Diversity and evolution of nitric oxide reduction

Ranjani Murali[1,a*], Laura A. Pace[2], Robert A. Sanford[3], Lewis M. Ward[4], Mackenzie Lynes[5], Roland Hatzenpichler[5,6], Usha F. Lingappa[7], Woodward W. Fischer[8], Robert B. Gennis[9], James Hemp[2,8,a*]

[1]Division of Biology and Biological Engineering, California Institute of Technology, Pasedena, CA, USA

[2]Metrodora Institute, Salt Lake City, UT, USA

[3]Department of Geology, University of Illinois at Urbana-Champaign, Urbana, IL, USA

[4]Smith College, Northampton, MA, USA

[5]Department of Chemistry and Biochemistry, Thermal Biology Institute, and Center for Biofilm Enginering, Montana State University, MT, USA

[6]Department of Microbiology and Cell Biology, Montana State University, MT, USA

[7]Department of Plant and Microbial Biology, University of California, Berkeley, CA, USA

[8]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, USA

[9]Department of Biochemistry, University of Illinois, Urbana-Champaign, Urbana, USA

[a]James Hemp and Ranjani Murali contributed equally to the work in this manuscript.

[*] Corresponding authors: E-mail: james.hemp@metrodora.org, m.ranjani@gmail.com

Nitrogen is an essential element for life, with the availability of fixed nitrogen limiting productivity in many ecosystems. The return of oxidized nitrogen species to the atmospheric N_2 pool is predominately catalyzed by microbial denitrification ($NO_3^- \rightarrow NO_2^- \rightarrow NO \rightarrow N_2O \rightarrow N_2$)[1]. Incomplete denitrification can produce N_2O as a terminal product, leading to an increase in atmospheric N_2O, a potent greenhouse and ozone-depleting gas[2]. The production of N_2O is catalyzed by nitric oxide reductase (NOR) members of the heme-copper oxidoreductase (HCO) superfamily[3]. Here we use phylogenomics to identify a number of previously uncharacterized HCO families and propose that many of them (eNOR, sNOR, gNOR, and nNOR) perform nitric oxide reduction. These families have novel active-site structures and several have conserved proton channels, suggesting that they might be able to couple nitric oxide reduction to energy conservation. We isolated and biochemically characterized a member of the eNOR family from *Rhodothermus marinus*, verifying that it performs nitric oxide reduction both *in vitro* and *in vivo*. These newly identified NORs exhibit broad phylogenetic and environmental distributions, expanding the diversity of microbes that can perform denitrification. Phylogenetic analyses of the HCO superfamily demonstrate that nitric oxide reductases evolved multiple times independently from oxygen reductases, suggesting that complete denitrification evolved after aerobic respiration.

The HCO superfamily is extremely diverse, with members playing crucial roles in both aerobic (oxygen reductases) and anaerobic (nitric oxide reductases) respiration. The superfamily currently consists of at least three oxygen reductase families (A, B and C) and three NOR families (cNOR, qNOR, and qCu_ANOR)[4]. The oxygen reductases

catalyze the reduction of O_2 to water ($O_2 + 4e_{out}^- + 4H_{in}^+ + nH_{in}^+ \rightarrow 2H_2O + nH_{out}^+$) and share a conserved reaction mechanism[5,6], where three of the electrons required to reduce O_2 are provided by the active-site metals, heme-Fe and Cu_B, while the fourth electron is derived from a redox-active cross-linked tyrosine cofactor[7] (**Figure 1**). The free energy of the reaction is converted to a transmembrane proton electrochemical gradient by two different mechanisms, charge separation across the membrane and proton pumping[8–10]. Both the chemical and pumped protons are taken up from the electrochemically negative side of the membrane (bacterial cytoplasm) via conserved proton-conducting channels that are comprised of conserved polar residues and internal water molecules. The different oxygen reductase families exhibit differential proton pumping stochiometries (n=4 for the A-family, and n=2 for the B and C-families)[8–10], depending on the details of their conserved proton channels. The oxygen reductases also vary in their secondary subunits that function as redox relays from electron donors to the active-site, with the A and B-families utilizing a Cu_A-containing subunit[11–13] and the C-family containing one or more cytochrome *c* subunits[14] (**Figure 1**).

Nitric oxide reductases (NORs) catalyze the reduction of nitric oxide to nitrous oxide ($2NO + 2H_{out}^+ + 2e_{out}^- + nH_{in}^+ \rightarrow N_2O + H_2O + nH_{in}^+$). Nitric oxide reduction is only a 2-electron reaction, so it does not require the cross-linked tyrosine cofactor for catalysis. There are currently three known HCO NORs, the cNOR, qNOR, and qCu_ANOR. The cNOR and qNOR families have a four amino acid coordinated Fe_B ion in their active-sites, in contrast to the three amino acid coordinated Cu_B found in oxygen reductases[15,16]. The cNOR and qNOR families are closely related to the C-family oxygen reductases and

71 likely evolved from this oxygen reductase family. In accordance with this relationship,
72 the cNORs have a secondary cytochrome *c*-containing subunit, while in the qNORs the
73 two subunits corresponding to those in the cNORs have been fused to a single subunit
74 that lacks the heme *c* binding. Although, the qNOR from *N. meningitidis*[16] is proposed to
75 take up protons from the cytoplasm for NO reduction, neither the cNOR nor qNOR have
76 conserved residues that could form a proton channel from the cytoplasm, therefore it is
77 unlikely that they conserve energy via either charge separation or proton pumping (n=0
78 for the cNOR and qNOR families). Phylogenomic analysis shows that the qCu_ANOR
79 from *Bacillus azotoformans*[17,18] is unrelated to the qNOR family, and has been
80 reclassified here as the bNOR family. The bNOR active-site structure is fundamentally
81 different than those from the cNOR and qNOR families **(Figure 1)**. It also contains a
82 conserved proton channel that is very similar to that found in the B-family oxygen
83 reductases (**Figure 1)**, and has been shown to be electrogenic[18]. This has important
84 consequences for the efficiency of energy conservation associated with denitrification[19].
85

86 **Novel heme-copper oxygen reductase homologs**

87 Phylogenomic analyses of genomic and metagenomic data have identified at least six
88 new families belonging to the HCO superfamily (**Figure 2**). All of these families are
89 missing the active-site tyrosine, suggesting that they do not catalyze O_2 reduction.
90 Furthermore, their active-sites exhibit structural features never seen before within the
91 superfamily (**Figure 1**). One of these families is closely related to qNOR and has been
92 proposed to be a nitric oxide dismutase contributing to O_2 production in '*Candidatus*
93 Methylomirabilis oxyfera'[20]. Another family is closely related to cNOR and might be a

94 sulfide and acetylene insensitive nitrous oxide reductase[21,22]. The remaining four families
95 (eNOR, sNOR, nNOR and gNOR) are closely related to the B-family of oxygen
96 reductases (**Figure 2**) and encode for homologs of the Cu_A-containing secondary
97 subunits, consistent with this evolutionary relationship (**Figure 1, Table S1**). Based on
98 modeled active-site structures and genomic context we propose that these four families
99 perform nitric oxide reduction (**Figure 1**).
100

101 **Biochemical characterization of eNOR**

102 To validate these predictions we isolated and biochemically characterized a member of
103 the eNOR family from *Rhodothermus marinus* DSM 4252, a thermophilic member of the
104 Bacteroidetes phylum. *Rhodothermus marinus* was originally classified as a strict
105 aerobe[23], but its genome encodes a periplasmic nitrate reductase (NapA), two nitrite
106 reductases (NirS and NirK), and a nitrous oxide reductase (NosZ), suggesting that it may
107 also be capable of denitrification (**Extended Data Figure 1**). Denitrification was not
108 observed under strictly anaerobic conditions, however under microoxic conditions
109 isotopically labeled $^{15}NO_3^-$ was converted to $^{30}N_2$ (**Extended Data Figure 2**)
110 demonstrating that *R. marinus* DSM 4252 is capable of complete aerobic denitrification
111 ($NO_3^- \rightarrow N_2$). Blockage of the nitrous oxide reductase with acetylene led to the
112 accumulation of N_2O (**Figure 3**), suggesting that a nitric oxide reductase was also present
113 in *R. marinus* DSM 4252. No known NORs (cNOR, qNOR, qCuANOR) were found in
114 the genome. However, *R. marinus* DSM 4252 does encode for a member of the eNOR
115 family (**Extended Data Figure 1**).
116

117 Isolation and biochemical characterization of the *R. marinus* DSM 4252 eNOR protein
118 verified that it catalyzes NO reduction (at 25°C, k_{cat} = 0.68 ± 0.21 $NO.s^{-1}$ (n = 4))
119 (**Figure 3**). eNOR was unable to catalyze O_2 reduction using various electron donors
120 (**Extended Data Figure 3**), clearly showing that it functions solely as a nitric oxide
121 reductase. UV-Vis spectroscopy and heme characterization via mass spectrometry
122 demonstrated that the *R. marinus* DSM 4252 eNOR contains a novel modified heme *a*
123 that is used in both heme sites (**Figure 3 and Extended Data Figures 3 and 4**). A
124 member of the eNOR family was previously isolated from the aerobic denitrifier
125 *Magnetospirillum magnetotacticum* MS-1[16,17], however its function was never
126 determined. The UV-Vis spectra of the *M. magnetotacticum* eNOR[17] is identical to the *R.*
127 *marinus* eNOR, suggesting that the modified heme *a* is a general feature of the family.
128 Mass spectrum analysis of the hemes extracted from eNOR suggest that this heme is A_s, a
129 previously isolated heme *a* with a hydroxyethylgeranylgeranyl side chain first identified
130 in the B-family oxygen reductase from *Sulfolobus acidocaldarius*[24]. Many eNOR operons
131 contain a CtaA homolog, an O_2-dependent enzyme that converts heme *o* to heme *a*[25].
132 This is consistent with the observation that eNOR requires microoxic conditions for
133 expression and suggests that aerobic denitrification might be common in nature.
134

135 **Active-site features of novel NORs**

136 In addition to the experimental evidence that both eNOR and bNOR enzymes are NO
137 reductases, there is good evidence that the other newly identified families also perform
138 nitric oxide reduction. The sNOR family has the same active-site structure as the bNOR
139 family, strongly suggesting that it also performs nitric oxide reduction. However, the

140 sNOR and bNOR families are not closely related, providing an example of convergent
141 evolution of active-site structures within the HCO superfamily (**Figures 1 and 2**).
142 Another example of convergent evolution is the nNOR family. It has the same conserved
143 active-site residues as the cNOR and qNOR families (**Figure 1**), but is very distantly
144 related to them (nNOR is related to the B-family, while cNOR and qNOR are related to
145 the C-family). Interestingly, the low-spin heme in nNOR is ligated by a histidine and
146 methionine, which likely raises its redox potential by ~150 mV[26]. This is similar to a
147 modification found in some eNORs, where the low spin heme is ligated by histidine and
148 lysine. The gNOR is the first example of a HCO family that has replaced one of the
149 active-site histidines, residues completely conserved in all other families. A bioinorganic
150 mimic of the gNOR active-site exhibited nitric oxide reduction capability[27], suggesting
151 that it is likely a functional NOR *in vivo*. The gNOR has a secondary subunit with a
152 cupredoxin fold that is missing the residues required to bind Cu_A, similar to the quinol
153 oxygen reductase from *E. coli*. Conserved residues that could bind quinol have been
154 identified in gNOR, suggesting that it is a quinol nitric oxide reductase similar to qNOR.
155
156 The biochemically characterized (eNOR and bNOR) and proposed (sNOR and gNOR)
157 HCO NORs have active-sites that differ significantly from those utilized by cNOR and
158 qNOR (**Figure 1**). This demonstrates that while oxygen reduction chemistry is
159 constrained to require a redox active tyrosine cofactor, multiple HCO active-site
160 structures are compatible with nitric oxide reduction chemistry. Interestingly, in the
161 currently characterized HCOs Cu_B is utilized for O_2 reduction chemistry and Fe_B for NO
162 reduction chemistry. If this pattern is verified for the other predicted NOR families it

163 would suggest that the chemistry performed by HCOs is partially determined by the
164 electronic properties of the active-site metal.
165
166 **Bioenergetics of novel denitrification pathways**
167 Although both denitrification and aerobic respiration are highly exergonic processes,
168 most of the enzymes in the denitrification pathway are not coupled to energy
169 conservation, making denitrification significantly less efficient than aerobic respiration[28].
170 In the HCO oxygen reductases conserved proton channels deliver protons from the
171 cytoplasm to the active-site for chemistry. These same channels are used to pump protons
172 to the periplasmic side. The cNORs and qNORs do not have conserved proton channels
173 from the cytoplasm[15,29], making them incapable of conserving energy. The eNOR has
174 conserved residues that closely resemble those found in the proton-conducting K-channel
175 within the B-family of oxygen reductases[30] (**Supplementary Table 1, Extended Data**
176 **Figure 5**). The sNOR family also has conserved residues in the K-channel region,
177 however they are different than those found in the eNOR and bNOR families.
178 Interestingly, the nNOR family, which has the same active site as cNOR and qNOR, also
179 has a conserved proton channel (**Supplementary Table 1, Extended Data Figure 5**),
180 suggesting that the lack of a proton channel in the cNOR and qNOR may not be due to
181 energy constraints[31]. The conserved proton channels in the eNOR, bNOR, sNOR, and
182 nNOR families would allow them to conserve energy via charge separation, and
183 potentially by proton pumping. Characterization of these new NOR families will be
184 critical for a full understanding of the mechanism of proton pumping in the HCO
185 superfamily, one of the most important unanswered question in bioenergetics.

186 **Distribution and environmental relevance of NORs**

187 The new HCO NOR families have broad phylogenetic and environmental distributions
188 (**Table 1, Supplementary Tables 2, 3**). The eNOR, sNOR, gNOR, and nNOR families
189 are found in both Bacteria and Archaea. To date, the bNOR family has been found only
190 in the Bacillales order of Firmicutes (**Supplementary Table 2**). Phylogenetic analysis of
191 metagenomic data shows that the majority of eNOR, sNOR, gNOR, and nNOR enzymes
192 are from uncharacterized organisms, suggesting that many more organisms are capable of
193 nitric oxide reduction than previously suspected. Furthermore, the new HCO NOR
194 families are found in a wide variety of environments (**Table 1, Supplementary Table 2**).
195 The sNOR are found in the majority of ammonia oxidizing bacteria sequenced to date
196 and likely play a role in this process. The gNORs are predominantly found in sulfidic
197 environments and may be an adaptation that allows for denitrification in the presence of
198 sulfide, which inhibits other NOR families. The eNOR family is very common in nature,
199 having a broad distribution similar to the cNOR and qNOR families. The eNOR family
200 appears to play key roles in a number of important microbiological processes. They are
201 found in many strains of *Candidatus Accumulibacter phosphatis*, a microbe utilized for
202 phosphate accumulation in wastewater treatment plants during enhanced biological
203 phosphorus removal. The eNOR is highly expressed in transcriptomic datasets from
204 these facilities, demonstrating that *Accumulibacter phosphatis* is capable of complete
205 denitrification *in situ*[32]. eNOR has also been found in microbes capable of performing
206 autotrophic nitrate reduction coupled to Fe(II) oxidation (NRFO). *Gallionellaceae* KS
207 and related strains express an eNOR under denitrifying conditions, suggesting that an
208 individual organism is capable of complete NRFO[33]. eNOR is also commonly found in

209 hypersaline environments (**Supplementary Table 2**) where is might play a role in the
210 adaptation of denitrification to high salt conditions.
211
212 Many organisms encode NORs from multiple families (e.g., *Ca.* Methylomirabilis
213 oxyfera has a qNOR, sNOR and gNOR, and *Bacillus azotoformans* has a qNOR, sNOR,
214 and bNOR). This suggests that selection for different enzymatic properties (NO affinity,
215 enzyme kinetics, energy conservation, or sensitivity to inhibitors) or the concentration of
216 O_2 may be important factors in determining their distribution, similar to what is observed
217 for the HCO oxygen reductase families[8]. Analysis of the presence of denitrification genes
218 (nitrite reductases, nitric oxide reductases, and nitrous oxide reductases) within
219 sequenced genomes indicates that many more organisms are capable of complete
220 denitrification than previously realized. Our current understanding of the diversity of
221 organisms capable of performing denitrification is far from complete.
222
223 Our evolutionary analysis shows that nitric oxide reductases have evolved many times
224 independently from oxygen reductases (**Figure 2**). The current data show that NORs
225 have originated from both the B and C-families of oxygen reductases, enzymes that are
226 adapted to low O_2 environments. These oxygen reductases can reduce NO at high
227 concentrations *in vitro*[34] so it is not surprising that small evolutionary modifications
228 would lead to enzymes capable of NO reduction at the lower NO concentrations
229 produced during denitrification. The fact that NO reductases are derived from oxygen
230 reductases strongly suggests that complete denitrification evolved after aerobic

respiration. This places important constraints on the nitrogen cycle before the rise of oxygen.

Competing Interests

The authors declare no competing interests.

Acknowledgment

We would like to thank NIH for funding this research (Grant # , Principle Investigator: Dr. Robert Gennis). We thank Sylvia Choi for providing pure ba_3 oxygen reductase from *Thermus thermophilus* to use as a control for oxygen reductase assays and for heme extraction, Lici Schurig-Briccio for guidance in performing nitric oxide reductase assays with the Clark Electrode and Peter Yau at the University of Illinois' Mass spectrometric facility for protein identification. We thank Alon Philosof and Connor Skennerton for valuable discussions on bioinformatics analysis. A portion of this research was performed under the Facilities Integrating Collaborations for User Science (FICUS) initiative (award 503546 to R.H.) and used resources at the DOE Joint Genome Institute and the Environmental Molecular Sciences Laboratory, which are DOE Office of Science User Facilities. Both facilities are sponsored by the Office of Biological and Environmental Research and operated under Contract Nos. DE-AC02-05CH11231 (JGI) and DE-AC05-76RL01830 (EMSL).

References

1. Thamdrup, B. New Pathways and Processes in the Global Nitrogen Cycle. *Annu. Rev. Ecol. Evol. Syst.* **43**, 407–428 (2012).

255 2. Klotz, M. G. & Stein, L. Y. Nitrifier genomics and evolution of the nitrogen cycle.
256 *FEMS Microbiology Letters* **278**, 146–156 (2008).
257 3. Zumft, W. G. Nitric oxide reductases of prokaryotes with emphasis on the
258 respiratory, heme–copper oxidase type. *J. Inorg. Biochem.* **99**, 194–215 (2005).
259 4. Hemp, J. & Gennis, R. B. Diversity of the Heme–Copper Superfamily in Archaea:
260 Insights from Genomics and Structural Modeling. *Bioenergetics* 1–31 (2008).
261 5. Pereira, M. M., Sousa, F. L., Veríssimo, A. F. & Teixeira, M. Looking for the
262 minimum common denominator in haem–copper oxygen reductases: Towards a
263 unified catalytic mechanism. *Biochim. Biophys. Acta BBA - Bioenerg.* **1777**, 929–934
264 (2008).
265 6. Hemp, J., Christian, C., Barquera, B., Gennis, R. B. & Martínez, T. J. Helix
266 Switching of a Key Active-Site Residue in the Cytochrome cbb_3 Oxidases.
267 *Biochemistry* 44, 10766–10775 (2005).
268 7. Hemp, J. *et al.* Evolutionary migration of a post-translationally modified active-Site
269 residue in the proton-pumping heme-copper oxygen reductases. *Biochemistry* **45**,
270 15405–15410 (2006).
271 8. Han, H. *et al.* Adaptation of aerobic respiration to low O_2 environments. *Proc. Natl.*
272 *Acad. Sci.* **108**, 14109–14114 (2011).
273 9. Hemp, J. *et al.* Comparative Genomics and Site-Directed Mutagenesis Support the
274 Existence of Only One Input Channel for Protons in the C-Family (cbb3 Oxidase) of
275 Heme−Copper Oxygen Reductases. *Biochemistry* **46**, 9963–9972 (2007).
276 10. Chang, H.-Y., Hemp, J., Chen, Y., Fee, J. A. & Gennis, R. B. The cytochrome ba_3
277 oxygen reductase from *Thermus thermophilus* uses a single input channel for proton

278 delivery to the active site and for proton pumping. *Proc. Natl. Acad. Sci.* **106**, 16169–
279 16173 (2009).
280 11. Qin, L., Hiser, C., Mulichak, A., Garavito, R. M. & Ferguson-Miller, S. Identification
281 of conserved lipid/detergent-binding sites in a high-resolution structure of the
282 membrane protein cytochrome *c* oxidase. *Proc. Natl. Acad. Sci.* **103**, 16117–16122
283 (2006).
284 12. Soulimane, T. *et al.* Structure and mechanism of the aberrant ba_3-cytochrome *c*
285 oxidase from *Thermus thermophilus*. *EMBO J.* **19**, 1766–1776 (2000).
286 13. Tiefenbrunn, T. *et al.* High resolution structure of the ba_3 cytochrome *c* oxidase from
287 *Thermus thermophilus* in a lipidic environment. *PLOS ONE* **6**, 1-12 (2011).
288 14. Buschmann, S. *et al.* The Structure of cbb_3 Cytochrome Oxidase Provides Insights
289 into Proton Pumping. *Science* **329**, 327–330 (2010).
290 15. Hino, T. *et al.* Structural basis of biological N_2O generation by bacterial nitric oxide
291 reductase. *Science* **330**, 1666–1670 (2010).
292 16. Gonska, N. *et al.* Characterization of the quinol-dependent nitric oxide reductase
293 from the pathogen *Neisseria meningitidis*, an electrogenic enzyme. *Sci. Rep.* **8**, 3637
294 (2018).
295 17. Suharti, Strampraad, M. J., Schröder, I. & de Vries, S. A novel copper A containing
296 menaquinol NO reductase from *Bacillus azotoformans. Biochemistry* **40**, 2632–2639
297 (2001).
298 18. Al-Attar, S. & de Vries, S. An electrogenic nitric oxide reductase. *FEBS Lett.* **589**,
299 2050–2057 (2015).

300 19. Strohm, T. O., Griffin, B., Zumft, W. G. & Schink, B. Growth Yields in Bacterial
301 Denitrification and Nitrate Ammonification. *Appl. Environ. Microbiol.* **73**, 1420–
302 1424 (2007).
303 20. Hanson, B. T. & Madsen, E. L. In situ expression of nitrite-dependent anaerobic
304 methane oxidation proteins by *Candidatus* Methylomirabilis oxyfera co-occurring
305 with expressed anammox proteins in a contaminated aquifer. *Environ. Microbiol.*
306 *Rep.* **7**, 252–264 (2015).
307 21. Jones, A. M., Hollocher, T. C. & Knowles, R. Nitrous oxide reductase of *Flexibacter*
308 *canadensis*: a unique membrane-bound enzyme. *FEMS Microbiol. Lett.* **92**, 205–209
309 (1992).
310 22. Jones, A. M., Adkins, A. M., Knowles, R. & Rayat, G. R. Identification of a
311 denitrifying gliding bacterium, isolated from soil and able to reduce nitrous oxide in
312 the presence of sulfide and acetylene, as *Flexibacter canadensis*. *Can. J. Microbiol.*
313 **36**, 765–770 (1990).
314 23. Alfredsson, G. A., Kristjansson, J. K., Hjrleifsdottir, S. & Stetter, K. O.
315 *Rhodothermus marinus*, gen. nov., sp. nov., a Thermophilic, Halophilic Bacterium
316 from Submarine Hot Springs in Iceland. *J. Gen. Microbiol.* 299–306 (1988).
317 24. Lübben, M. & Morand, K. Novel prenylated hemes as cofactors of cytochrome
318 oxidases. Archaea have modified hemes A and O. *J. Biol. Chem.* **269**, 21473–21479
319 (1994).
320 25. Brown, K. R., Allan, B. M., Do, P. & Hegg, E. L. Identification of Novel Hemes
321 Generated by Heme A Synthase: Evidence for Two Successive Monooxygenase
322 Reactions. *Biochemistry* **41**, 10906–10913 (2002).

26. Fufezan, C., Zhang, J. & Gunner, M. R. Ligand preference and orientation in *b*- and *c*-type heme-binding proteins. *Proteins Struct. Funct. Bioinforma.* **73**, 690–704 (2008).

27. Lin, Y.-W. *et al.* Introducing a 2-His-1-Glu nonheme iron center into myoglobin confers nitric oxide reductase activity. *J. Am. Chem. Soc.* **132**, 9970–9972 (2010).

28. Chen, J. & Strous, M. Denitrification and aerobic respiration, hybrid electron transport chains and co-evolution. *Biochim. Biophys. Acta BBA - Bioenerg.* **1827**, 136–144 (2013).

29. Matsumoto, Y. *et al.* Crystal structure of quinol-dependent nitric oxide reductase from *Geobacillus stearothermophilus*. *Nat. Struct. Mol. Biol.* **19**, 238–245 (2012).

30. Chang, H.-Y. *et al.* Exploring the proton pump and exit pathway for pumped protons in cytochrome ba_3 from *Thermus thermophilus*. *Proc. Natl. Acad. Sci.* **109**, 5259–5264 (2012).

31. Blomberg, M. R. A. & Siegbahn, P. E. M. Why is the reduction of NO in cytochrome *c* dependent nitric oxide reductase (cNOR) not electrogenic? *Biochim. Biophys. Acta BBA - Bioenerg.* **1827**, 826–833 (2013).

32. Camejo, P. Y., Oyserman, B. O., McMahon, K. D. & Noguera, D. R. Integrated Omic Analyses Provide Evidence that a "*Candidatus* Accumulibacter phosphatis" Strain Performs Denitrification under Microaerobic Conditions. *mSystems* **4**, e00193-18.

33. Huang, Y.-M., Straub, D., Blackwell, N., Kappler, A. & Kleindienst, S. Meta-omics Reveal *Gallionellaceae* and *Rhodanobacter* Species as Interdependent Key Players for Fe(II) Oxidation and Nitrate Reduction in the Autotrophic Enrichment Culture KS. *Appl. Environ. Microbiol.* **87**, e00496-21.

346 34. Loullis, A. & Pinakoulaki, E. Probing the nitrite and nitric oxide reductase activity of
347 *cbb$_3$* oxidase: resonance Raman detection of a six-coordinate ferrous heme–nitrosyl
348 species in the binuclear *b$_3$*/Cu$_B$ center. *Chem. Commun.* **51**, 17398–17401 (2015).

349

350 **List of Figures**

351 **Figure 1. Comparison of HCO active sites.** a) Active-site and proton channel properties
352 of the five characterized HCO families (A-family, B-family, C-family, cNOR, and
353 qNOR). The oxygen reductases all have an active-site composed of high-spin heme, a
354 redox-active cross-linked tyrosine cofactor, and a copper (Cu$_B$) ligated by three
355 histidines. The A-family has two conserved proton channels, whereas the B and C-
356 families only have one. The active-sites of the nitric oxide reductases are composed of a
357 high-spin heme and an iron (Fe$_B$) that is ligated by three histidines and a glutamate.
358 Notably they are missing the tyrosine cofactor. The cNOR and qNOR are also missing
359 conserved proton channels, making them non-electrogenic. b) Sequence alignment of the
360 active-sites of the newly discovered HCO families that are related to the B-family. c)
361 Predicted active-sites and proton channels for the new HCO families. The eNOR, bNOR,
362 sNOR, and nNOR families contain completely conserved proton channels shown here as
363 arrows. The putative proton channel in the bNOR and eNOR families are highly similar
364 to the K-channel from the B-family oxygen reductase and are colored in red. The K-
365 channel in the B-family is also similar to the K-channel in the A-family oxygen reductase
366 which is colored in dark red. The proton channel in the C-family is different from these
367 channels and is marked in yellow. The putative proton channels in sNOR and nNOR are
368 marked in cyan and differentiated from the other channels with a dashed black outline.

369

370 **Figure 2. Evolution of nitric oxide reductases.** Phylogenetic tree of HCO families. All
371 of the new NOR families are derived from oxygen reductase ancestors. Oxygen
372 reductases are in shades of blue, whereas nitric oxide reductases are in shades of yellow,
373 green and red.

374 **Figure 3. Biochemical Characterization of the eNOR from *Rhodothermus marinus*.**
375 a) UV-Vis spectrum of isolated eNOR indicates the presence of an unusual heme *a*
376 signature at 589 nm. b) NO reductase activity was measured with the use of a Clark
377 electrode in the presence of TMPD and ascorbate as electron donor. c) N_2O accumulation
378 observed in a culture of *Rhodothermus marinus*, in the presence of acetylene d) N_2O
379 production by eNOR from *R. marinus* and qNOR from *Persephonella marina*.

380 **Extended Data Figure 1: Genome of *R. marinus* encodes for the complete**
381 **denitrification pathway.** a. The genes for NapAB, the periplasmic nitrate reductase
382 (Rmar_0413), the nitrite reductases nirK (Rmar_1208) and nirS (Rmar_0652) and nitrous
383 oxide reductase, nosZ (Rmar_2012) are encoded in the *R. marinus* genome. b. The gene
384 neighborhood of eNOR in *R. marinus*. c. The gene neighborhood of eNOR in *Halovivax*
385 *ruber* includes ctaB, an enzyme involved in the biosynthesis of heme *a*.

386

387 **Extended Data Figure 2: *Rhodothermus marinus* does perform complete**
388 **denitrification.** a. *R. marinus* converts $^{15}NO_3^-$ to $^{30}N_2$. Ratio of $^{30}N_2$ to $^{28}N_2$ for each
389 sample. Air is ambient atmosphere as a standard. C0 is a nitrate-free control. 14 0-17 are
390 cultures grown with unlabeled nitrate, transferred to sealed vials after 0-17 hours,
391 respectively. 15 0-17 are the equivalent samples grown with ^{15}N-labeled nitrate. Error
392 bars represent two standard deviations from three replicate GC/MS measurements. $^{30}N_2$
393 enrichments from the ^{15}N-labeled samples are over 30-60x higher than background
394 atmospheric ratios, while unlabeled samples have no significant enrichment over
395 background. b. Growth of *R. marinus*, measured using OD_{600nm} over 39 hours. NO_3^-
396 utilization was established by measuring the concentrations of nitrate in the media using a
397 calorimetric assay. c. *R. marinus* growth in rich media was compared under denitrifying
398 and non-denitrifying conditions using OD_{600nm}. d. Phenotypic differences of *R. marinus*
399 cultures, under denitrifying and non-denitrifying conditions.
400
401 **Extended Data Figure 3: Characteristics of eNOR from *Rhodothermus marinus*** a.
402 SDS-PAGE gel electrophoresis of eNOR shows two bright bands which are estimated to
403 subunits of I and II of the complex. Both subunits appear to run faster than their
404 estimated molecular weight. This is typical for membrane proteins. For comparison, an
405 SDS-PAGE gel of cytochrome bo_3 oxidase from *E. coli* is included. b. Mass
406 spectrometric identification of eNOR is confirmed by LC/MS/MS analysis. c,d. Absence
407 of O_2 reduction by *R. marinus* eNOR, in comparison to robust O_2 reduction by *T.*
408 *thermophilus* ba_3-type oxygen reductase. e. UV-visible spectrum of eNOR f. Pyridine
409 hemochrome-spectra of extracted hemes from eNOR showing a peak which is atypical of
410 hemes *a, b* or *c*.
411
412 **Extended Data Figure 4: Identification of hemes extracted from eNOR.** Comparing
413 the elution profile of extracted hemes from partially purified *R. marinus* eNOR to bovine
414 cytochrome *c* oxidase (A-type, t=16 min) , *T. thermophilus* ba_3-type oxygen reductase (*b*-
415 and A_s-type hemes, t=12 min and t=19 min) reveals that the heme is most likely an A_s-
416 type heme. Mass spectra of the peak at ~19 min from the eNOR hemes elution profile
417 confirms that the heme is an A_s-type heme with a molecular weight of 920 Da.
418
419 **Extended Data Figure 5: Proton channel in eNOR of *Rhodothermus marinus* and in**
420 **the NOR families bNOR, sNOR and nNOR. a.** eNOR contains conserved residues in
421 Helix VII, similar to the location of K-proton channel residues in *T. thermophilus* ba_3-
422 type oxygen reductase. b. A multiple sequence alignment of the NOR families eNOR,
423 bNOR, sNOR and nNOR show conserved amino acids in analogous location to the K-
424 channel in the B-type oxygen reductase. Some conserved residues are also identified in
425 gNOR and may indicate the presence of a conserved proton channel but they do not map
426 to corresponding residues in the B-type oxygen reductase.
427
428 **Extended Data Figure 6: Conserved amino acids in the eNOR family of enzymes.**
429 Multiple sequence alignment of 23 eNOR sequences from various taxonomically
430 divergent organisms reveals conserved residues that correspond to the active site ligands,
431 proton channel residues and other sequence features that are unique to eNOR. The active

site residues are highlighted in maroon while the proton channel residues are highlighted in blue.

List of Tables

Table 1. Environmental distribution of the HCO NOR families. Distribution of NOR families in sequenced genomes versus environmental datasets. The newly discovered NOR families account for approximately 2/3 of currently known diversity and 1/2 of the abundance of NORs in nature.

Supplementary Table 1. Putative proton channels in the new NOR families – eNOR, bNOR, sNOR, nNOR. A list of conserved residues is noted in the table for each family with a reference sequence according to which the residues are numbered. These conserved residues are compared with amino acids found in analogous positions in the B-type oxygen reductase.

Supplementary Table 2. Distribution of NOR families by phylum in GTDB. A distribution of all the NOR families within various bacterial and archaeal phyla within the genomes in release 202 of GTDB was analyzed using HMMs that are specific to each NOR family.

Supplementary Table 3. Distribution of NOR families in various ecosystems as per the IMG database. A distribution of various NOR families in over 2000 metagenomes on the IMG database was evaluated, and then tabulated according to the environment from which each metagenome is sourced.

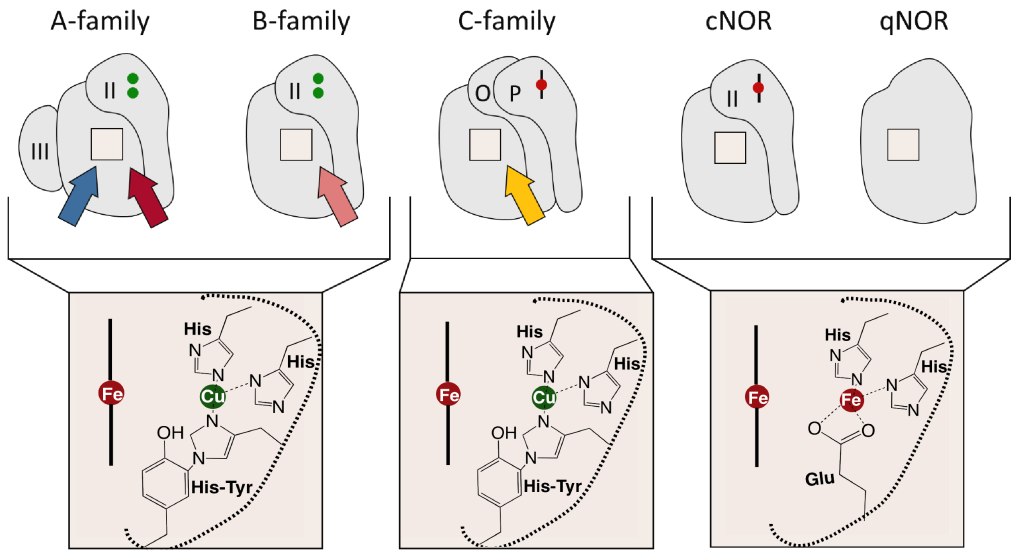
Supplementary Table 4. Denitrification pathways in bacteria and archaea. An analysis of denitrification pathways in bacterial genomes and archaeal genomes in release 202 of GTDB was performed by searching for the presence and absence of NarGHI, NapAB, NirK, NirS, NosZ, NosD and the NORs in each genome using curated HMMs for each of the proteins.

Table 1. Environmental distribution of the HCO NOR families. Distribution of NOR families in sequenced genomes versus environmental datasets. The newly discovered NOR families account for approximately 2/3 of currently known diversity and 1/2 of the abundance of NORs in nature.

	NCBI-Genomes	IMG-metagenomes	GTDB-genomes
A-family	20290	102368	45135
B-family	1238	4683	2021
C-family	13976	23015	14981
qNOR	4388	7680	3458
cNOR	2801	4824	2594
eNOR	68	2709	547
sNOR	95	872	344
bNOR	51	12	200
nNOR	6	289	32
gNOR	10	913	156
NOD	8	539	108
N2O	25	597	n.a.

Figure 1. Comparison of HCO active sites. a) Active-site and proton channel properties of the five characterized HCO families (A-family, B-family, C-family, cNOR, and qNOR). The oxygen reductases all have an active-site composed of high-spin heme, a redox-active cross-linked tyrosine cofactor, and a copper (Cu_B) ligated by three histidines. The A-family has two conserved proton channels, whereas the B and C-families only have one. The active-sites of the nitric oxide reductases are composed of a high-spin heme and an iron (Fe_B) that is ligated by three histidines and a glutamate. Notably they are missing the tyrosine cofactor. The cNOR and qNOR are also missing conserved proton channels, making them non-electrogenic. b) Sequence alignment of the active-sites of the newly discovered HCO families that are related to the B-family. c) Predicted active-sites and proton channels for the new HCO families. The eNOR, bNOR, sNOR, and nNOR families contain completely conserved proton channels shown here as arrows. The putative proton channel in the bNOR and eNOR families are highly similar to the K-channel from the B-family oxygen reductase and are colored in red. The K-channel in the B-family is also similar to the K-channel in the A-family oxygen reductase which is colored in dark red. The proton channel in the C-family is different from these channels and is marked in yellow. The putative proton channels in sNOR and nNOR are marked in cyan and differentiated from the other channels with a dashed black outline.

A



B

```
A-family_1M56   PVLYQHILWFFGHPEVYIIVLPAFGIVSHVIATFAK--GYLPMVYAMVAIGVLGFVVWAHHMYT
B-family_3S8F   PLVARTLFWWTGHPIVYFWLLPAYAIIYTILPKQAG--SDPMARLAFLLFLLLSTPVGFHHQFA
eNOR            AAWYRQMYWIIGHGSQQINLAAMITVWYFLTHVVGG--SEKLSRTAFILYLFFINMGAAHHLLA
bNOR            VMVARTLFWAFGHTAVNIWYLTAVSAWYVIVPKIIG--SDMLTRVVIIALVIMNITGGFHHQII
sNOR            PLLSKNLIYAFGHIFANSIIYMGVIAVYEIFPKYTG--VYGNFLIAWNASTLFTMIIYPHHLLM
gNOR            ALLYKNVYWWGLDLIADGLVLIYVAGSWYLLAMLLT--MQHIARAALFVELVVSWFVWSHHLLS
nNOR            PWPFNVAFWLFAHNLMEAMGIMALAAVYALVPLYTR--SPGMGVLAVGLYTLSAIPAFGHHLYT
C-family_3MK7   GATDAMVQWWYGHNAVGFFLTAGFLGIMYYFVPKQA--SYRLSIVHFWALITVYIWAGPHHLHY
cNOR_3O0R       LTRDKFYWWWVVHLWVEGVWELIMGAILAFVLVKIT--IEKWLYVIIAMALISGIIGTGHHYFW
qNOR_3AYF       FTMADFWRWWIIHLWVEGIFEVFAVVVIGFLLVQLR--TVRALYFQFTILLGSGVIGIGHHYYY
                --------------Helix VI--------------  ---------Helix VII--------
```

C

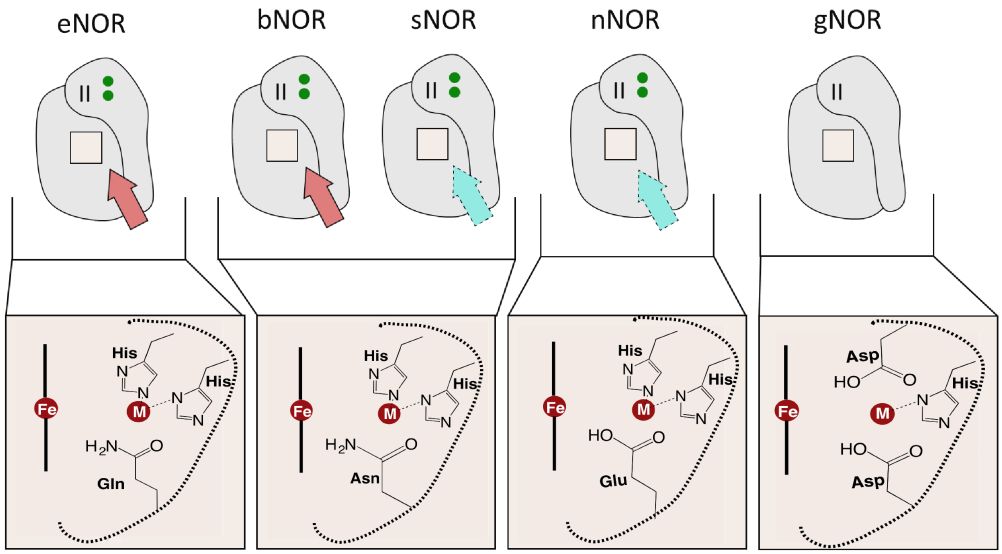


Figure 2. Evolution of nitric oxide reductases. Phylogenetic tree of HCO families. All of the new NOR families are derived from oxygen reductase ancestors. Oxygen reductases are in shades of blue, whereas nitric oxide reductases are in shades of yellow, green and red. The biochemically characterized eNOR from *Rhodothermus marinus* is marked with a black star.

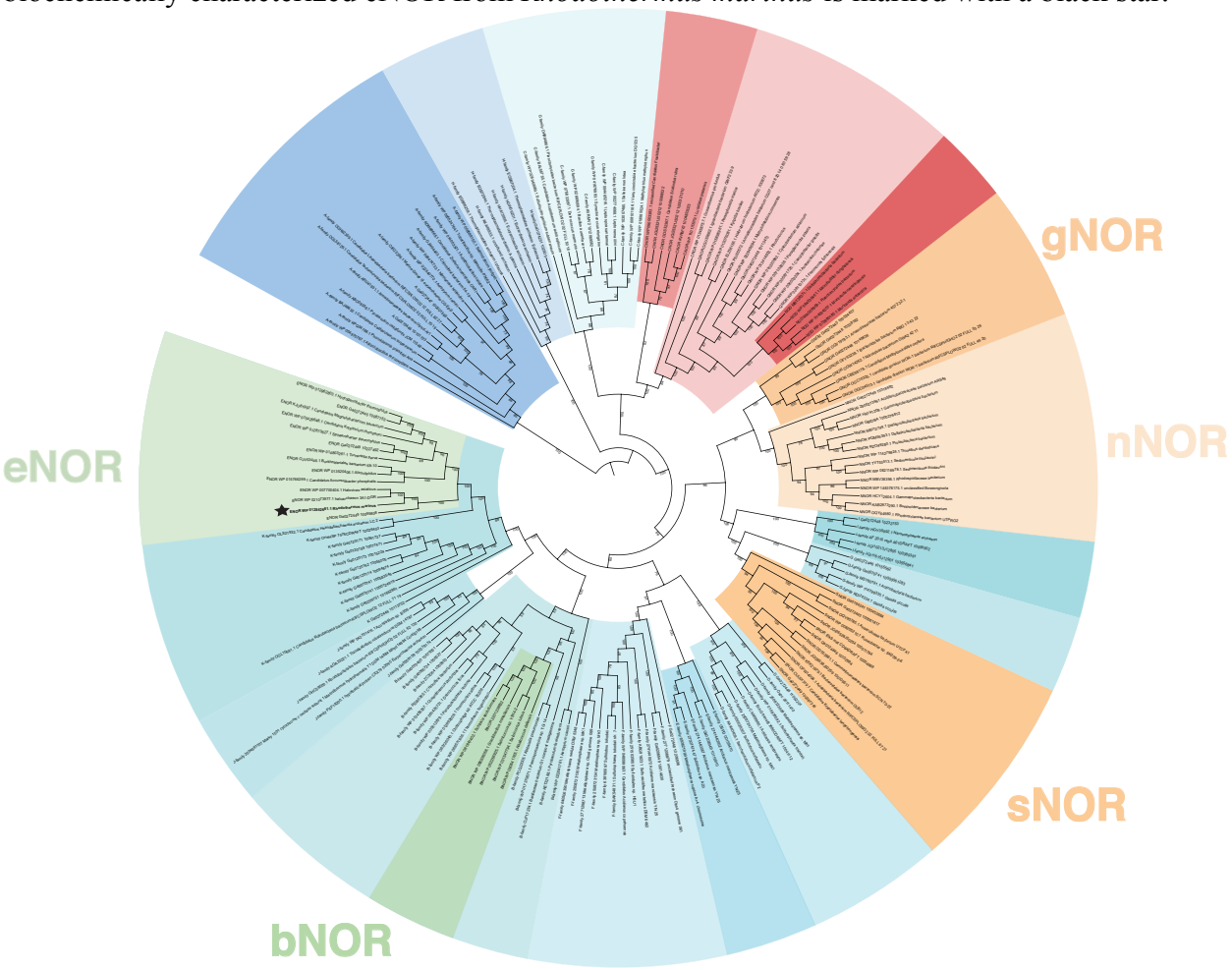
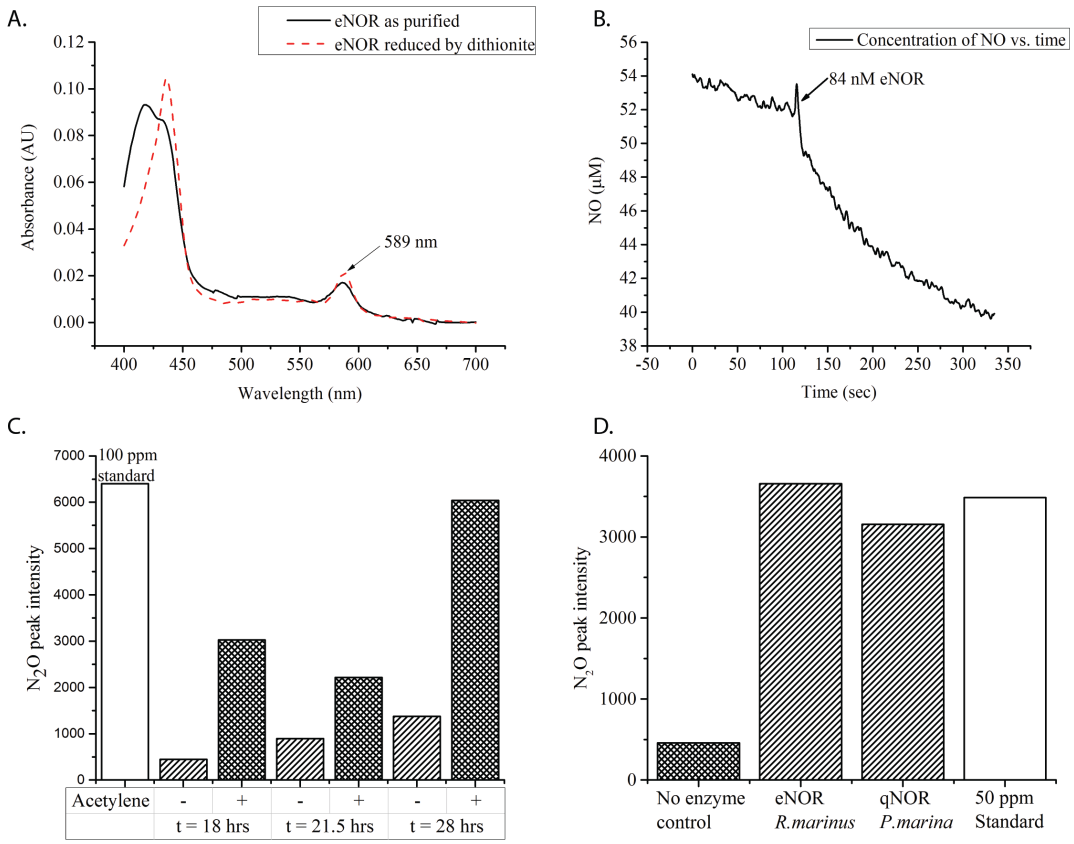
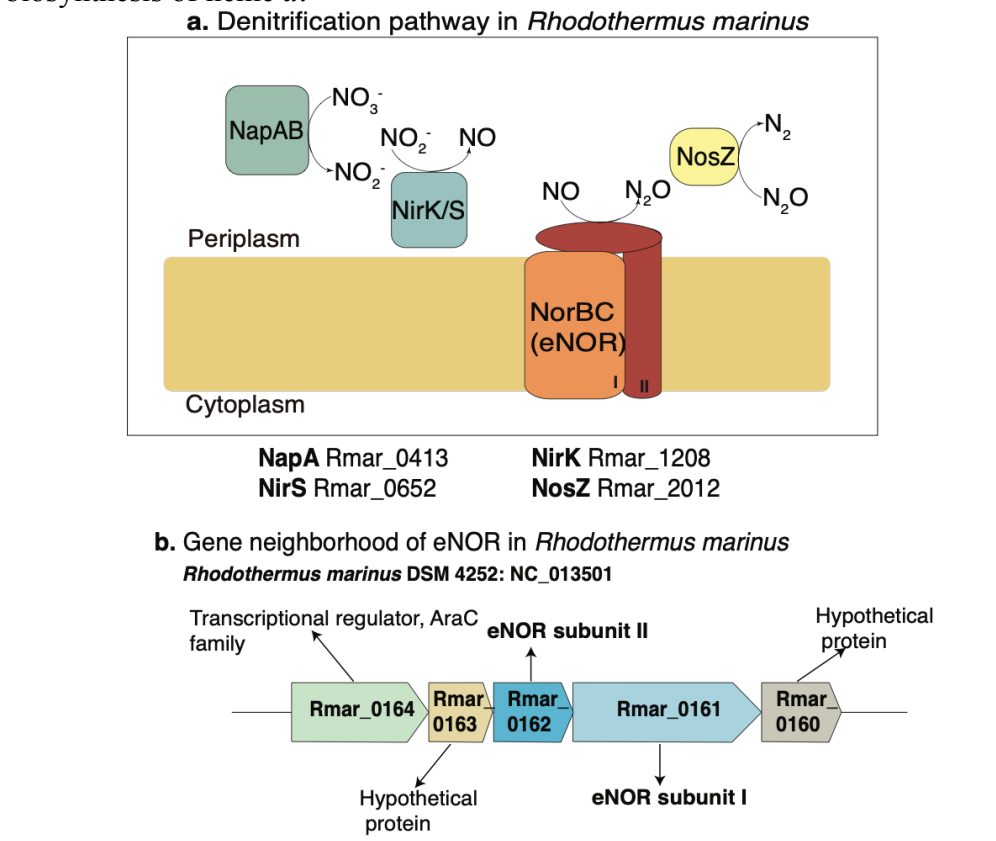


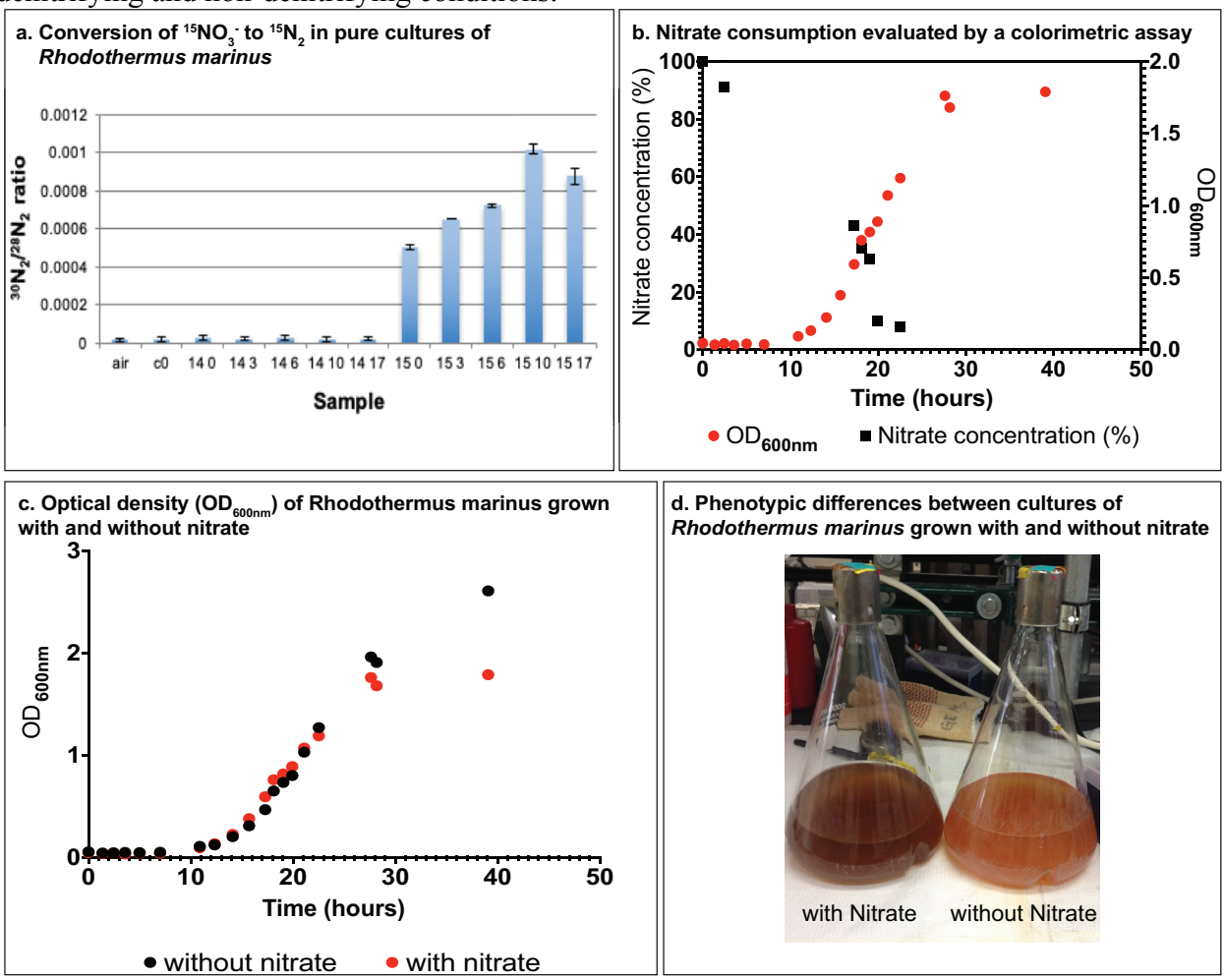
Figure 3. Biochemical Characterization of the eNOR from *Rhodothermus marinus*. a) UV-Vis spectrum of isolated eNOR indicates the presence of an unusual heme *a* signature at 589 nm. b) NO reductase activity was measured with the use of a Clark electrode in the presence of TMPD and ascorbate as electron donor. c) N_2O accumulation observed in a culture of *Rhodothermus marinus*, in the presence of acetylene d) N_2O production by eNOR from *R. marinus* and qNOR from *Persephonella marina*.



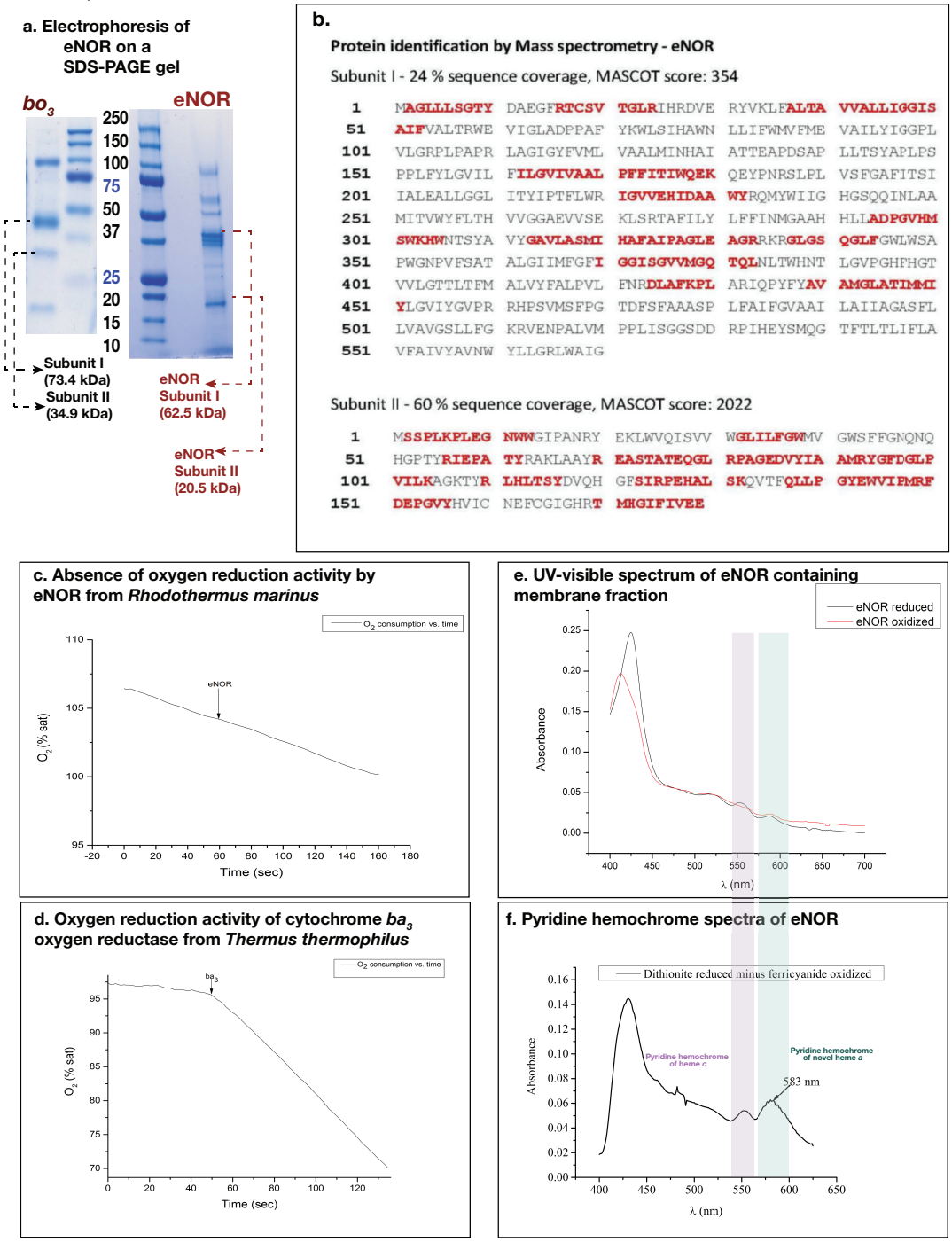
Extended Data Figure 1: Genome of *R. marinus* encodes for the complete denitrification pathway. a. The genes for NapAB, the periplasmic nitrate reductase (Rmar_0413), nitrite reductases nirK (Rmar_1208) and nirS (Rmar_0652), nitric oxide reductase eNOR(Rmar_0161) and nitrous oxide reductase, nosZ (Rmar_2012) are encoded in the *R. marinus* genome. b. The gene neighborhood of eNOR in *R. marinus*. c. The gene neighborhood of eNOR in *Magnetospirillum magneticum AMB-1* includes ctaA and ctaB, enzymes involved in the biosynthesis of heme *a*.



Extended Data Figure 2: *Rhodothermus marinus* does perform complete denitrification. a. *R. marinus* converts $^{15}NO_3^-$ to $^{30}N_2$. Ratio of $^{30}N_2$ to $^{28}N_2$ for each sample. Air is ambient atmosphere as a standard. C0 is a nitrate-free control. 14 0-17 are cultures grown with unlabeled nitrate, transferred to sealed vials after 0-17 hours, respectively. 15 0-17 are the equivalent samples grown with ^{15}N-labeled nitrate. Error bars represent two standard deviations from three replicate GC/MS measurements. $^{30}N_2$ enrichments from the ^{15}N-labeled samples are over 30-60x higher than background atmospheric ratios, while unlabeled samples have no significant enrichment over background. b. Growth of *R. marinus*, measured using OD_{600nm} over 39 hours. NO_3^- utilization was established by measuring the concentrations of nitrate in the media using a calorimetric assay. c. *R. marinus* growth in rich media was compared under denitrifying and non-denitrifying conditions using OD_{600nm}. d. Phenotypic differences of *R. marinus* cultures, under denitrifying and non-denitrifying conditions.

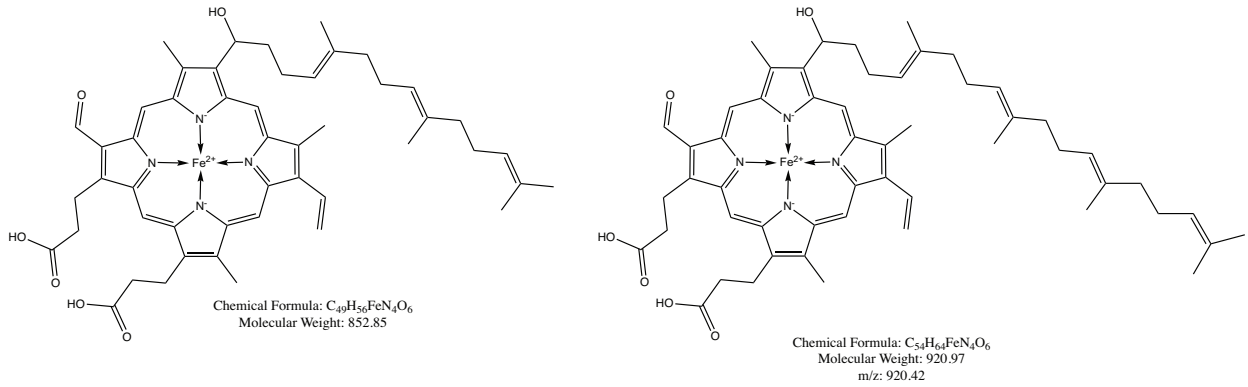


Extended Data Figure 3: Characteristics of eNOR from *Rhodothermus marinus* a. SDS-PAGE gel electrophoresis of eNOR shows two bright bands which are estimated to subunits of I and II of the complex. Both subunits appear to run faster than their estimated molecular weight. This is typical for membrane proteins. For comparison, an SDS-PAGE gel of cytochrome bo_3 oxidase from *E. coli* is included. b. Mass spectrometric identification of eNOR is confirmed by LC/MS/MS analysis. c,d. Absence of O_2 reduction by *R. marinus* eNOR, in comparison to robust O_2 reduction by *T. thermophilus* ba_3-type oxygen reductase. e. UV-visible spectrum of eNOR f. Pyridine hemochrome-spectra of extracted hemes from eNOR showing a peak which is atypical of hemes *a, b* or *c*.

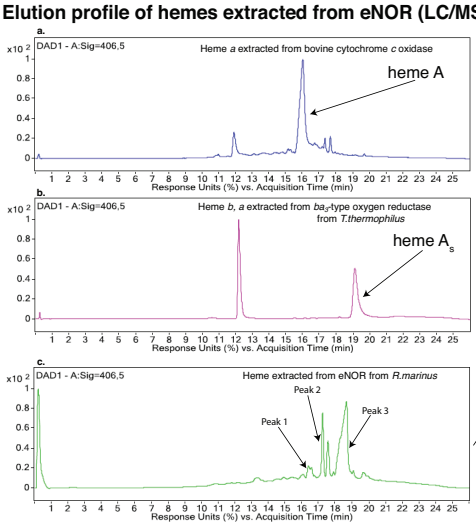


Extended Data Figure 4: Identification of hemes extracted from eNOR. Comparing the elution profile of extracted hemes from partially purified *R. marinus* eNOR to bovine cytochrome *c* oxidase (A-type, t=16 min) , *T. thermophilus* ba_3-type oxygen reductase (*b*- and A_s-type hemes, t=12 min and t=19 min) reveals that the heme is most likely an A_s-type heme. Mass spectra of the peak at ~19 min from the eNOR hemes elution profile confirms that the heme is an A_s-type heme with a molecular weight of 920 Da.

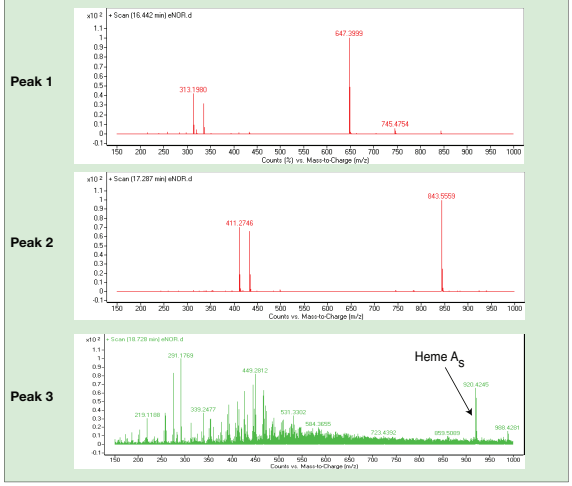
a. Chemical structures of hemes A and heme As.



b. Elution profile of hemes extracted from eNOR (LC/MS)

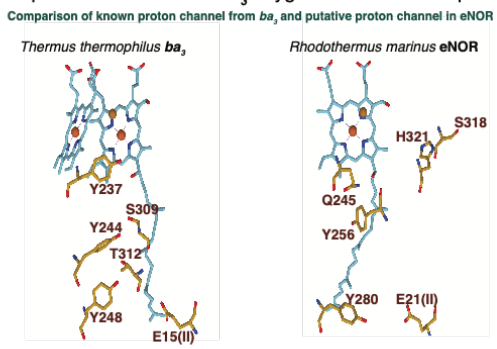


c. Mass spectra corresponding to prominent peaks in the eNOR elution profile



Extended Data Figure 5: Proton channel in eNOR of *Rhodothermus marinus* and in the NOR families bNOR, sNOR and nNOR. a. eNOR contains conserved residues in Helix VII, similar to the location of K-proton channel residues in *T. thermophilus* ba_3-type oxygen reductase. b. A multiple sequence alignment of the NOR families eNOR, bNOR, sNOR and nNOR show conserved amino acids in analogous location to the K-channel in the B-type oxygen reductase. Some conserved residues are also identified in gNOR and may indicate the presence of a conserved proton channel but they do not map to corresponding residues in the B-type oxygen reductase.

a. Comparison of known proton channel from ***ba$_3$*** oxygen reductase and putative proton channel in eNOR

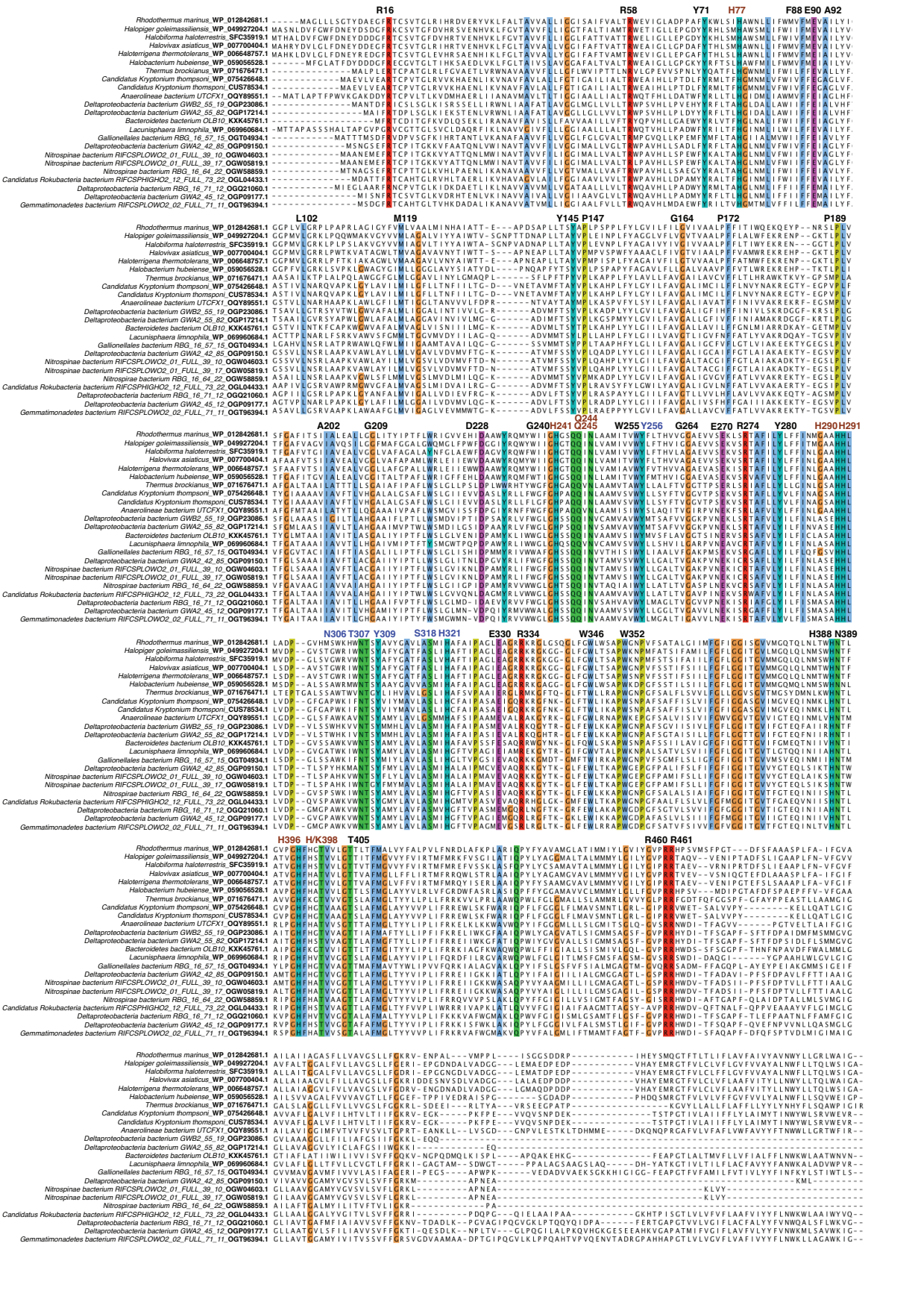


b. Conserved residues forming putative proton channels in eNOR, sNOR, bNOR and nNOR. Conserved proton channel residues are marked in black, while the active site residues are marked in **red**. Each group is numbered according to the protein whose accession number is in **blue**.

```
                                   H255 E259       Y270   Y276      Y284                 H307                                T338 T341 T345
nNOR
Thermomicrobium sp002898255 GBD20489.1      FWLFAHNLMFAMGIMTLGAIYAIVPRYTR-SGQLY-SPRAAVVAMILYTMAAIPAFGHHLYTWVTGNPEVLQNVSR-STSWATGFIAATLTAFNVGLTV----------------WRNGL
Sedimenticola selenatireducens WP_084609916.1 FYIFAHNLMFAMAIMVISAVYATLPLYLADGTRKLYSDKLANLALWILLVTSVTSFFHHFYTTNPGLPSALAY-HGNFMSWATGVGAAL-STFTILATI----------------WKHGI
Sedimenticola selenatireducens PLX61751.1   FYIFAHNLMFAMAIMVISAVYATLPLYLADGTRKLYSDKLANLALWILLVTSVTSFFHHFYTTNPGLPSALAY-HGNFMSWATGVGAAL-STFTILATI----------------WKHGI
Rhodocyclaceae bacterium UTPRO2 OQY64980.1  FFFFAHNLMFAMAIMVASAIYATLPLYLADGSRKLFSDKMANLALWILLVTSITSGLHHFITFYPNQPAALSYWGS-IMSWGTGIGAAL-SIFTVFATI----------------WQHGL
gNOR                               D285 D289       Y301                 H335                                H373
Sulfurimonas autotrophica WP_013326534.1    WWGLDLVADGLVLIYVAGSWYLLATLIT--GQKLF-MENVARAALMLELLVSWMVWSHHLLA-DQGQPEMMKLISGEMVTAFELLTQGL-ALFITLVTL----------------WKARP
Ignavibacteria bacterium GWA2_36_19 OGU38604.1 WWGLDLIADGLVLIFVAGTWYLLATLIT--GKKLF-MENWARAALFVEMVVSWTVWSHHLMS-DQAQPGILKILSGEMVTAFELITQGL-AFFITLATL----------------WSARP
Ignavibacteria bacterium CG2_30_36_16 OIP63421.1 WWGLDLIADGLVLIFVAGTWYLLATLIT--GKKLF-MENWARAALFVEMVVSWTVWSHHLMS-DQAQPGILKILSGEMVTAFELITQGL-AFFITLATL----------------WSARP
candidate division WOR-1 bacterium OGW14262.1 WWGLDLIADGLVLIYVAGTWYLLATLIS--GKELY-MRNVARAALLVELIVSWNVWAHHLLS-DQAQPNIMKIISGEMVTAFELVTMGI-AIFITLKTL----------------WEARP
candidate division WOR-1 bacterium OGC04573.1 WWGLDLIADGLVLIYVAGTWYLLAMIIT--GRQIF-MQNFARAALFVELVVSWFVWSHHLLS-DQTQPVMMRIFSGEMITAFELVTSGI-AVFLTLATL----------------WQARP
sNOR                               H232 N236 Y240 Y247 E248                 H288                                T319
Bacteroidetes bacterium 37-13 OJV27025.1    TFFFGHTIANEALYLGLATLYELLPEVS--GRPKFKTTWYVALGWNCAIIFILGAFFHHLYM-DFVQPKGFQIFGQ-IASYFATIPSVVVTIISIVTLL----------------YNNKI
Geobacillus sp. WP_023634191.1              IYAFGHIFANSVIYMGVIAVYELLPKYT--NRP-WKSYKIFLIAWNMSTLFTIIIYPHHLLM-DFVMPK-WMLIIGQVFSYLNGLPVLVVTAFGALMIV----------------YRSGI
Rhodanobacter denitrificans WP_015448124.1  IYWFGHMVINATIYMGVIAVYELLPRYT--GRP-YGISRPFLWSWAASTVFVIIVFPHHLLM-DYAEPR-WMLVMGQIISYAAGFPVFLVTAYGVLTNI----------------HRSGL
Nitrosococcus halophilus WP_013031504.1     IYFFGHVFINATIYASVTAVYELLPRYT--GRP-WKTSKVFYAAWLAIVFMVMAVYPHHLMM-DFAMPP-WALIVGQVLSYGSGVPVMVVTGYGALMIV----------------YRSGI
Thioalkalivibrio sp. ALRh WP_019592254.1    TYFFGHVFINATIYMAVIGVYELLPRYT--GRP-WKVSRVFLAAWAASTVMVLLVYPHHLLM-DFSQPTSLHVLGQ-VISYTSGLPVLLVTAWGALTNV----------------YRSGI
bNOR                               H233 N237 Y240    Y248                 H281                      S309 T312
Salinicoccus qingdaonensis WP_092985759.1   FWAFGHTLVNVWYLVAVSAWYLVVPKVI--GGKLF-SDSLARAVVILIVVLNVPGGFHHQIV-DPGFTEGLKFMHL-FMSLAIAVPSLL-TAFALFATLERTGRRRG-G-KGLLGWFWKL
Salinicoccus sp. YB14-2 WP_092985759.1      FWAFGHTLVNVWYLVAVSAWYLVVPKVI--GGKLF-SDSLARAVVILIVVLNVPGGFHHQIV-DPGFTEGLKFMHL-FMSLAIAVPSLL-TAFALFATLERTGRRRG-G-KGLLGWFWKL
Jeotgalicoccus psychrophilus WP_026860023.1 FWSFGHTLVNVWYLVAVSAWYIVLPKVI--GGKIF-SDSLARLVVILIVILNVPGGFHHQIV-DPGFTEGLKFMHL-FMSLAIGFPSLM-TAFALFATLERAGRNKG-G-KGLFGWFFKL
Virgibacillus dakarensis WP_088049698.1     FWSFGHTLVNIWYIVATSAWYVVVPKII--GGRVF-SDKLARLVVVLLVILNIPGGFHHQII-DPGISESVKFLHV-FMSISIAFPSLM-TAFAMFAVFERAGRKLG-G-KGLLGWFKKL
Sporosarcina sp. HYO08 KXH87068.1           FWSFGHTLVNIWYLTAVSAWYTIVPKII--GGRRF-SDTLTRVVVILLVITNIPGGFHHQIV-DPGMGEALKYMHV-FMSLAIGFPSLM-TAYAMFSVFERTGRLKG-G-KGLLGWFKKL
Sporosarcina ureilytica WP_075529285.1      FWSFGHTLVNIWYLTAVSAWYVIVPKII--GGKRF-SDTLTRVVVAMLVITNIPGGFHHQIV-DPAMGPALKYMHV-FMSLSIAFPSLM-TAFAMFYVFERTGRAKG-G-KGLFGWLKKL
Bacillus sp. FJAT-27445 WP_059173536.1      FWAFGHTLVNIWYLTAVSAWYVIIPKLM--GGRRW-SDTLTRVVIIALVVMNITGGFHHQIV-DPGITESVKFMHV-FMSLAIGFPSLM-TAYALFAVFEKTARRNG-G-KGILGWYKKM
Bacillus sp. EB01 WP_043932172.1            FWAFGHTLVNIWYLTATSAWYVVIPKLM--GGKRW-SDTLTRVVIISLVVMNITGGFHHQIV-DPGISEAVKYMHV-FMSLAIGFPSLM-TAYALFATFEKTARRNG-G-KGLVGWYKKM
Thalassobacillus cyri WP_093046123.1        FWAFGHTLVNIWYMTAISAWYVFVPKII--DGRRW-SDMLTRIVVIALVIMNITGGFHHQII-DPGISESVKYMHV-FMSLAIGFPSLM-TAYAMFAVFEKTARKKG-G-RGLVGWYKKM
Neobacillus bataviensis WP_007083088.1      FWAFGHTLVNIWYLTAVSAWYVIVPKII--SGKRW-SDTLTRVVIIALVVMNITGGFHHQIV-DPGISPSVKFMHV-FMSLAIGFPSLM-TAWAMFAVFERTARKQG-G-KGLIGWYKKM
eNOR                               H241Q245 N247    Y256                 Y280      H290                      S318 H321
Rhodothermus marinus WP_012842681.1         YWIIGHGSSQQINLAAMITVWYFLTHVVG--GAEVV-SEKLSRTAFILYLFFINMGAAHHLLA-DPGVHMSWKHWNTSYAVYGAVLASMI-HAFAIPAGLEAGRRKRGLGSQGLFGWLWSA
Alicycliphilus denitrificans WP_013520406.1 WWAFGHSSQQINVAAHISIWYAVAAIAF--GAKPM-SERVSRGAFLLYILFLQLASAHHILA-DPGVSTEWKIVNTSYFMYFAVLGSMI-HALSIPGAMEVAQRAKGYN-KGLFEWLRKA
Candidatus Kryptonium thompsoni WP_075426648.1 FWGFGHPAQQINLAAMVSVWYLLSYFTV--GGVTP-SEKVSRTAFLLYVLFINAGSAHHLLV-DPGFGAPWKIFNTSYVIYMAVLASLI-HCFAIPASAEIGQRKRGFN-KGLFTWLIKA
Bacteroidetes bacterium OLB11 KXK43645.1    WWGLGHSSQQINVAAMVSIWYMVSFLAV--GGTSI-NEKVSRSAFVLYILFICLASAHHLLT-DPGVSSAWKVWNTSYAMYLAVLASMI-HAFAVPSSFESAQRRWGYN-KGLFEWLSKA
Gemmatimonadetes bacterium SCN 70-22 ODT01569.1 WWGLGHSSQQINVAAMVAIWYMLAALTV--GGVVL-NEKVSRSAFVLYILFISMASAHHLLV-DPGFGPAWKIVNTSYFMYMAVLASMI-HGFTVPAGMELGMRLRGYT-DGLFGWLRRA
```

Extended Data Figure 6: Conserved amino acids in the eNOR family of enzymes. Multiple sequence alignment of 23 eNOR sequences from various taxonomically divergent organisms reveals conserved residues that correspond to the active site ligands, proton channel residues and other sequence features that are unique to eNOR. The active site residues are highlighted in maroon while the proton channel residues are highlighted in blue.

a. Sequence alignment of eNOR - Subunit I showing conserved amino acid residues according to *R.marinus* eNOR numbering



Materials and Methods

Growth and Expression Conditions

Rhodothermus marinus DSM 4252 was inoculated from frozen stock and grown in 5 ml of DSM Medium 630 with 10 g/L NaCl at 60 °C for 36 hrs. It was then inoculated into a larger secondary culture and grown overnight. 25 ml of the culture was inoculated into 1 L of medium with 30 mM nitrate added. The cells were shaken at 75 rpm and grown at 60 °C. The cells were pelleted by centrifugation at 8000 rpm. The cell pellet was either directly used for protein purification or frozen at -80 °C until the time of use.

Labeled ^{15}NO experiments

We used labeled nitrate ($^{15}NO_3^{2-}$) to verify that *Rhodothermus marinus* DSM 4252 was capable of complete denitrification (NO_3^- to N_2) using eNOR as the sole nitric oxide reductase. Cultures were inoculated into flasks containing media with either $^{14}NO_3^{2-}$, $^{15}NO_3^{2-}$, or no nitrate. The cultures were then allowed to grow microaerobically for a period of time before being subsampled for transfer to sealed vials in order to allow accumulation of gaseous end products. Samples were taken from each media composition after 0, 3, 6, 10, and 17 hours. The headspace was sampled after 20 hours of growth in sealed vials via gastight GC syringe and immediately injected into a Hewlett Packard 5972 gas chromatograph/mass spectrometer. Chromatogram peaks corresponding to isotopologues of NO, N_2O, and N_2 were identified by their mass spectra and peak areas were quantified relative to ambient air. As ^{15}N cultures were grown in isotopically pure $^{15}NO_3^{2-}$, complete denitrification should result in accumulation of $^{30}N_2$ at a 1:2 ratio relative to nitrate consumption. $^{30}N_2$ should only accumulate if eNOR is functioning as part of a complete denitrification pathway. If eNOR does not function effectively as a nitric oxide reductase, then ^{15}NO should be seen to accumulate. Instead, only the $^{30}N_2$ peak was observed, indicating the eNOR

functioned effectively as a nitric oxide reductase for denitrification. Over the course of incubations, $^{30}N_2$ was seen to accumulate to more than 50x background. ^{14}N samples showed no significant accumulation of $^{30}N_2$ above background, confirming that the $^{30}N_2$ in ^{15}N samples was due to denitrification of labeled nitrate. NO was not seen to accumulate in any of the cultures. These results demonstrate that eNOR is a functional nitric oxide reductase and can be used as part of a complete denitrification pathway.

Purification of eNOR

The culture of *Rhodothermus marinus*, once harvested, was re-suspended in 100 mM Tris-HCl, pH 8 with 10 mM $MgCl_2$ and 50 µg/ml DNase, using a Bamix homogenizer. The resulting solution was spun down at 42000 rpm in a Beckman Ultracentrifuge. The membrane pellet was collected and re-suspended in 20 mM Tris-HCl, pH 7.5, 1 % CHAPS (Affymetrix) to a final concentration of 40-50 mg/ml. The solution was stirred at 4 °C for 1 hr. In this step a lot of peripheral membrane proteins appear to be solubilized and the remaining protein is pelleted by spinning down at 42000 rpm for 1 hr. The remaining pellet is then solubilized in 20 mM Tris-HCl, pH 7.5, 1 % DDM (Affymetrix) at a final protein concentration of around 5-10 mg/ml. The DDM solubilized fraction was once again centrifuged at 42000 rpm to pellet down protein that was not solubilized.

The solubilized protein was then loaded on a DEAE CL-6B (Sigma) column, pre-equilibrated in 20 mM Tris-HCl, pH 7.5, 0.05 % DDM, and subjected to a linear gradient spanning from 0 to 500 mM NaCl. The fraction containing the eNOR, identified using a peak at 591nm, corresponding to the peak of cytochrome 'a1' in *Magnetospirillum magnetotacticum*[1], eluted at around 200 mM salt. This fraction was then loaded on a Q Sepharose High Performance (GE Healthcare) column, pre-equilibrated with 20 mM Tris-HCl, pH 7,5, 0.05 % DDM and then eluted in a gradient from 0 to 1 M NaCl. The eNOR containing fraction was eluted at around 250 mM salt and the eluted

fraction was then loaded on a Chelating Sepharose (GE Healthcare) column, loaded with Cu^{2+} and equilibriated with 20 mM Tris-HCl, 500 mM NaCl, as previously described for cytochrome *caa*$_3$ from *Rhodothermus marinus*[2]. The eNOR fraction was once again loaded on a Q Sepharose High performance column, and a gradient was run between 0 and 300 mM NaCl at low flow rates (0.5 ml/min) and the first peak was found to be the eNOR.

Gel Electrophoresis

The purified eNOR was run on a Tris-Hepes 4-20 % acrylamide gel (NuSep) in the recommended Tris-Hepes-SDS running buffer at 120 V for ~1 hr. The protein was visualized and compared to the Precision Plus Protein™ Dual Color Standards (BIO-RAD).

UV-Visible Spectroscopy

All spectra were recorded on a HP Agilent 8453 UV-Vis spectrophotometer using a quartz cuvette from Starna Cells (No. 16.4-Q-10/Z15). Potassium Ferricyanide was used to obtain the oxidized spectrum, and dithionite was used to obtain the reduced spectrum.

Pyridine Hemochrome Assay

The hemes in eNOR were analyzed using a pyridine hemochrome assay[3]. A stock solution of 200 mM NaOH with 40 % pyridine was prepared. The stock solution was mixed 1:1 with the protein and an oxidized spectrum was obtained by adding 3 µl of 100 mM $K_3Fe(CN)_6$. A reduced spectrum was similarly prepared by adding a few crystals of sodium dithionite. The reduced minus oxidized spectrum was used to identify the heme co-factors

Heme extraction and HPLC Analysis

The hemes from eNOR were extracted and analyzed using an HPLC elution profile according to established protocols[4,5]. 50 µl of eNOR was mixed with 0.45 ml of acetone / HCl (19:1) and incubated for 20 minutes at room temperature after shaking. The mixture was centrifuged at 14,000 rpm for 2 minutes, followed by addition of 1 ml of ice cold water, and 0.3 ml of 100% ethyl acetate to the supernatant. The water/ethyl acetate mixture was vortexed and centrifuged again for 2 minutes. The ethyl acetate phase was recovered and concentrated using a speed vac.

The extracted hemes were analyzed using an Agilent 1290 Infinity LC attached to an Agilent 6230 TOF LC/MS equipment by separation using an Agilent Eclipse Plus C18 column (2.1x300 mm, 1.8 µm, 600 bar) and an acetonitrile (0.05 %TFA) / water (0.05 % TFA) gradient from 20 to 95 %.

NO reductase activity verification using GC

Anaerobic reaction conditions were set up in a 5 mL clear serum vial (Voigt Global Distribution, Inc) sealed with a 20 mm rubber stopper, by passing N_2 through 2 ml of 20 mM KPi, 0.05 % DDM, pH 7.5 with 1 mM TMPD, 5 mM Ascorbate. A control was performed by adding only 50 µM NO. Sample reactions were begun by adding eNOR to a final concentration of 100 nM. The reaction was incubated at 42 °C for half an hour before the headspace was injected into an HP Agilent 5890 Series GC, fitted with a TCD and ECD (SRI Instruments) for verification of N_2O production.

Turnover measurement using a Clark electrode

A sealed chamber fitted with an ISO-NO (World Precision Instruments) electrode was used for NO reductase activity measurements. 1 mM TMPD or 100 µM PMS and 4 mM Ascorbate were was added to 2 ml 50 mM Citrate, pH 6, 0.05 % DDM in the reaction chamber and all traces of oxygen were removed by passing water-saturated Argon for 20 minutes through the solution. This

is similar to the protocol described for cNOR from *Thermus thermophilus*[6]. The buffer system also contained an oxygen scavenging system constituting 100 nM catalase, 35 nM Glucose oxidase and 90 nM Glucose. The NO reduction traces were recorded using a Duo-18 (World Precision Instruments), and activities calculated from the slope of the traces.

LC/MS/MS analysis

Mass spectrometric analysis was conducted at the Protein Sciences Facility, Roy J Carver Biotechnology Center, University of Illinois, Urbana, IL 61801 using a Thermo LTQ Velos ETD pro mass spectrometer**.** For liquid samples, the samples were cleaned up using G-Biosciences Perfect Focus (St. Louis MO) prior to digestion with trypsin. Digestion was done using proteomics grade trypsin 1:20 (G-Biosciences, St. Louis, MO) and a CEM Discover Microwave Reactor (Mathews, NC) for 15 minutes at 55˚ C at 50 Watts. Digested peptides were extracted 3X using 50% acetonitrile containing 5% formic acid, pooled and dried using a Speedvac (Thermo Scientific). The dried peptides were suspended in 5% acetonitrile containing 0.1% formic acid and applied to LC/MS.

HPLC for the trypsin digested peptides was performed with a Thermo Fisher Dionex 3000 RSLCnano using Thermo Acclaim PepMap RSLC column (75 µm x 15 cm C-18, 2 µm, 100Å) and a Thermo Acclaim PepMap 100 Guard column (100 µm x 2 cm, C-18, 5 µm, 100Å), solvents were water containing 0.1% formic acid (A) and acetonitrile containing 0.1% formic acid (B) at a flow rate of 300 nanoliters per minute at 40° C. Gradient was from 100% A to 60% B in 60 minutes. The effluent from the UHPLC was infused directly into a Thermo LTQ Velos ETD Pro mass spectrometer.

Control and data acquisition of the mass spectrometer was done using Xcalibur 2.2 under data dependent acquisition mode, after an initial full scan, the top five most intense ions were subjected to MS/MS fragmentation by collision induced dissociation. The raw data were processed by Mascot Distiller (Matrix Sciences, London, UK) and then by Mascot version 2.4. The result was searched against NCBI NR Protein database.

Analysis of heme-copper oxygen reductase phylogeny and distribution in environmental datasets

We performed a large-scale analysis of heme-copper oxygen reductase (HCO) protein sequences in the NCBI and IMG databases with BLASTP using an e-value of $1e^{-3}$ to generate a database of HCO sequences that had at least some of the conserved amino acids previously identified in subunit I[7]. We then used the database of HCOs, filtered it with a sequence cut-off of 50% to generate the multiple sequence alignment, **MSA1**. A phylogenetic tree (**Figure 2**) was inferred using IQ-TREE 2[8] with the Dayhoff substitution matrix, Gamma model of rate heterogeneity and 1000 ultrafast bootstraps. Using the curated HMMs for each of the HCO family oxygen reductases[9], we probed release 202 of the Genome Taxonomy Database[10] for distribution of the NOR families – eNOR, bNOR, sNOR, nNOR, gNOR, cNOR and qNOR – in bacteria and archaea. Curated HMMs for the nitrate reductases (NapAB, NarGH), nitrite reductases (NirK, NirS) and nitrous oxide reductases (NosD and NosZ) were sourced from the HMMs database of MagicLamp[11].

Analysis of HCO distribution in various ecosystems were performed using the metagenomes in the IMG database. Approximately 2300 metagenomes were identified which were sourced from 44 environments identified by IMG. The number of different HCOs in each of these environments were extracted using BLASTP and query sequences that belong to each of the different HCO families.

References

1. Tamegai, H., Yamanaka, T. & Fukumori, Y. Purification and properties of a 'cytochrome a1'-like hemoprotein from a magnetotactic bacterium, Aquaspirillum magnetotacticum. *Biochim. Biophys. Acta BBA - Gen. Subj.* **1158**, 237–243 (1993).
2. Pereira, M. M. *et al.* The caa3 terminal oxidase of the thermohalophilic bacterium Rhodothermus marinus: a HiPIP:oxygen oxidoreductase lacking the key glutamate of the D-channel. *Biochim. Biophys. Acta BBA - Bioenerg.* **1413**, 1–13 (1999).
3. Berry, E. A. & Trumpower, B. L. Simultaneous determination of hemes a, b, and c from pyridine hemochrome spectra. *Anal. Biochem.* **161**, 1–15 (1987).
4. Sone, N. & Fujiwara, Y. Haem O can replace haem A in the active site of cytochrome c oxidase from thermophilic bacterium PS3. *FEBS Lett.* **288**, 154–158 (1991).
5. Reinhold-Hurek, B. & Zhulin, I. B. Terminal oxidases of Azoarcus sp. BH72, a strictly respiratory diazotroph. *FEBS Lett.* **404**, 143–147 (1997).
6. Schurig-Briccio, L. A. *et al.* Characterization of the nitric oxide reductase from Thermus thermophilus. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 12613–12618 (2013).
7. Hemp, J. & Gennis, R. B. Diversity of the Heme–Copper Superfamily in Archaea: Insights from Genomics and Structural Modeling. *Bioenergetics* 1–31 (2008).

8. Minh, B. Q. *et al.* IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
9. Murali, R., Hemp, J., Orphan, V. & Bisk, Y. FIND: Identifying Functionally and Structurally Important Features in Protein Sequences with Deep Neural Networks. *bioRxiv* 592808 (2019) doi:10.1101/592808.
10. Parks, D. H. *et al.* A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **36**, 996–1004 (2018).
11. Garber, A. *MagicLamp: toolkit for annotation of 'omics datasets using curated HMM sets.* (2020).