10 mM) or low (0.5–2 mM) concentrations of nitrate as the sole nitrogen source. As shown in **Fig. 2C**, there was no significant difference in the growth rate between these two complemented strains when incubated in 10 mM nitrate-containing minimal medium. However, when they were inoculated in the minimal medium containing low concentrations of nitrate (0.5–2 mM), the *nasAC*-complemented strain showed a growth delay, *i.e.*, an increase in the culturing time to reach the maximum cell density, compared to that of the *nasN*-complemented strain (**Fig. 2D**). Strikingly, in comparison to the *nasN*-complemented strain, the duration of the lag phase for the *nasAC*-complemented strain became longer when the concentration of nitrate used in the medium decreased. These results

suggest that NasN confers a growth advantage to bacteria over that of NasAC under nitrate-limiting conditions.

We further performed western blot analysis to determine the concentrations of these two NAS proteins *in vivo*. The tested strains were incubated in minimal medium with 1 mM ammonia as the sole nitrogen source at 37 °C for 36 h, which ensured that all tested strains would grow to an OD₆₀₀ of ~1.5 (approximately 2×10⁸ cells/ml) when ammonia in culture media was completely depleted (confirmed by measuring the ammonia concentration in the culture). Then 1 mM nitrate was added to the culture for further incubation at 37 °C (**Fig. 2E**). It is clear that the addition of nitrate supported the further growth of all tested strains to OD₆₀₀ of ~3.0 (approximately 2×10⁹ cells/ml), in contrast with no nitrate addition. However, compared to the *nasN*-complemented strain and the wild-type strain, the *nasAC*-complemented strain exhibited a 12-h growth delay (**Fig. 2E**) with a 2-fold slower growth rate (**Fig. 2F**) and a significantly lower consumption rate of nitrate (**Fig. 2G**) during the period of further incubation. As shown in **Fig. 2H**, after the shift in nitrogen source, the expression levels of NasAC and NasN at various time points were similar to each other. We also measured the growth rate at a lower temperature (30 °C), as this is the optimum temperature for *Ame* NasAC (9). However, a similar growth delay during the further incubation period was observed in the *nasAC*-complemented strain, along with a slower nitrate consumption rate (**Fig. S2A-C**). These results indicate that although both NasN and the heterooligomeric canonical NAS (represented by NasAC) play the same role in nitrate assimilation *in vivo*, the NasN-mediated nitrate reduction seems more efficient physiologically, conferring a growth advantage to bacteria under nitrate-limiting conditions.

NasN is an NADPH-dependent, diflavin-containing monomeric NAS

Sequence similarity searching of the NasN protein by HMMER (25) revealed that its N-terminal region (734 aa) contains the highly conserved CA domain required for nitrate reduction, sharing 53.2%, 44.9%, and 47.1% similarity (calculated by the BLSM62 algorithm)

with the catalytic subunits of three characterized NASs from *S. elongatus*, *Ame*, and *B. subtilis*, respectively (**Fig. S3A-B**). Unlike the canonical catalytic subunit, NasN has an additional C-terminal region (617 aa) which contains a unique FMN-FAD/NAD domain (**Fig. S4**), known to function in receiving/transferring electrons from NAD(P)H to the final physiological acceptor characterized in the diflavin reductase family (26,27) and thus hereafter designated DF domain. Moreover, the typical [2Fe-2S] cluster-binding domain contained in the canonical NASs is not identified in NasN. The purified His-tagged NasN protein (148.9 kDa) heterologously expressed in *Escherichia coli* showed a single band at ~150 kDa on SDS-PAGE (**Fig. 4A**). When subjected to analytical size exclusion chromatography, an apparent molecular mass of ~147 kDa was observed for this protein (**Fig. 4B**), suggesting that the native form of NasN in solution is a monomer. The UV-visible spectrum of purified NasN showed absorbance maxima at 378 and 455 nm (solid line in **Fig. 4C**), which is typical for diflavin reductases and is ascribed to the DF domain (28,29). However, the holoprotein exhibited no apparent spectral features typical of iron-sulfur proteins. Thus we attempted to overexpress the N-terminal region of NasN to prevent the possible spectral interference between flavins and iron-sulfur clusters as described previously (30-32). The truncated NasN (N-NasN) was co-expressed and purified with a His-tag to near homogeneity (**Fig. 4A**). Following anaerobic reconstitution of the iron-sulfur cluster, N-NasN showed a brown color with UV-visible absorbance maxima at ~410 nm (**Fig. 4C**), which is characteristic of the presence of the iron-sulfur cluster. The reconstituted N-NasN was determined to be 3.7 ± 0.6 mol irons and 3.8 ± 0.2 mol sulfurs per mol of N-NasN monomers, and the reconstituted holoprotein NasN contained 3.4 ± 0.2 mol irons and 3.1 ± 0.7 mol sulfurs per mol of NasN monomers, indicating the presence of one [4Fe-4S] cluster in each NasN monomer.

We measured the cytoplasmic NaR activities of the crude cell extracts from *Msm* wild-type, *narG*-null, and *nasN*-null mutant strains under anaerobic conditions and found

that there was no detectable activity in any of cytoplasmic fractions of these three strains when reduced methyl viologen (MV, an artificial chemical reductant) or NADH was used as an electron donor (**Fig. 4D**). However, when NADPH was used as an electron donor, both cytoplasmic fractions from *Msm* wild-type and *narG*-null mutant showed similar detectable NaR activity, but not that of the *nasN*-null mutant (**Fig. 4D**), suggesting that NasN rather than NarGHJI is most likely an active cytoplasmic NaR present in *Msm* during aerobic growth and that it is likely a novel NADPH-dependent NaR. In addition, MV-dependent NaR activity of the purified NasN was readily detected under anaerobic assay conditions (**Fig. 4E**), confirming that NasN is biochemically active *in vitro*. Among the physiological electron donors tested for the purified NasN, NADPH was preferred over NADH, as shown by a 26-fold greater activity with NADPH relative to NADH (**Fig. 4E**), further indicating that NasN is a novel NADPH-dependent NaR.

NasN is a monomeric enzyme lacking the typical [2Fe-2S] cluster involved in electron transfer in the canonical NASs. It is likely that the C-terminal DF domain of NasN functions in electron transfer with a preference for NADPH as its electron donor, which is typical of the diflavin reductases (26,27,33). We, therefore, measured the NaR activity of purified NasN in the presence of FMN and/or FAD and found that both flavins could significantly enhance the NADPH-dependent activity of NasN (**Fig. 4F**). Higher activity of ~10-fold was achieved by adding an appropriate amount of FMN and FAD to the assay cocktail, compared to that without the addition of flavins. This result supports the proposition that the DF domain plays a crucial role in electron transfer for NasN-mediated nitrate reduction. In addition, it further confirms that NasN is an NADPH-dependent, diflavin-containing monomeric NAS. All this is consistent with its unique primary structure, domain organization, and biochemical properties, which are distinct from any of the previously characterized NASs. NasN is thus a novel prokaryotic NAS.

The biochemical properties of NasN were further characterized. NasN was shown to be

oxygen-sensitive, as air-exposure of the purified enzyme resulted in a significant decrease in both NADPH- and NADH-dependent NaR activities (**Fig. 4E**), and thus all further experiments were performed under anaerobic assay conditions. NasN exhibited a maximum reaction rate at pH 7.5, 30 °C in 100 mM phosphate buffer (**Fig. S5A-B**). Although the activity of NasN rapidly declined at 40 °C, the enzyme was stable at temperatures below 30 °C, retaining over 85% activity after incubation for 10 min (**Fig. S5C**).

NasN exhibits greater NAD(P)H-dependent NaR activity and tolerance to oxygen than the canonical NADH-dependent NAS

The steady-state kinetics of NasN against its nitrate substrate fit well to the Michaelis-Menten model, with $K_m = 12 \mu\text{M}$ for nitrate, $V_{max} = 860 \text{ nmol/min/mg}$ and a k_{cat} value of $2,100 \text{ min}^{-1}$ (**Fig. 4G**). Although the K_m value for nitrate of NasN is 2-fold greater than that of cyanobacterial Fd-NarB (~6 μM) (8), it is much less than those of previously reported NADH-dependent NASs with values ranging from 17 to 950 μM (12,13), indicating a stronger binding affinity of NasN to nitrate than that of the canonical NADH-dependent NASs.

The biochemical properties of NasN were compared to those of a canonical NADH-dependent NAS (heterodimeric *Ame* NasAC), by quantitatively determining the NaR activities of cytoplasmic fractions of the crude cell extracts from the complemented strains. Under standard anaerobic assay conditions, the NaR activity present in the NasAC-containing cytoplasmic fraction using either NADPH or NADH as an electron donor was readily observed, whereas no NaR activity could be detected in the NasA-containing cytoplasmic fraction (**Table 1**), demonstrating that the electron receiving/transferring *via* FAD/NAD and [2Fe-2S] cluster is indispensable for NAD(P)H-dependent nitrate reduction catalyzed by the canonical NADH-dependent NASs *in vitro*. In contrast, the NasN-containing cytoplasmic fraction exhibited higher NAD(P)H-dependent NaR activity than that of NasAC-containing cytoplasmic fraction. The NADPH-dependent activity of NasN was especially high with a ~380-fold greater value

than that of NasAC (**Table 1**). The k_{cat} values of NasN were ~450-fold and ~2-fold higher than those of NasAC when using NADPH and NADH as an electron donor, respectively, resulting in ~890-fold and ~1.5-fold greater catalytic efficiencies (k_{cat}/K_m) of NasN than those of NasAC. Therefore, NasN is more efficient than NasAC in NAD(P)H-dependent nitrate reduction *in vitro*. On the other hand, when testing the NaR activity employing MV as an electron donor, NasN showed a much lower activity and catalytic efficiency than did NasAC (**Table 1**). Given that MV can activate NaR activity by donating electrons either directly or via [4Fe-4S] cluster to the catalytic site of NaR (13,34-36), this result indicates that the unique DF domain of NasN may largely impair the electron transfer from MV to its catalytic site compared to transfer towards the much simpler catalytic subunit of other NASs.

All NAS enzymes contain an N-terminal [4Fe-4S] cluster at the catalytic subunit, which can be oxidized into [2Fe-2S] (37,38) or [3Fe-4S] clusters (36) in the presence of oxygen. To determine whether the oxidation of the [4Fe-4S] cluster can affect the enzyme activity of NASs, the NaR activities of air-exposed NasN- or NasAC-containing cytoplasmic fractions were measured. We found that the air-exposure of these anaerobically prepared cytoplasmic fractions resulted in a time-dependent decrease of NAD(P)H-dependent NaR activity, whereas no significant decrease of NaR activity was observed when cytoplasmic fractions were kept anaerobic (**Fig. 4I-J**). This is consistent with previous findings that the oxidation of [4Fe-4S] cluster-containing proteins results in protein instability (38-40). It is worth mentioning that for NasAC, both NADPH- and NADH-dependent activities declined rapidly in air within 30 min, while for NasN, although its NADH-dependent activity was lost rapidly within one hour, 25% of its NADPH-dependent NaR activity was retained after air-exposure for 4 hours (**Fig. 4I-J**). This is consistent with the above assay employing the purified NasN (**Fig. 4E**). No significant decrease in MV-dependent NaR activity was observed in any of the enzymes tested after air-exposure (**Fig. 4H**). Although based on this result, it may be assumed that MV-mediated electron transfer is

independent of the [4Fe-4S] cluster, however, another scenario could be that the strong reducing capacity of MV reconstituted the oxidized [4Fe-4S] cluster to make it reactive. Nonetheless, these results highlight an essential role of the N-terminal [4Fe-4S] cluster in transferring electrons to the MoCo catalytic cavity during NAD(P)H-dependent nitrate reduction by NASs and demonstrate that the NAD(P)H-dependent NAS enzymes are generally sensitive to oxygen. Moreover, the NADPH-dependent, diflavin-containing monomeric NasN has been shown to be more tolerant to oxygen oxidation than the NADH-dependent heterooligomeric NAS enzymes which utilize their [2Fe-2S] cluster for electron transfer, such as *Ame* NasAC tested in this study.

NasN evolved more recently than other NASs and experienced multiple independent losses in the slow-growing pathogenic mycobacteria evolving from their fast-growing ancestors

To characterize the phylogenetic distribution of *nasN* in all organisms, a total of 7007 genomes were obtained *via* tBLASTn search and protein domain analysis, with multiple co-existing or singleton status of nucleotide sequences homologous to *nasN* (the CA, DF or intact *nasN* homologs) from 6307 species (unique NCBI taxonomy ID) belonging to the domains of Archaea (267 species), Bacteria (5668 species) and Eukaryota (372 species) (**Fig. 5A** and **Table S1**).

In contrast to the widely distributed canonical NASs (represented by the CA homologs) identified in 267 species of Archaea, 4838 species of Bacteria, and 3 species of Eukaryota (filamentous fungi, *i.e.*, *Aspergillus oryzae*, *Aspergillus glaucus*, and *Leptosphaeria biglobosa*), the intact *nasN* homologs are distributed merely in a limited number of bacterial species (579 species), wherein 210 species have both the intact *nasN* and CA homologs and 369 species only contain the intact *nasN* homologs (**Fig. 5A** and **Table S1**). In particular, the intact *nasN* homologs are mainly present in the phyla Actinobacteria and Proteobacteria and focused in very limited classes, with 325 species in the class *Actinobacteria* of Actinobacteria and 231 species in three classes of Proteobacteria (22

from *Alphaproteobacteria*, 128 from *Betaproteobacteria* and 81 from *Gammaproteobacteria*). The remaining intact *nasN* homologs are identified in the phyla Firmicutes (12 species), Acidobacteria (5 species), Verrucomicrobia (5 species) and Deinococcus-Thermus (1 species). Notably, most of the intact *nasN* homologs (78%) are present in 11 genera which contain at least 10 *nasN*-containing species (**Fig. 5A**). Intriguingly, among the 369 *nasN*-containing species which do not contain CA homologs, some environmental bacteria such as *Rhodococcus* species (41,42) of Actinobacteria, as well as *Acidobacterium ailaau* (43,44) and *Granulicella mallensis* (44,45) of Acidobacteria, are known to possess assimilatory nitrate reductase activity, strongly supporting the existence of NasN as an alternative pathway, in addition to the canonical NASs, for nitrate assimilation.

The nucleotide sequences homologous to *nasN*, encoding the N-terminal CA domain or C-terminal DF domain of NasN can be individually identified in many canonical NAS enzymes and diflavin reductases, respectively (2,13,26). The wide distribution of the CA homologs in many bacteria and archaea species (**Fig. 5A**, bluish shading) suggests that the canonical NAS containing only the CA domain in the catalytic subunit (see **Fig. 1**), is distributed most prevalently, and maybe the earliest NAS of all prokaryotes. Three phylogenetic trees were constructed based on nucleotide sequences homologous to *nasN*, respectively (**Fig. 5B-D**). The significant similarity in both clustering and phylogenetic branching of the *nasN*-derived CA and DF homologs in their corresponding trees (**Fig. 5B-C**) indicates a similar evolutionary history or even a coevolution event shared by these two homologs. On the other hand, *nasN*-derived CA and DF homologs are distributed as narrowly as is the intact *nasN* (**Fig. 5B-D**). Based on the integration of their similar evolutionary histories and incongruent distribution, one may hypothesize that *nasN* evolved via the fusion of the homologous nucleotide sequences for CA and DF domains relatively late in the history of evolution. This resembles the evolution of some diflavin reductase family proteins, such as the

cytochrome P450 reductase CPRBM3, formed by the fusion of the DF domain to their final physiological electron acceptors (26,46) (see **Fig. 1**).

We noticed that the intact *nasN* homolog is absent in 23 of 153 mycobacterial species with available genome sequences (**Fig. 6**, middle panel and **Fig. S6**). Intriguingly, most of the *nasN*-lacking mycobacteria are slow-growing pathogenic species, such as all members of the tuberculosis-related *Mtb* complex (MTBC) and the leprosy-related *Mycobacterium leprae* group, as well as a few members of the opportunistic infection-related *Mycolicibacter terrae* complex (MTC). Even the three fast-growing species lacking the *nasN* gene are all conditional pathogens, *i.e.*, the goat infection-related *Mycolicibacter algericum* (47) and the human infection-related *Mycolicibacter thermoresistibile* and *Mycolicibacter insubricum* (48,49).

To further delineate these losses which are unanimously found in the pathogenic mycobacteria, we analyzed the *nasN*-related neighboring genomic regions with lengths of maximum 100 genes from the *nasN*-lacking species and their *nasN*-containing close relatives (**Fig. 6**, right panel). Although no known insertion sequences (50) were found within these regions, the difference of the remaining genes after the respective *nasN*-related deletion events indicates that most of these events likely occurred independently. By comparing the *nasN*-related neighboring regions between the closely related genomes, especially in the members of MTBC and MTC, at least eight deletion events can be observed (**Fig. 6**, right panel). Among these events, all of the *nasN*-lacking mycobacterial species in events 1-4 have highly similar remaining genes in the *nasN*-related genomic regions after *nasN* deletion occurred, indicating that they share the same gene loss pattern. On the other hand, we observed that the gene loss patterns of events 5-8 were different from those of events 1-4 (**Fig. 6**, right panel). Thus, the dispersed independent loss of *nasN* along the tree of mycobacterial evolution revealed various mechanisms of *nasN* deletion. In conclusion, our analysis suggests that mycobacterial *nasN* deletion occurred repeatedly and independently at least several

times along the evolutionary pathway and that all of the *nasN*-lacking species are pathogenic (some conditional) and mostly slow-growing.

DISCUSSION

The NADPH-dependent, diflavin-containing monomeric NasN characterized by this study represents a novel type of NAS, which differs from the canonical NASs in its structure, electron transferring mechanism, enzymatic properties, and phylogenetic distribution. The holoenzyme of the canonical NAS enzymes is hetero-oligomeric in nature, composed of one conserved catalytic subunit and at least one electron transfer subunit (which may be shared with the nitrite reductase) or one electron donor subunit (ferredoxin or flavodoxin) (Fig. 1), wherein the [2Fe-2S] cluster plays a crucial role, universally, in intermolecular electron transfer for nitrate reduction (8,13). In contrast, although the N-terminal region of NasN contains a canonical CA domain, it is NADPH-dependent and lacks the [2Fe-2S] cluster-binding domain. Instead, NasN contains a unique DF domain at its C-terminus, which may mediate the intra- and inter-molecular electron transfer from NAD(P)H to its CA domain. Both the physiological and the accompanying biochemical data (Table 1, Figs. 2D-H and 4H-J) indicate that the diflavin-containing monomeric NasN is advantageous over the canonical heterodimeric NasAC in regard to low concentrations of nitrate *in vivo* and tolerance to oxygen. Considering that an increased catalytic activity of a cytochrome P450 system was exhibited by fusion of the DF domain with its catalytic subunit (46,51), the better performance of the monomeric NasN lacking the [2Fe-2S] cluster provides additional evidence to demonstrate the higher efficiency of the direct electron receiving/transferring mechanism mediated by the DF domain.

MV is the most commonly used artificial electron donor for *in vitro* NaR activity assays in many previous studies (6,8-10,12,13). However, so far there have been few reports detecting the NaR activity of bacterial NASs employing their physiological electron donors. This technical obstacle might be attributed to the previous assumption that the canonical

NADH-dependent NASs cannot employ NADPH as their physiological electron donor, even though they all theoretically bear the binding sites for both NADH and NADPH (9-13). Our data, on the other hand, indicated that the reason for lack of reported NADPH-dependent activity of NaR *in vitro* may instead be due to oxidation of the enzyme samples during preparation (Fig. 4I-J). In other words, air-exposure of NAS enzymes during their preparation and activity assay could cause the originally low activity of the oxygen-sensitive NADH-dependent NAS to become undetectable. Another possible factor that should be mentioned is that the presence of large quantities of membrane-bound oxidases in the crude cell extract always consuming the reductive NAD(P)H (52). In our experiments, an excess supply of NAD(P)H as the physiological electron donors was shown to be essential for a successful assay, particularly when the cytoplasmic fraction of the crude cell extract was used as the crude enzyme.

Employing our optimized assay system, NasN not only demonstrated its high level of NADPH-dependent NaR activity in both crude cell extracts and the purified form but also was able to use NADH as an electron donor *in vitro* for catalysis, albeit the resulting activity was much lower than NADPH-dependent NaR activity (Table 1, Fig. 4E-J). We also found that, under the same assay conditions, there was detectable NADPH-dependent NaR activity present in the NasAC-containing cytoplasmic fraction (Table 1), confirming that the FAD/NAD domain of the NADH-dependent NAS functions in NADPH binding and electron transfer. These results also show that the NADPH-dependent NasN is different from the NADH-dependent NAS in their preference of the physiological electron donors, *i.e.*, NADPH and NADH, respectively.

Generally, taxonomic distribution combined with phylogenetic analysis has been adopted for inferring evolutionary events on a large time scale (53-55). The hypotheses concerning the origin of NasN, including CA-DF fusion and late-evolving in bacterial evolution, are jointly supported by the extremely narrow distribution of the intact *nasN* homologs limited to a few taxa of phyla

Actinobacteria and Proteobacteria (**Fig. 5A, D**), and the restricted nucleotide diversity of NasN-related CA and DF homologs, respectively (**Fig. 5B-C**). Furthermore, according to the metadata from biosample of *nasN*-containing bacterial species, most of these bacteria (mainly belonging to Actinobacteria and Proteobacteria) are ubiquitously isolated from the environment, such as soil and water (**Table S2**), implying that the high-efficiency of NasN-mediated nitrate assimilation in the environmental bacteria might enforce a probable strong selection favoring the expansion of the *nasN*-encoding population in the corresponding species. The incongruent distribution of *nasN* between the fast-growing and slow-growing mycobacteria (**Figs. 6 and S6**) coincides well with the hypothesis that the slow-growing mycobacteria originated from the ancestral fast-growing environmental mycobacteria (56-58). The *nasN*-containing bacteria may have a growth advantage by efficiently assimilating nitrate when competing with other bacteria under nitrate-limiting conditions, as indicated by our *in vivo* growth studies (**Fig. 2E**) and by the observation that most of the *nasN*-containing bacteria are isolated from the environment (**Table S2**). Therefore, these widely distributed multiple independent *nasN*-loss events mainly in the species of slow-growing pathogenic mycobacteria seem unlikely to be evolutionarily random. In fact, many intracellular bacterial pathogens such as *Mtb* have evolved contrasting lifestyles within the host cell for persistence in the host and for infecting new hosts, rather than the inorganic nitrogen source-dependent rapid growth physiology (59). In addition, *Mtb* has evolved multiple strategies to utilize organic nitrogen sources derived from their hosts rather than the rare nitrate (60).

NarGHJI of *Mtb* strongly reduces nitrate under both anaerobic and aerobic conditions (16,61,62) and thus, is the sole NaR responsible for both nitrate assimilation and respiration (16). This enzyme is also found to be involved in the adaptation of *Mtb* to multiple intracellular stresses (60). In contrast, the NarGHJI homolog in *Msm* has no detectable activity in a rapid anaerobic dormancy model (61) and is believed to exhibit weak activity in a hypoxia dormancy

model induced by slow oxygen consumption (17). Furthermore, our data confirm that *Msm* expresses NasN rather than NarGHJI for highly efficient nitrate assimilation under aerobic conditions (**Fig. 4D**). These major differences in nitrate metabolism *in vivo* between the pathogenic and environmental mycobacteria may shed light on our traditional research strategy employing *Msm* as a laboratory model for *Mtb* physiology studies, particularly helping to adjust to utilizing more appropriate model systems regarding nitrate metabolism, both assimilatory and respiratory.

Experimental procedures

Bacterial strains, plasmids, and culture conditions

The strains and plasmids used in this study are summarized in **Table S3**. *E. coli* strains were grown aerobically at 37 °C in Luria-Bertani (LB) broth. *Msm* strains were grown aerobically at 37 °C in Middlebrook 7H9 broth or 7H10 agar plate supplemented with 0.2% v/v glycerol and 0.05% v/v Tween-80. *Ame* U32 was grown aerobically at 37 °C in Bennet medium as described previously (9). To analyse nitrogen assimilation in *Msm* strains, a nitrogen-free modified *M. phlei* medium (63) was prepared (MPLN minimal medium at pH 6.6: 36.7 mM KH₂PO₄, 7.7 mM sodium citrate, 5 mM MgSO₄, 2% v/v glycerol, 0.05% v/v Tween-80, supplemented with trace elements solution containing 60 μM FeCl₃, 20 μM NaMoO₄, 3.4 μM CaCl₂, 3.5 μM ZnSO₄ and 4 μM CuSO₄), to which different nitrogen sources were added when needed: NaNO₃ (0.5-10 mM), NaNO₂ (1 mM), NH₄Cl (1-10 mM), or L-glutamine (5 mM), respectively. When required, antibiotics were used at the following concentrations: kanamycin, 25 and 50 μg/ml; hygromycin, 50 and 100 μg/ml for *Msm* and *E. coli*, respectively. All chemicals used in this study were obtained from Sigma-Aldrich, USA.

Construction of *Msm* mutant and complemented strains

The *nasN* gene (*Msmeg_4206*) was disrupted in *Msm* mc²155 and replaced with the hygromycin-resistance cassette by allelic exchange as described previously (64). Briefly, two fragments of 1.2-kb and 1.0-kb containing the upstream and downstream regions of *nasN*,

respectively, were amplified from *Msm* mc²155 genomic DNA by PCR. The primers for gene knockout are listed in **Table S4**. The PCR products digested with HindIII/NheI and AvrII/AflII were subsequently inserted into the corresponding sites of the pYUB854 vector, resulting in the knockout plasmid pYUB854-*nasN*. Then mc²155 competent cells were transformed and screened for hygromycin-resistant colonies, and further verified by PCR and DNA sequencing. *Msm* mc²155Δ*narG* and mc²155Δ*glnR* mutant strains were constructed in the same way as the mc²155Δ*nasN* mutant strain (**Fig. S7A-F**).

For complementation experiments, the *Msm nasN* gene together with its native promoter was cloned into an integrating vector pMV306. pMV306-*nasN*_{*Msm*} was used as the template for site-directed mutagenesis of *nasN*, yielding plasmid pMV306-*nasNdelG*_{*Msm*} with a deletion of nucleotide G₁₈₅₀ in *nasN*. To constitutively express proteins in *Msm*, pMV306H was constructed by inserting the *hsp60* promoter from pMV261 into pMV306. For the quantitative measurement of the NaR activity, His-tagged *Msm nasN*, *Ame nasAC*, and *nasA* were inserted into pMV306H, respectively. The complementary recombinants were screened by hygromycin- and kanamycin-resistant screening after introducing the DNA sequencing-verified plasmids to the mc²155Δ*nasN* strain.

Measurement of growth and nitrogen utilization

Msm strains were grown aerobically at 37 °C in 7H9 broth with shaking at 180 rpm overnight until the culture reached the stationary growth phase. The overnight cultures were washed twice with the nitrogen-free MPLN medium by centrifuging at 3,000 × g for 10 min and then adjusted by using MPLN medium to an OD600 of ~1.0 (corresponding to ~1×10⁸ cells/ml) as the seeding cultures. For the spot dilution assay, the inocula were prepared by 10-fold serial dilutions. Two microliters of dilutions ranging from 10⁻¹ to 10⁻⁵ were spotted onto MPLN agar plates containing 10 mM NaNO₃ and then incubated at 37 °C for 4 days. For the growth phenotype test, the seeding cultures were inoculated (1:200, v/v) into 50 ml

7H9 broth or liquid MPLN medium containing different nitrogen sources. Then the cultures were incubated aerobically at 37 °C with shaking at 100 rpm. The growth of these cultures was monitored by measuring OD600 and the growth rates were determined by regression analysis. Nitrate and nitrite concentrations in cultures were measured by using the Nitrite/Nitrate Assay Kit (Roche, Germany) with a detection limit of 0.02 mg/l nitrate/nitrite. Ammonia and glutamine concentrations in cultures were measured by using the Ammonia Assay Kit and the Glutamine/Glutamate Determination Kit (Sigma-Aldrich, USA) with detection limits of 0.2 mg/l ammonia and 0.02 mg/l glutamine, respectively.

Enzyme preparation and protein concentration determination

To purify *Msm NasN*, the full-length *nasN* gene was PCR amplified from *Msm* mc²155 genomic DNA using the primers listed in **Table S4** and cloned into the pET28b(+) vector. The recombinant plasmid pET-*nasN* was transformed into *E. coli* BL21(DE3) for expressing N-terminal His-tagged NasN protein. The recombinant *E. coli* cells were grown aerobically at 37 °C in LB broth supplemented with 50 µg/ml kanamycin and trace elements solution as mentioned above. Protein expression was induced by adding isopropyl β-D-thiogalactoside (IPTG) into the culture to a final concentration of 0.5 mM when the cell density reached an OD600 of 0.6–0.8. After continuous cultivation for an additional 12 h at 16 °C, the cells were harvested by centrifugation at 5,000 × g for 15 min at 4 °C and resuspended in a lysis buffer pH 7.4, containing 50 mM sodium phosphate, 0.5 M NaCl, 10% v/v glycerol, 20 mM imidazole and 1 mM phenylmethylsulfonyl fluoride (PMSF). It should be noted that all operations were carried out under anaerobic conditions from initial sample preparation to the final analysis of the purified protein. Cells were disrupted by sonication on ice followed by centrifugation at 15,000 × g for 60 min to remove cell debris and the insoluble fraction. Then the supernatant was loaded onto a Ni-Sepharose Fast Flow column (GE Healthcare, Sweden) pre-equilibrated with

lysis buffer. After washing with lysis buffer supplemented with 50 mM imidazole, recombinant NasN protein was eluted with lysis buffer supplemented with 300 mM imidazole and 1 mM DTT. Fractions containing NasN were combined and concentrated using 10-kDa cutoff ultrafilter (Merck, Ireland) and then loaded onto a Hiload 16/60 Superdex 200 prep grade column on an AKTA FPLC system (GE Healthcare, Sweden) for further purification by size exclusion chromatography at a flow rate of 1 ml/min and eluted with elution buffer pH 7.2, containing 50 mM phosphate, 0.15 M NaCl, 5% v/v glycerol. The purified protein was concentrated to ~5 mg/ml maintained in the elution buffer by ultrafiltration

and stored in small working aliquots at -70 °C. For gel filtration analysis, the Gel Filtration Calibration Kits (GE Healthcare, UK), were used. These standard proteins including Aprotinin (6.5 kDa), Ovalbumin (44 kDa), Conalbumin (75 kDa), Aldolase (158 kDa) and Ferritin (440 kDa) were dissolved in the same buffer as purified NasN protein and loaded onto the column with the same volume. The apparent molecular mass and the oligomeric state of NasN were determined from its elution volume relative to those of the standard proteins plotted against the logarithm of their molecular masses. The N-terminal region of NasN (N-NasN) was prepared in a similar way as the holoprotein.

For the preparation of cytoplasmic fractions, *Msm* strains were grown aerobically at 37 °C in MPLN medium containing 10 mM nitrate and trace elements solution as mentioned above, to which 5 mM L-glutamine was added when needed. The cells of late-log phase cultures with OD₆₀₀ 0.8-1.0 were harvested by centrifugation at 5,000 × g for 10 min at 4 °C, washed and resuspended using 100 mM sodium phosphate buffer (pH 7.5) containing 10% v/v glycerol and 1 mM PMSF, and sonicated under anaerobic conditions, then the cytoplasmic cell fractions were collected as the crude enzyme extracts after centrifugation at 16,000 × g for one hour at 4 °C.

The Bradford method was used to determine the concentration of the purified recombinant NasN protein and the total protein concentration of cytoplasmic fractions of the

Msm strains. The concentration of His-tagged protein in cytoplasmic fractions was determined by ELISA using an anti-His tag monoclonal antibody, with recombinant NasN protein as a standard. Primary antibodies against His-tag (1:2,000 dilution; mouse monoclonal, M20001, Abmart, Shanghai, China), GroEL (1:5,000 dilution; mouse monoclonal, ab20519, Abcam, Massachusetts, USA), and HRP-conjugated goat anti-mouse secondary antibody (1:5,000 dilution; M21001, Abmart, Shanghai, China) were used.

Cofactor analysis

UV-visible absorbance spectrum (anaerobic, sealed cuvette) was obtained using a NanoDrop ND1000 spectrophotometer (Thermo Scientific, USA). To increase [4Fe-4S] cluster content in protein, an iron-sulfur cluster reconstitution was performed as described previously (65). Iron and acid-labile sulfide contents were measured as described by Pierik et al (66).

Nitrate reductase activity assay

Reaction conditions including temperature, pH, cofactor and amount of enzyme were optimized. Standard NAD(P)H-dependent NaR activity was assayed at 30 °C for 10 min under anaerobic assay conditions in a reaction mixture of 0.5 ml volume which included 100 mM phosphate buffer (pH 7.5), 10 mM NaNO₃, and appropriate amounts of purified or crude enzymes, in the presence of electron donors of 400 μM NADPH or 400 μM NADH, plus 25 μM FMN and 25 μM FAD. For the determination of MV-dependent NaR activity, 150 μM MV and 12 mM dithionite solution (Na₂S₂O₄/NaHCO₃) were added to the assay mixture to start the reaction under anaerobic conditions. After 10 min of reaction time, the reaction was stopped by vigorous stirring under aerobic conditions until the blue color had disappeared in the reaction mixture. The concentration of accumulated nitrite was determined by using the Nitrite/Nitrate Assay Kit (Roche, Germany). NaR specific activity is expressed as nmol of nitrite per min per mg of proteins (nmol/min/mg). For NaR kinetics experiments, the initial velocities of nitrate reduction were plotted against nitrate concentrations and analyzed using non-linear regression to the Michaelis-Menten equation

using GraphPad Prism 5.0 program.

Electrophoretic mobility shift assay (EMSA) and DNase I footprinting assay

Msm GlnR (*Msmeg_5784*) was expressed and purified as described by Lin et al (67). The 294-bp promoter region and 302-bp coding region of *Msm nasN* were amplified by PCR and cloned into the HincII site of the pUC18H vector (Tolo Biotech., Shanghai, China), using the primers listed in **Table S4**. Then the obtained plasmids were verified by DNA sequencing and used as the templates for preparation of the EMSA probes by PCR using primers 6-carboxyfluorescein (FAM)-labeled M13F (-47) and M13R (-48). To investigate the binding of GlnR to the labeled probes, EMSA was performed as described previously (21). Briefly, in a 20 μ l-volume reaction system, 50 ng of FAM-labeled probe was incubated at 30 °C with varying amounts of purified GlnR protein for 30 min in a binding buffer containing 50 mM Tris-HCl (pH 8.0), 100 mM KCl, 2.5 mM MgCl₂, 1 mM DTT, 10% glycerol and 100 μ g/ μ l sheared salmon sperm DNA. The FAM-labelled probe from the coding region of *nasN* was used as a negative control. After electrophoresis, gels were scanned directly with an ImageQuantTM LAS 4000 Imaging system (GE Healthcare, UK).

DNase I footprinting assay was carried out by Tolo Biotech., following the same procedure and condition as previously described (68). Specifically, the labeled-probe was prepared in the same way as described above in the procedure of EMSA. Then, 400 ng of the FAM labeled-probe was incubated with different amounts of recombinant GlnR in a total volume of 40 μ l. After digestion and purification, the samples were loaded into an ABI 3130 sequencer and analyzed with PeakScanner software v1.0 (Applied Biosystems, USA).

Bioinformatics

The nucleotide sequences homologous to *nasN* were obtained from the NCBI nt nucleotide database and the RefSeq database

(Bacteria and Archaea) by tBLASTn analysis (e-value < 0.01, max-target-sequence: 20K) using the complete sequence of *Msm nasN* as a query. The tBLASTn result was further filtered to exclude false-positive matching which was aligned with an inconsistent Pfam domain (69) compared with the domains of *Msm NasN*. Among these remaining sequences, three kinds of nucleotide sequences homologous to *nasN* (the CA, DF or intact *nasN* homologs) were assigned to each sequence based on annotated domains. To reveal the evolution pattern of *nasN*, these nucleotide sequences were deduplicated and then aligned with MAFFT v7 (70). The generated MSA (multiple sequence alignment) was used to construct a phylogenetic tree by FastTree v2 (71). The unique taxonomic unit (taxonomy ID) of the genome containing a homologous sequence was retrieved and summarized into the taxonomic unit level. The NCBI taxonomy tree was constructed based on the whole taxonomy ID retrieved using ETE 3 (72) and the taxdump database of NCBI and visualized with GraPhlAn (73).

The complete 16S rRNA gene sequences of 153 genome-sequenced mycobacterial species listed by the LPSN database (<http://www.bacterio.net/index.html>) were collected from the NCBI database and trimmed to approximately 1400 nucleotides corresponding to the position 53-1460 of *E. coli* 16S rRNA gene sequence as previously described (58). Then the trimmed sequences were aligned using MAFFT v7 (70) and used to construct the phylogenetic tree by RAxML v8 (74). The locus_tags of the used 16S rRNA gene sequences were summarized in **Table S5**.

Statistical analysis

All data were expressed as means \pm SD where appropriate. Student's *t* test (unpaired, two-tailed) and one-way ANOVA with Tukey corrections were performed for statistical analysis by using the GraphPad Prism 5.0 program.

Acknowledgments. We thank Yuan-Yuan Liu (Tolo Biotech.), Yong-Xing Lei and Ni-Sha He for technical help, and Heather Smith Xie for improving the English of this manuscript.

Conflict of interest: The authors declare no competing financial interests.

Author contributions. Y. X., X.-Y. Z., Z.-H. S., L.-D. L., and G.-P. Z. designed the research; W. T., T.-H. L., Y. Y., Y.-C. W. performed the research; W. T., T.-H. L., J. W., H.-K. Z., Y. X., X.-Y. Z., Z.-H. S., L.-D. L., and G.-P. Z. analyzed the data; W. T., T.-H. L., Y. X., X.-Y. Z., L.-D. L., and G.-P. Z. wrote and revised the paper.

References

1. Kuypers, M. M. M., Marchant, H. K., and Kartal, B. (2018) The microbial nitrogen-cycling network. *Nat. Rev. Microbiol.* **16**, 263-276
2. Morozkina, E. V., and Zvyagilskaya, R. A. (2007) Nitrate reductases: structure, functions, and effect of stress factors. *Biochemistry (Mosc.)* **72**, 1151-1160
3. Maia, L. B., and Moura, J. J. (2014) How biology handles nitrite. *Chem. Rev.* **114**, 5273-5357
4. Moreno-Vivian, C., Cabello, P., Martinez-Luque, M., Blasco, R., and Castillo, F. (1999) Prokaryotic nitrate reduction: Molecular properties and functional distinction among bacterial nitrate reductases. *J. Bacteriol.* **181**, 6573-6584
5. Richardson, D. J., Berks, B. C., Russell, D. A., Spiro, S., and Taylor, C. J. (2001) Functional, biochemical and genetic diversity of prokaryotic nitrate reductases. *Cell. Mol. Life Sci.* **58**, 165-178
6. Gangeswaran, R., and Eady, R. R. (1996) Flavodoxin 1 of *Azotobacter vinelandii*: characterization and role in electron donation to purified assimilatory nitrate reductase. *Biochem. J.* **317**, 103-108
7. Flores, E., Frias, J. E., Rubio, L. M., and Herrero, A. (2005) Photosynthetic nitrate assimilation in cyanobacteria. *Photosynth. Res.* **83**, 117-133
8. Srivastava, A. P., Hardy, E. P., Allen, J. P., Vaccaro, B. J., Johnson, M. K., and Knaff, D. B. (2017) Identification of the ferredoxin-binding site of a ferredoxin-dependent cyanobacterial nitrate reductase. *Biochemistry* **56**, 5582-5592
9. Shao, Z., Gao, J., Ding, X., Wang, J., Chiao, J., and Zhao, G. (2011) Identification and functional analysis of a nitrate assimilation operon *nasACKBDEF* from *Amycolatopsis mediterranei* U32. *Arch. Microbiol.* **193**, 463-477
10. Lin, J. T., Goldman, B. S., and Stewart, V. (1994) The *nasFEDCBA* operon for nitrate and nitrite assimilation in *Klebsiella pneumoniae* M5al. *J. Bacteriol.* **176**, 2551-2559
11. Ogawa, K., Akagawa, E., Yamane, K., Sun, Z. W., LaCelle, M., Zuber, P., and Nakano, M. M. (1995) The *nasB* operon and *nasA* gene are required for nitrate/nitrite assimilation in *Bacillus subtilis*. *J. Bacteriol.* **177**, 1409-1413
12. Chu, S., Zhang, D., Wang, D., Zhi, Y., and Zhou, P. (2017) Heterologous expression and biochemical characterization of assimilatory nitrate and nitrite reductase reveals adaption and potential of *Bacillus megaterium* NCT-2 in secondary salinization soil. *Int. J. Biol. Macromol.* **101**, 1019-1028
13. Gates, A. J., Luque-Almagro, V. M., Goddard, A. D., Ferguson, S. J., Roldan, M. D., and Richardson, D. J. (2011) A composite biochemical system for bacterial nitrate and nitrite assimilation as exemplified by *Paracoccus denitrificans*. *Biochem. J.* **435**, 743-753
14. Hedgecock, L. W., and Costello, R. L. (1962) Utilization of nitrate by pathogenic and saprophytic mycobacteria. *J. Bacteriol.* **84**, 195-205
15. Gupta, R. S., Lo, B., and Son, J. (2018) Phylogenomics and comparative genomic studies robustly support division of the genus *Mycobacterium* into an emended genus *Mycobacterium* and four novel genera. *Front. Microbiol.* **9**, 67
16. Malm, S., Tiffert, Y., Micklinghoff, J., Schultze, S., Joost, I., Weber, I., Horst, S., Ackermann, B., Schmidt, M., Wohlleben, W., Ehlers, S., Geffers, R., Reuther, J., and Bange, F. C. (2009) The roles of the nitrate reductase NarGHJI, the nitrite reductase NirBD and the response regulator GlnR in nitrate assimilation of *Mycobacterium tuberculosis*. *Microbiology* **155**, 1332-1339
17. Khan, A., and Sarkar, D. (2006) Identification of a respiratory-type nitrate reductase and its role for survival of *Mycobacterium smegmatis* in Wayne model. *Microb. Pathog.* **41**, 90-95
18. Narrandes, N. C. (2013) *Functional Characterisation of Molybdopterin Synthase-encoding Genes in Mycobacteria*. Master's thesis, University of the Witwatersrand, Johannesburg, South Africa

19. Jeßberger, N. (2011) *The GlnR-dependent Nitrogen Regulatory Network of Mycobacterium smegmatis*. Ph.D. thesis, Friedrich-Alexander University Erlangen-Nuremberg, Erlangen, Germany
20. Gray, T. A., Palumbo, M. J., and Derbyshire, K. M. (2013) Draft genome sequence of MKD8, a conjugal recipient *Mycobacterium smegmatis* strain. *Genome Announc.* **1**, e0014813
21. Wang, J., Wang, Y., and Zhao, G. P. (2015) Precise characterization of GlnR Box in actinomycetes. *Biochem. Biophys. Res. Commun.* **458**, 605-607
22. Amon, J., Titgemeyer, F., and Burkovski, A. (2010) Common patterns - unique features: nitrogen metabolism and regulation in Gram-positive bacteria. *FEMS Microbiol. Rev.* **34**, 588-605
23. Jessberger, N., Lu, Y., Amon, J., Titgemeyer, F., Sonnewald, S., Reid, S., and Burkovski, A. (2013) Nitrogen starvation-induced transcriptome alterations and influence of transcription regulator mutants in *Mycobacterium smegmatis*. *BMC Res. Notes* **6**, 482
24. Jenkins, V. A., Barton, G. R., Robertson, B. D., and Williams, K. J. (2013) Genome wide analysis of the complete GlnR nitrogen-response regulon in *Mycobacterium smegmatis*. *BMC Genomics* **14**, 301
25. Eddy, S. R. (2011) Accelerated profile HMM searches. *PLoS Comput. Biol.* **7**, e1002195
26. Aigrain, L., Fatemi, F., Frances, O., Lescop, E., and Truan, G. (2012) Dynamic control of electron transfers in diflavin reductases. *Int. J. Mol. Sci.* **13**, 15012-15041
27. Murataliev, M. B., Feyereisen, R., and Walker, F. A. (2004) Electron transfer by diflavin reductases. *Biochim. Biophys. Acta* **1698**, 1-26
28. Netz, D. J., Stumpfig, M., Dore, C., Muhlenhoff, U., Pierik, A. J., and Lill, R. (2010) Tah18 transfers electrons to Dre2 in cytosolic iron-sulfur protein biogenesis. *Nat. Chem. Biol.* **6**, 758-765
29. Olteanu, H., and Banerjee, R. (2003) Redundancy in the pathway for redox regulation of mammalian methionine synthase: reductive activation by the dual flavoprotein, novel reductase 1. *J. Biol. Chem.* **278**, 38310-38314
30. Willistein, M., Bechtel, D. F., Muller, C. S., Demmer, U., Heimann, L., Kayastha, K., Schunemann, V., Pierik, A. J., Ullmann, G. M., Ermler, U., and Boll, M. (2019) Low potential enzymatic hydride transfer via highly cooperative and inversely functionalized flavin cofactors. *Nat. Commun.* **10**, 2074
31. Kawada, Y., Goshima, T., Sawamura, R., Yokoyama, S. I., Yanase, E., Niwa, T., Ebihara, A., Inagaki, M., Yamaguchi, K., Kuwata, K., Kato, Y., Sakurada, O., and Suzuki, T. (2018) Daidzein reductase of *Eggerthella* sp. YY7918, its octameric subunit structure containing FMN/FAD/4Fe-4S, and its enantioselective production of R-dihydroisoflavones. *J. Biosci. Bioeng.* **126**, 301-309
32. McGoldrick, H. M., Roessner, C. A., Raux, E., Lawrence, A. D., McLean, K. J., Munro, A. W., Santabarbara, S., Rigby, S. E., Heathcote, P., Scott, A. I., and Warren, M. J. (2005) Identification and characterization of a novel vitamin B12 (cobalamin) biosynthetic enzyme (CobZ) from *Rhodobacter capsulatus*, containing flavin, heme, and Fe-S cofactors. *J. Biol. Chem.* **280**, 1086-1094
33. Simtchouk, S., Eng, J. L., Meints, C. E., Makins, C., and Wolthers, K. R. (2013) Kinetic analysis of cytochrome P450 reductase from *Artemisia annua* reveals accelerated rates of NADH-dependent flavin reduction. *FEBS J.* **280**, 6627-6642
34. Anderson, L. J., Richardson, D. J., and Butt, J. N. (2001) Catalytic protein film voltammetry from a respiratory nitrate reductase provides evidence for complex electrochemical modulation of enzyme activity. *Biochemistry* **40**, 11294-11307
35. Gates, A. J., Richardson, D. J., and Butt, J. N. (2008) Voltammetric characterization of the aerobic energy-dissipating nitrate reductase of *Paracoccus pantotrophus*: exploring the activity of a redox-balancing enzyme as a function of electrochemical potential. *Biochem. J.* **409**, 159-168

36. Srivastava, A. P., Hirasawa, M., Bhalla, M., Chung, J. S., Allen, J. P., Johnson, M. K., Tripathy, J. N., Rubio, L. M., Vaccaro, B., Subramanian, S., Flores, E., Zabet-Moghaddam, M., Stittle, K., and Knaff, D. B. (2013) Roles of four conserved basic amino acids in a ferredoxin-dependent cyanobacterial nitrate reductase. *Biochemistry* **52**, 4343-4353
37. Crack, J. C., Munnoch, J., Dodd, E. L., Knowles, F., Al Bassam, M. M., Kamali, S., Holland, A. A., Cramer, S. P., Hamilton, C. J., Johnson, M. K., Thomson, A. J., Hutchings, M. I., and Le Brun, N. E. (2015) NsrR from *Streptomyces coelicolor* is a nitric oxide-sensing [4Fe-4S] cluster protein with a specialized regulatory function. *J. Biol. Chem.* **290**, 12689-12704
38. Khoroshilova, N., Popescu, C., Munck, E., Beinert, H., and Kiley, P. J. (1997) Iron-sulfur cluster disassembly in the FNR protein of *Escherichia coli* by O₂: [4Fe-4S] to [2Fe-2S] conversion with loss of biological activity. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 6087-6092
39. Koppel, N., Bisanz, J. E., Pandelia, M. E., Turnbaugh, P. J., and Balskus, E. P. (2018) Discovery and characterization of a prevalent human gut bacterial enzyme sufficient for the inactivation of a family of plant toxins. *Elife* **7**, e33953
40. Lessner, F. H., Jennings, M. E., Hirata, A., Duin, E. C., and Lessner, D. J. (2012) Subunit D of RNA polymerase from *Methanosarcina acetivorans* contains two oxygen-labile [4Fe-4S] clusters: implications for oxidant-dependent regulation of transcription. *J. Biol. Chem.* **287**, 18510-18523
41. Iino, T., Miyauchi, K., Kasai, D., Masai, E., and Fukuda, M. (2013) Characterization of nitrate and nitrite utilization system in *Rhodococcus jostii* RHA1. *J. Biosci. Bioeng.* **115**, 600-606
42. Fournier, D., Hawari, J., Halasz, A., Streger, S. H., McClay, K. R., Masuda, H., and Hatzinger, P. B. (2009) Aerobic biodegradation of N-nitrosodimethylamine by the propanotroph *Rhodococcus ruber* ENV425. *Appl. Environ. Microbiol.* **75**, 5088-5093
43. Myers, M. R., and King, G. M. (2016) Isolation and characterization of *Acidobacterium ailaui* sp. nov., a novel member of Acidobacteria subdivision 1, from a geothermally heated Hawaiian microbial mat. *Int. J. Syst. Evol. Microbiol.* **66**, 5328-5335
44. Echorst, S. A., Trojan, D., Roux, S., Herbold, C., Rattei, T., and Woebken, D. (2018) Genomic insights into the Acidobacteria reveal strategies for their success in terrestrial environments. *Environ. Microbiol.* **20**, 1041-1063
45. Mannisto, M. K., Rawat, S., Starovoytov, V., and Haggblom, M. M. (2012) *Granulicella arctica* sp. nov., *Granulicella mallensis* sp. nov., *Granulicella tundricola* sp. nov. and *Granulicella sapmiensis* sp. nov., novel acidobacteria from tundra soil. *Int. J. Syst. Evol. Microbiol.* **62**, 2097-2106
46. Urlacher, V. B., and Girhard, M. (2019) Cytochrome P450 monooxygenases in biotechnology and synthetic biology. *Trends Biotechnol.* **37**, 882-897
47. Sahraoui, N., Ballif, M., Zelleg, S., Yousfi, N., Ritter, C., Friedel, U., Amstutz, B., Yala, D., Boulahbal, F., Guetarni, D., Zinsstag, J., and Keller, P. M. (2011) *Mycobacterium algericum* sp. nov., a novel rapidly growing species related to the *Mycobacterium terrae* complex and associated with goat lung lesions. *Int. J. Syst. Evol. Microbiol.* **61**, 1870-1874
48. Suy, F., Carricajo, A., Grattard, F., Cazorla, C., Denis, C., Girardin, P., Lucht, F., and Botelho-Nevers, E. (2013) Infection due to *Mycobacterium thermoresistibile*: a case associated with an orthopedic device. *J. Clin. Microbiol.* **51**, 3154-3156
49. Varghese, B., Enani, M., Shoukri, M., AlThawadi, S., AlJohani, S., and Al-Hajoj, S. (2017) Emergence of rare species of nontuberculous mycobacteria as potential pathogens in Saudi Arabian clinical setting. *PLoS Negl. Trop. Dis.* **11**, e0005288
50. Mahillon, J., and Chandler, M. (1998) Insertion sequences. *Microbiol. Mol. Biol. Rev.* **62**, 725
51. Schulz, S., Schumacher, D., Raszkowski, D., Girhard, M., and Urlacher, V. B. (2016) Fusion to hydrophobin HFBI improves the catalytic performance of a cytochrome P450 system. *Front. Bioeng. Biotechnol.* **4**, 57
52. Hajjar, C., Cherrier, M. V., Mirandela, G. D., Petit-Hartlein, I., Stasia, M. J., Fontecilla-Camps, J. C., Fieschi, F., and Dupuy, J. (2017) The NOX family of proteins is also present in

- bacteria. *Mbio* **8**, e01487-17
53. Lombard, J., Lopez-Garcia, P., and Moreira, D. (2012) The early evolution of lipid membranes and the three domains of life. *Nat. Rev. Microbiol.* **10**, 507-515
54. Zhi, X. Y., Yao, J. C., Tang, S. K., Huang, Y., Li, H. W., and Li, W. J. (2014) The fthalosine pathway played an important role in menaquinone biosynthesis during early prokaryote evolution. *Genome Biol. Evol.* **6**, 149-160
55. Novoa, E. M., Pavon-Eternod, M., Pan, T., and Ribas de Pouplana, L. (2012) A role for tRNA modifications in genome structure and codon usage. *Cell* **149**, 202-213
56. Devulder, G., de Montclos, M. P., and Flandrois, J. P. (2005) A multigene approach to phylogenetic analysis using the genus *Mycobacterium* as a model. *Int. J. Syst. Evol. Microbiol.* **55**, 293-302
57. Mignard, S., and Flandrois, J. P. (2008) A seven-gene, multilocus, genus-wide approach to the phylogeny of mycobacteria using supertrees. *Int. J. Syst. Evol. Microbiol.* **58**, 1432-1441
58. Tortoli, E. (2012) Phylogeny of the genus *Mycobacterium*: many doubts, few certainties. *Infect. Genet. Evol.* **12**, 827-831
59. Case, E. D., and Samuel, J. E. (2016) Contrasting lifestyles within the host cell. *Microbiol. Spectr.* **4**
60. Gouzy, A., Poquet, Y., and Neyrolles, O. (2014) Nitrogen metabolism in *Mycobacterium tuberculosis* physiology and virulence. *Nat. Rev. Microbiol.* **12**, 729-737
61. Weber, I., Fritz, C., Ruttkowski, S., Kreft, A., and Bange, F. C. (2000) Anaerobic nitrate reductase (*narGHJT*) activity of *Mycobacterium bovis* BCG in vitro and its contribution to virulence in immunodeficient mice. *Mol. Microbiol.* **35**, 1017-1025
62. Stermann, M., Sedlacek, L., Maass, S., and Bange, F. C. (2004) A promoter mutation causes differential nitrate reductase activity of *Mycobacterium tuberculosis* and *Mycobacterium bovis*. *J. Bacteriol.* **186**, 2856-2861
63. Khan, A., Akhtar, S., Ahmad, J. N., and Sarkar, D. (2008) Presence of a functional nitrate assimilation pathway in *Mycobacterium smegmatis*. *Microb. Pathog.* **44**, 71-77
64. Bardarov, S., Bardarov Jr, S., Jr., Pavelka Jr, M. S., Jr., Sambandamurthy, V., Larsen, M., Tufariello, J., Chan, J., Hatfull, G., and Jacobs Jr, W. R., Jr. (2002) Specialized transduction: an efficient method for generating marked and unmarked targeted gene disruptions in *Mycobacterium tuberculosis*, *M. bovis* BCG and *M. smegmatis*. *Microbiology* **148**, 3007-3017
65. Chen, M., Asai, S. I., Narai, S., Nambu, S., Omura, N., Sakaguchi, Y., Suzuki, T., Ikeda-Saito, M., Watanabe, K., Yao, M., Shigi, N., and Tanaka, Y. (2017) Biochemical and structural characterization of oxygen-sensitive 2-thiouridine synthesis catalyzed by an iron-sulfur protein TtuA. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 4954-4959
66. Pierik, A. J., Wolbert, R. B., Mutsaers, P. H., Hagen, W. R., and Veeger, C. (1992) Purification and biochemical characterization of a putative [6Fe-6S] prismane-cluster-containing protein from *Desulfovibrio vulgaris* (Hildenborough). *Eur. J. Biochem.* **206**, 697-704
67. Lin, W., Wang, Y., Han, X., Zhang, Z., Wang, C., Wang, J., Yang, H., Lu, Y., Jiang, W., Zhao, G. P., and Zhang, P. (2014) Atypical OmpR/PhoB subfamily response regulator GlnR of actinomycetes functions as a homodimer, stabilized by the unphosphorylated conserved Asp-focused charge interactions. *J. Biol. Chem.* **289**, 15413-15425
68. Wang, Y., Cen, X. F., Zhao, G. P., and Wang, J. (2012) Characterization of a new GlnR binding box in the promoter of *amtB* in *Streptomyces coelicolor* inferred a PhoP/GlnR competitive binding mechanism for transcriptional regulation of *amtB*. *J. Bacteriol.* **194**, 5237-5244
69. Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E. L., Studholme, D. J., Yeats, C., and Eddy, S. R. (2004) The Pfam protein families database. *Nucleic Acids Res.* **32**, D138-141
70. Katoh, K., and Standley, D. M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772-780

71. Price, M. N., Dehal, P. S., and Arkin, A. P. (2010) FastTree 2-approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, e9490
72. Huerta-Cepas, J., Serra, F., and Bork, P. (2016) ETE 3: reconstruction, analysis, and visualization of phylogenomic data. *Mol. Biol. Evol.* **33**, 1635-1638
73. Asnicar, F., Weingart, G., Tickle, T. L., Huttenhower, C., and Segata, N. (2015) Compact graphical representation of phylogenetic data and metadata with GraPhlAn. *PeerJ* **3**, e1029
74. Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312-1313

FOOTNOTES

This research was supported by the National Natural Science Foundation of China (31430004, 31830002, 81671988, 91751000, 91951000 to G.-P. Z., 91856112 to Y. X., 31970032, 81991532 to L.-D. L, and 31800123), the Research Unit Fund of Li Ka Shing Institute of Health Sciences (7103506 to G.-P. Z.), the Hong Kong Health and Medical Research Fund (12110622 to G.-P. Z.), the National Key R&D Program of China (2019YFA0905100, 2018YFD0500900, 2017YFC1201200), the National Science and Technology Major Program of China (2018ZX10302301 to L.-D. L.), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB19040200), the Key Research Program of the Chinese Academy of Sciences (KFZD-SW-219-5), and the International Partnership Program of the Chinese Academy of Sciences (153D31KYSB20170121).

The abbreviations used are: *Msm*, *Mycobacterium smegmatis*; *Mtb*, *Mycobacterium tuberculosis*; *Ame*, *Amycolatopsis mediterranei*; NAS, assimilatory nitrate reductase; NaR, nitrate reductase; CA domain, the nitrate reduction catalytic domain; DF domain, the FMN-FAD/NAD domain; Fv-NarB, flavodoxin-dependent heterodimeric NarB; Fd-NarB, ferredoxin-dependent heterodimeric NarB; MoCo, molybdopterin cofactor; N-NasN, N-terminal region (1-734 aa) of NasN; EMSA, electrophoretic mobility shift assay; MPLN, a nitrogen-free modified *M. phlei* medium; MV, methyl viologen.

Table 1. Electron donor-dependent enzymatic kinetic parameters of NasN and NasAC employing different electron donors.

Electron donor ^a	Kinetic parameters ^b	Enzyme		
		NasN	NasAC	NasA
NADPH	<i>K_m</i> (nitrate)	11.9 ± 1.2	21.2 ± 5.6	ND
	<i>V_{max}</i>	845 ± 13	2.20 ± 0.10	ND
	<i>k_{cat}</i>	2100 ± 30	4.70 ± 0.40	ND
	<i>k_{cat}/K_m</i>	176	0.222	ND
NADH	<i>K_m</i> (nitrate)	18.4 ± 5.1	12.7 ± 3.9	ND
	<i>V_{max}</i>	32.0 ± 1.5	17.0 ± 3.5	ND
	<i>k_{cat}</i>	80.2 ± 17.7	39.6 ± 1.2	ND
	<i>k_{cat}/K_m</i>	4.36	3.12	ND
MV	<i>K_m</i> (nitrate)	51.5 ± 9.5	14.2 ± 3.0	15.1 ± 2.4
	<i>V_{max}</i>	5.90 ± 0.90	590 ± 84	572 ± 60
	<i>k_{cat}</i>	15.1 ± 2.2	1400 ± 98	1390 ± 114
	<i>k_{cat}/K_m</i>	0.293	98.6	92.0

^aConcentrations of each electron donor, i.e., 400 μM NADPH or NADH, or 150 μM MV; ^b*K_m* (μM), *V_{max}* (nmol/min/mg), *k_{cat}* (min⁻¹), *k_{cat}/K_m* (min⁻¹μM⁻¹)

All enzyme samples used for this study were cytoplasmic fractions of the crude cell extracts anaerobically prepared from the mc²155ΔnasN::His-nasN_{Msm}^c, mc²155ΔnasN::His-nasAC_{Ame}^c and mc²155ΔnasN::His-nasA_{Ame}^c strains, respectively. The strains were grown aerobically at 37 °C in MPLN minimal medium containing 10 mM NaNO₃, except that 5 mM L-glutamine was used for the cultivation of the mc²155ΔnasN::His-nasA_{Ame}^c strain. Anti-His tag monoclonal antibody-based ELISA method was used to determine the concentrations of His-tagged proteins in enzyme samples. Reactions were conducted as described in Experimental Procedures. Data are expressed as the mean ± SD of three independent experiments. ND, not determined.

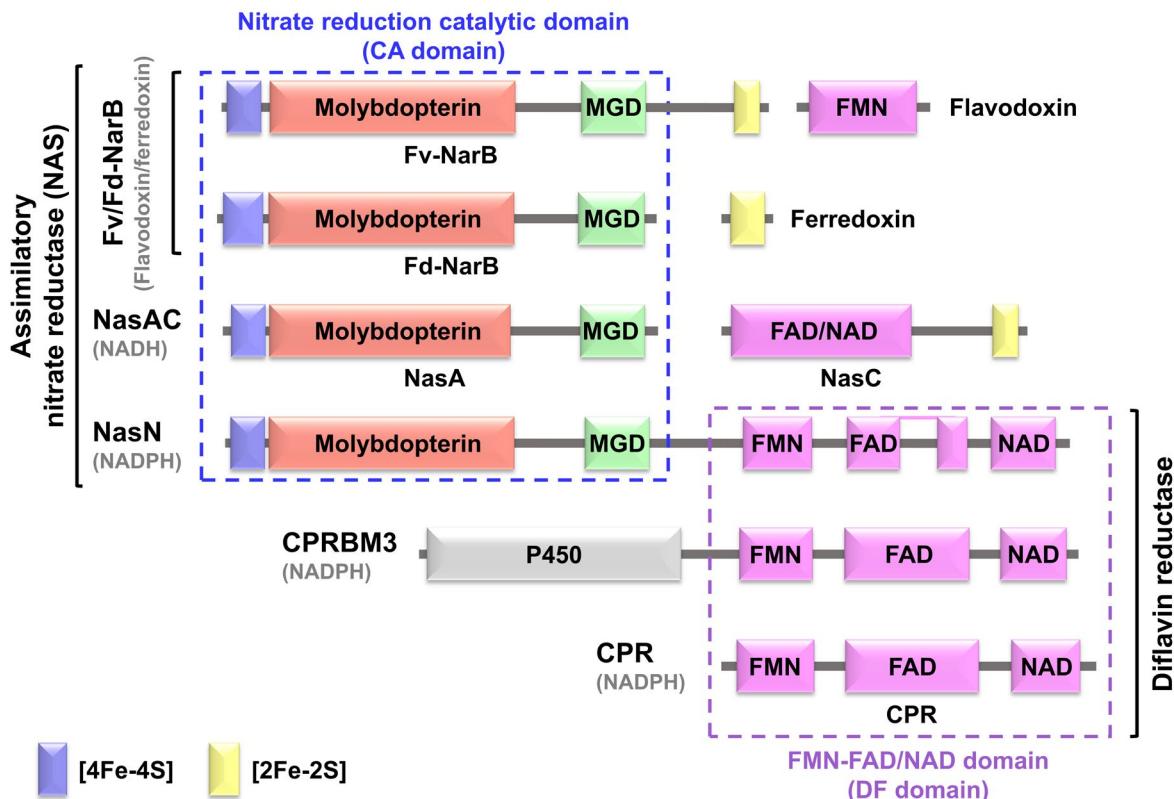


Figure 1. Domain architecture of various NAS enzymes and diflavin reductases. Schematic representation of the structural domains present in various recognized NAS enzymes and diflavin reductases from UniProt database: flavodoxin (Fv)- or ferredoxin (Fd)-dependent Fv/Fd-NarB from *Azotobacter vinelandii* (C1DHC8 and P52964) and *Synechococcus elongatus* (P39458 and P0A3D2); NADH-dependent heterodimeric NasAC from *Ame* (A0A0H3D089 and A0A0H3CWC1); *Msm* NasN (*Msmeg_4206*, this study); *Bacillus megaterium* CPRBM3, cytochrome P450 BM3 (P14779); *Mus musculus* CPR, cytochrome P450 reductase (P37040). The annotated binding domains are depicted with different colors: blue, [4Fe-4S] cluster; yellow, [2Fe-2S] cluster; red, molybdopterin; green, bis-molybdopterin guanine dinucleotide (MGD); pink, two flavins (FMN and FAD) and NAD; grey, cytochrome P450 (P450).

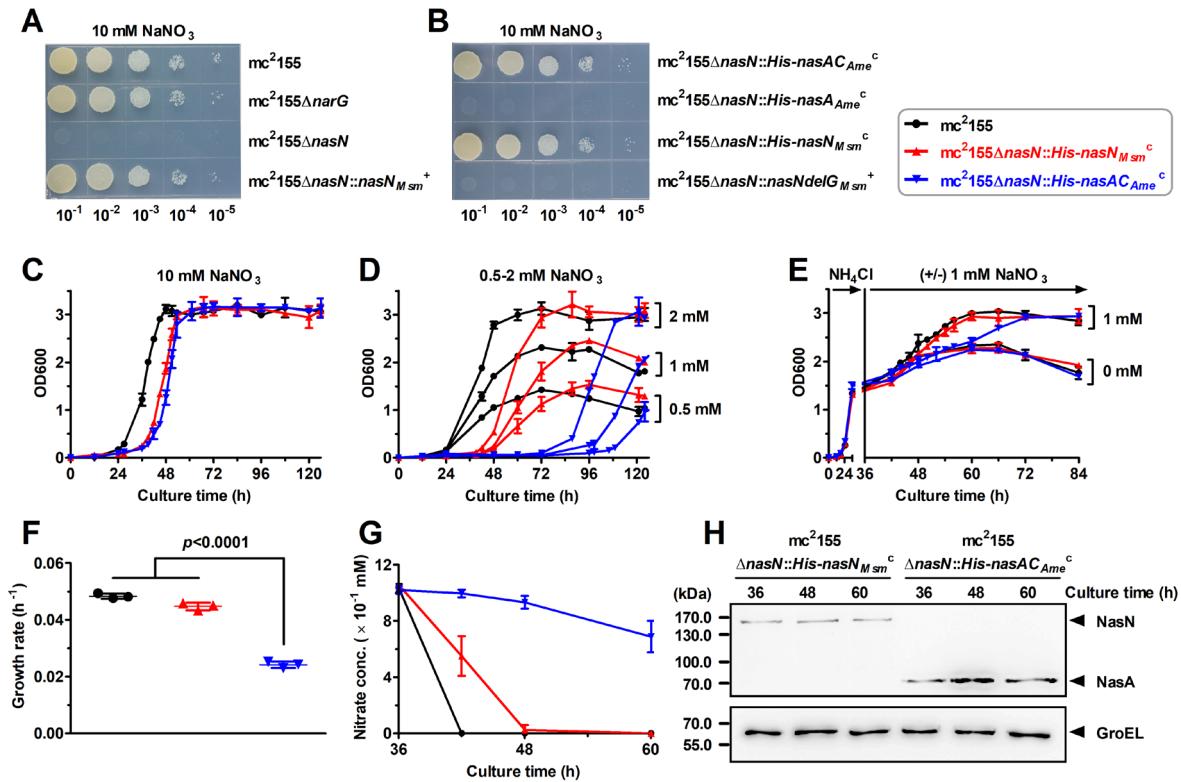


Figure 2. Genetic and physiological characterization of *Msmeg_4206* (*nasN*). (A-D) Growth phenotype of *Msm* strains on nitrate as the sole nitrogen source. *Msm* strains were cultured aerobically at 37 °C in the solid (A, B) or liquid (C, D) MPLN minimal medium containing 0.5-10 mM NaNO₃ as the sole nitrogen source. The plates for (A, B) were incubated for four days. OD₆₀₀, optical density at 600 nm. (E-H) *Msm* strains were cultured aerobically at 37 °C for 36 h in MPLN medium containing NH₄Cl (1 mM), then the culture was divided into two equal fractions and one was added with 1 mM NaNO₃ for further incubation at 37 °C. Growth curves (E), cell growth rates (F) and extracellular nitrate concentrations (G) were measured after nitrate addition. (H) Western blot analysis of total cell lysates of *Msm* strains expressing His-tagged *Msm* NasN or *Ame* NasA proteins at different times as described in (E). GroEL was used as a loading control. Data (C-G) are expressed as the mean ± SD of three biological and two technical replicates. Statistical analysis was performed using one-way ANOVA with Tukey corrections (F).

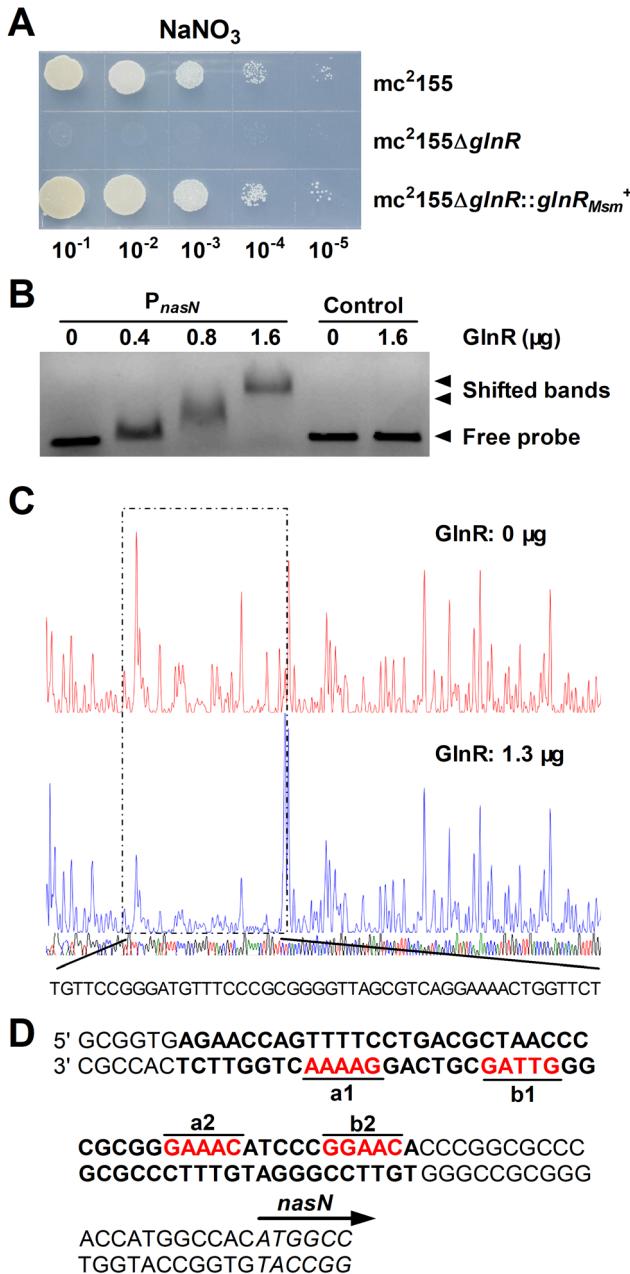


Figure 3. GlnR binds to the *nasN* promoter. (A) The growth defect of *Msm* *glnR*-null mutant on nitrate as the sole nitrogen source. Aliquots from serially diluted cultures were spotted onto MPLN solid medium containing 10 mM NaNO₃ and incubated for 4 days at 37 °C under aerobic conditions. (B) EMSA analysis of His-tagged GlnR protein with FAM-labeled DNA fragments of the *nasN* promoter region (P_{nasN}). A *nasN* CDS fragment was used as a negative control. Excess sheared salmon sperm DNA was added to the assay system to prevent nonspecific binding. (C) DNase I footprinting analysis of His-tagged GlnR binding to the *nasN* promoter region. The protected DNA sequences are shown at the bottom. (D) The predicted GlnR-binding sites (shown in red) in the *nasN* promoter region. GlnR protected DNA sequences are in bold.

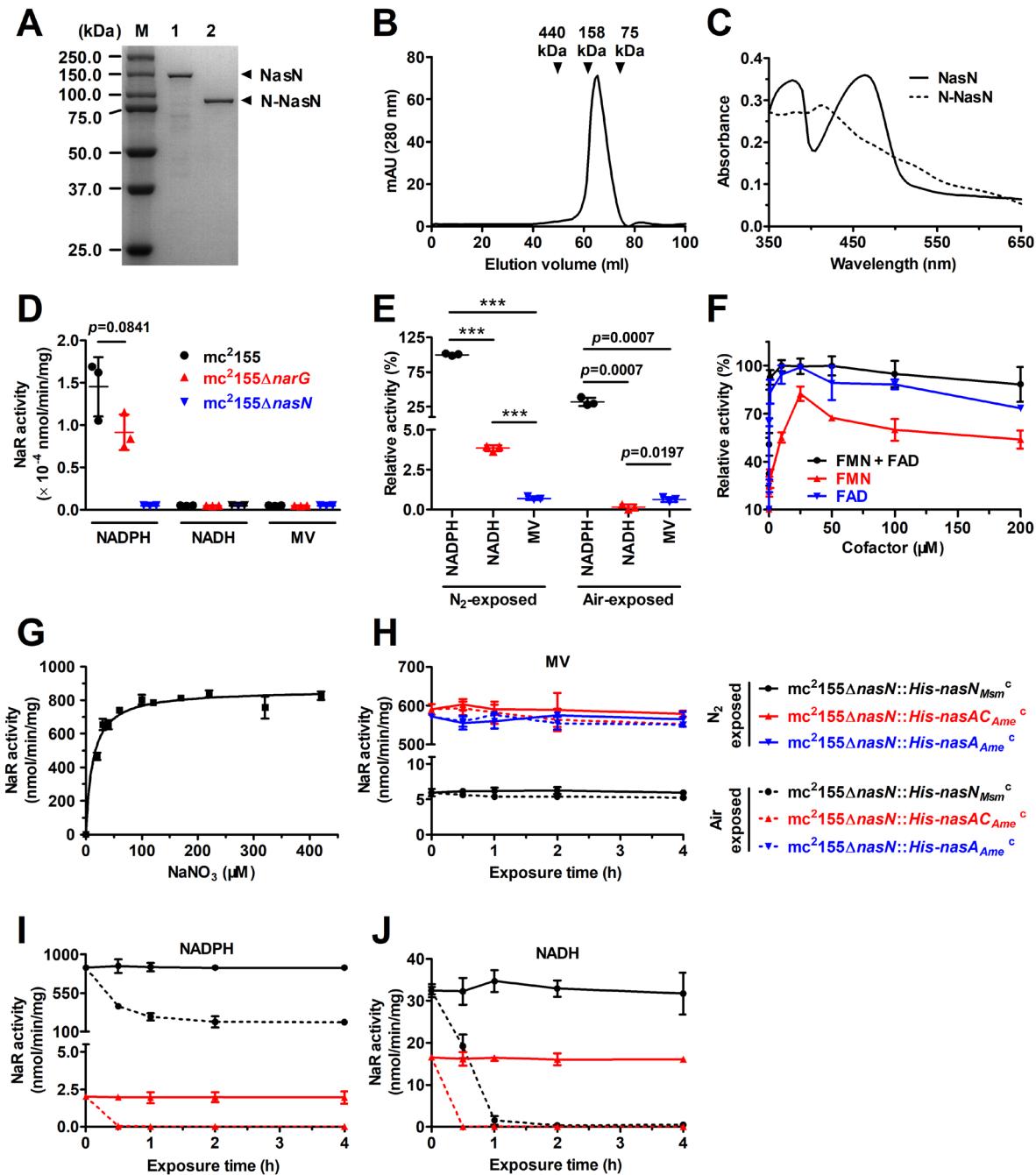


Figure 4. Biochemical and enzymatic properties of NasN. (A) The SDS-PAGE analysis of NasN and the N-terminal region (1-734 aa) of NasN (N-NasN). Lane M, molecular mass markers; lane 1, NasN sample obtained from Ni^{2+} -NTA affinity chromatography; lane 2, N-NasN sample obtained from size exclusion chromatography. (B) Gel filtration analysis of the purified NasN. The elution positions of Conalbumin (75 kDa), Aldolase (158 kDa) and Ferritin (440 kDa) standards are indicated by the arrows. (C) UV-visible absorption spectra of the reconstituted N-NasN, in comparison to the reconstituted holoprotein. (D) The specific NaR activities in cytoplasmic fractions of the crude cell extracts prepared anaerobically from *Msm* wild-type strain, $mc^2155\Delta narG$ and $mc^2155\Delta nasN$ strains.

The strains were grown aerobically at 37 °C in MPLN medium containing 10 mM NaNO₃, except that 5 mM L-glutamine was used for the cultivation of *Msm ΔnasN*. Reactions were conducted with electron donors of 400 μM NADPH or NADH, or 150 μM reduced methyl viologen (MV) at pH 7.5, 30 °C under anaerobic assay conditions as described in Experimental Procedures. (E) NasN activities employing different electron donors with or without pre-incubation in the air. For the air-exposed samples, the purified NasN was pre-incubated in the air for 1 h at 4 °C. The activities were expressed as relative to the highest activity (100%). (F) Effect of flavin cofactors upon NasN activity. Reactions were conducted in the presence of 400 μM NADPH. The activities were expressed as relative to the highest activity (100%). (G) Michaelis-Menten kinetics of *Msm NasN*. Reactions were carried out under anaerobic assay conditions with 400 μM NADPH, 25 μM FMN/FAD, at 30 °C for 10 min, which ensured zero-order kinetics for all reactions. (H-J) Effect of oxygen upon the activities of NasN and NasAC. All enzyme samples used for this study were cytoplasmic fractions of the crude cell extracts anaerobically prepared from the complemented strains expressing His-tagged *Msm NasN* or *Ame NasA*. These samples were subject to pre-incubation in N₂ (solid lines) or air (dashed lines) at 4 °C for 0 to 4 h before measuring the activity under the anaerobic assay conditions using different electron donors. Data (D-J) are expressed as the mean ± SD of three biological and two technical replicates. Statistical analysis was performed using Student's *t* test and one-way ANOVA with Tukey corrections (D and E). ****p* < 0.0001 (E).

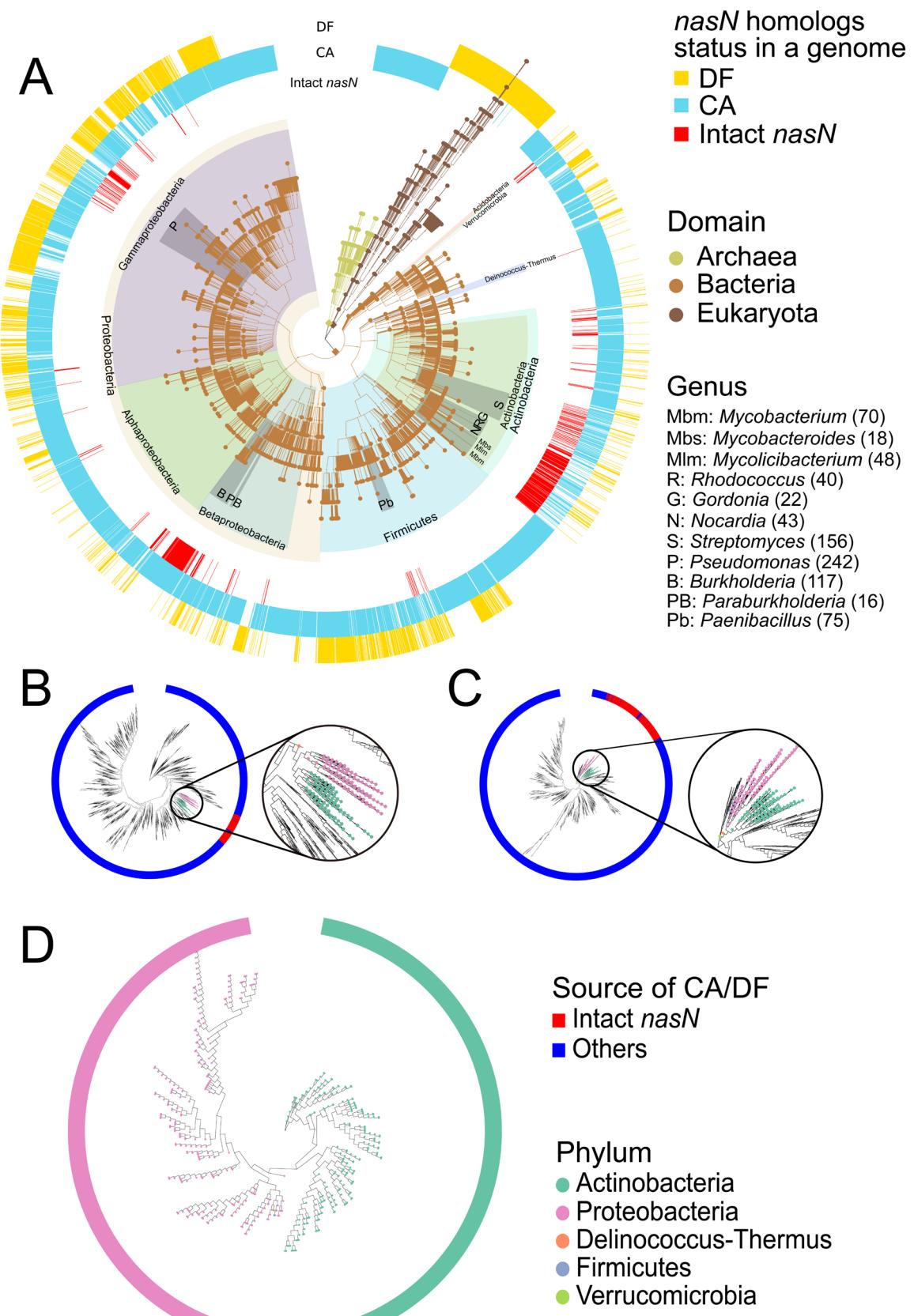


Figure 5. Profiling of nucleotide sequences homologous to *nasN* (the CA, DF or intact *nasN* homologs) in the three domains of life (A) and the phylogenetic trees of CA (B), DF (C) and intact *nasN* (D) homologs. (A) The NCBI taxonomy tree was constructed with 6307 species (unique NCBI taxonomy ID) belonging to the three domains of life and annotated at species level by colors of the circular ring outside the tree with the existing status of nucleotide sequences homologous to *nasN* in a genome: Intact *nasN*, homologs for the intact *nasN*; CA, homologs for the CA domain (N-terminal region of NasN); DF, homologs for the DF domain (C-terminal region of NasN). The grey-shaded region of the tree with labels is used to represent the genera with the total number of species containing the intact *nasN* homolog, except the genus whose number of *nasN*-containing species was less than 10. (B-C) The phylogenetic trees were constructed based on the nucleotide sequences of CA (9218) and DF (4230) homologs individually, which were extracted and deduplicated from genome sequences at NCBI. The original sequence sources including intact *nasN* or other homologs (CA or DF homologs) are illustrated with the colors of the circular ring outside the tree. The colorized endpoints of the branches also represent the sequences that originated from the intact *nasN* homologs. (D) The phylogenetic tree of the intact *nasN* homologs (549 deduplicated nucleotide sequences). The details of the used nucleotide sequences are listed in **Table S2**.

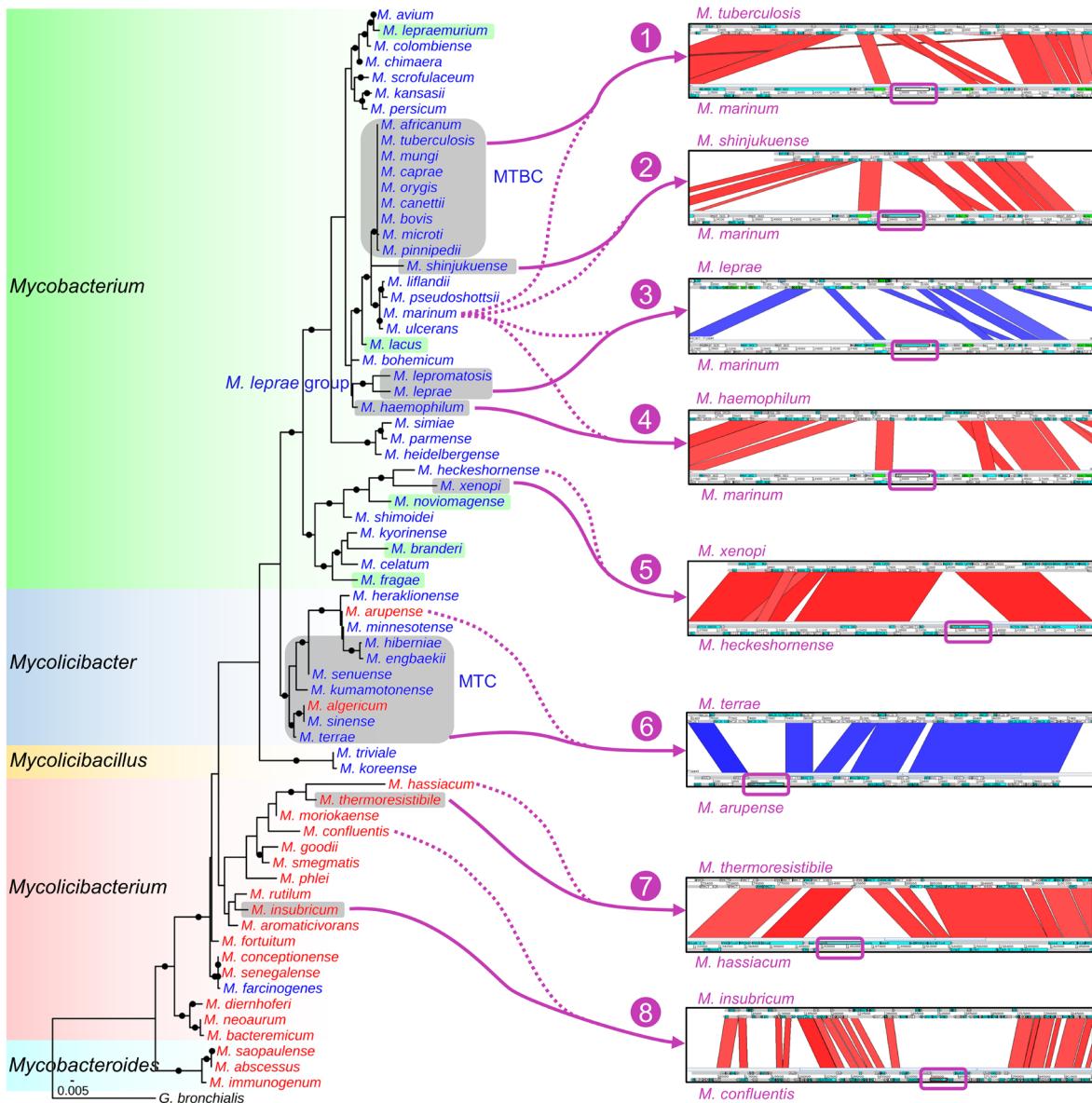


Figure 6. The loss of *nasN* in mycobacterial species during evolution. The maximum-likelihood phylogenetic tree was constructed based on the 16S rRNA gene sequences from 153 genome-sequenced mycobacterial species listed by the LPSN database, with the *Gordonia bronchialis* as an outgroup (Fig. S6). A pruned tree (middle panel) with 70 selected mycobacterial species is shown. All the *nasN*-lacking species have been listed, including the 23 confirmed species (shaded in grey), and 5 unconfirmed species (shaded in green) due to the possible dramatic genome reduction and rearrangement or their *nasN*-related segments locating at the ends of the sequenced contigs. The five newly proposed mycobacterial clades (15) are depicted in different colors (left panel). Three major clusters, including the *Mycobacterium tuberculosis* complex (MTBC), the *Mycolicibacter terrae* complex (MTC) and the *Mycobacterium leprae* group are shown. The fast-growing and slow-growing mycobacterial species are colored in red and blue, respectively. Bootstrap values (>50) are shown with closed circles. The scale bar indicates the genetic distance (5 substitutions per 1000 nucleotides).

The right panel shows the alignment of the *nasN*-related genomic region (~100 genes) between the selected *nasN*-lacking species (purple solid line) and their *nasN*-containing close relatives (purple dashed line) with the *nasN*-containing segment circled in the purple box. The red and blue color in the right panels shows the forward and reversed alignment in a genomic region.

**A recently evolved diflavin-containing monomeric nitrate reductase is responsible for
highly efficient bacterial nitrate assimilation**

Wei Tan, Tian-Hua Liao, Jin Wang, Yu Ye, Yu-Chen Wei, Hao-Kui Zhou, Youli Xiao,
Xiao-Yang Zhi, Zhi-Hui Shao, Liang-Dong Lyu and Guo-Ping Zhao

J. Biol. Chem. published online February 28, 2020

Access the most updated version of this article at doi: [10.1074/jbc.RA120.012859](https://doi.org/10.1074/jbc.RA120.012859)

Alerts:

- [When this article is cited](#)
- [When a correction for this article is posted](#)

[Click here](#) to choose from all of JBC's e-mail alerts