



HAL
open science

Resolving phylogenetic and biochemical barriers to functional expression of heterologous iron-sulphur cluster enzymes

Hélène Shomar, Pierre Simon Garcia, Elena Fernández-Fueyo, Francesca D'angelo, Martin Pelosse, Rita Rebelo Manuel, Ferhat Büke, Siyi Liu, Niels van den Broek, Nicolas Duraffourg, et al.

► **To cite this version:**

Hélène Shomar, Pierre Simon Garcia, Elena Fernández-Fueyo, Francesca D'angelo, Martin Pelosse, et al.. Resolving phylogenetic and biochemical barriers to functional expression of heterologous iron-sulphur cluster enzymes. 2021. hal-03380538

HAL Id: hal-03380538

<https://hal.science/hal-03380538>

Preprint submitted on 15 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

1 **Resolving phylogenetic and biochemical barriers to functional** 2 **expression of heterologous iron-sulphur cluster enzymes**

3 Helena Shomar^{a*}, Pierre Simon Garcia^{b,c*}, Elena Fernández-Fueyo^{a,1*}, Francesca D'Angelo^{c*},
4 Martin Pelosse^{d,2}, Rita Rebelo Manuel^{a,3}, Ferhat Büke^a, Siyi Liu^e, Niels van den Broek^{a,4}, Nicolas
5 Duraffourg^d, Carol de Ram^f, Martin Pabst^f, Simonetta Gribaldo^b, Beatrice Py^e, Sandrine Ollagnier
6 de Choudens^d, Gregory Bokinsky^{a***†}, Frédéric Barras^{c***†}

7 *These authors contributed equally and are listed in reverse alphabetical order.

8 **These senior authors contributed equally.

9 **Affiliations:** ^a Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of
10 Technology, Delft, The Netherlands; ^b Department of Microbiology, UMR Institut Pasteur-CNRS
11 2001, Unit Evolutionary Biology of the Microbial Cell, Paris, France; ^c Department of
12 Microbiology, UMR Institut Pasteur-CNRS 2001, Unit Stress Adaptation and Metabolism of
13 Enterobacteria, Paris, France; ^d Univ. Grenoble Alpes, CNRS, CEA, IRIG, Laboratoire de Chimie
14 et Biologie des Métaux, Grenoble, France; ^e Aix-Marseille Université-CNRS, Laboratoire de
15 Chimie Bactérienne, Institut de Microbiologie de la Méditerranée, Institut Microbiologie
16 Bioénergies Biotechnologie, Marseille, France; ^f Department of Biotechnology, Delft University of
17 Technology, Delft, The Netherlands

18 **Current addresses:** ¹ Institute for Translational Vaccinology (Intravacc), Process Development
19 Bacterial Vaccines, Bilthoven, The Netherlands; ² EMBL Grenoble, Grenoble, France; ³ Instituto
20 de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Oeiras,
21 Portugal; ⁴ Vaccine Process & Analytical Development, Janssen Vaccines & Prevention B.V.,
22 Leiden, The Netherlands

23 †Materials and correspondence: Frédéric Barras frederic.barras@pasteur.fr, Gregory Bokinsky
24 g.e.bokinsky@tudelft.nl

25

26

27 **Abstract**

28 Many of the most promising applications of synthetic biology, including engineering of
29 microbes for renewable chemical production, relies upon the ability of genetically-tractable hosts
30 to express heterologous enzymes from foreign species. While countless methods for facilitating
31 heterologous enzyme expression have been developed, comparable tools for facilitating
32 heterologous enzyme activity are generally lacking. Such tools are needed to fully exploit the
33 biosynthetic potential of the natural world. Here, using the model bacterium *Escherichia coli*, we
34 investigate why iron-sulphur (Fe-S) enzymes are often inactive when heterologously expressed.
35 By applying a simple growth complementation assay with collections of Fe-S enzyme orthologs
36 from a wide range of prokaryotic diversity, we uncover a striking correlation between phylogenetic
37 distance and probability of functional expression. Moreover, co-expression of a heterologous Fe-
38 S biogenesis pathway increases the phylogenetic range of orthologs that can be functionally
39 expressed. On the other hand, we find that heterologous Fe-S enzymes that require specific
40 electron carrier proteins within their natural host are rarely functionally expressed unless their
41 specific reducing partners are identified and co-expressed. We demonstrate *in vitro* that such
42 selectivity in part derives from a need for low-potential electron donors. Our results clarify how
43 phylogenetic distance and electron transfer biochemistry each separately impact functional
44 heterologous expression and provide insight into how these barriers can be overcome for
45 successful microbial engineering involving Fe-S enzymes.

46

47 Introduction

48 Microbes acquire new phenotypes *via* horizontal gene transfer, a process that reshapes microbial
49 evolution and ecology. Gene transfer is also an everyday laboratory technique that enables
50 expression of heterologous proteins and engineered biosynthetic pathways. In natural systems,
51 obstacles against foreign gene expression, which include restriction-modification systems and
52 differences in codon usage, can impede horizontal gene transfer from influencing host phenotype.
53 In the lab, such genetic obstacles are routinely avoided to ensure heterologous protein expression,
54 e.g. by using strains that lack genome defence systems or by optimizing codon usage. However,
55 heterologous proteins must not only be expressed but also must retain their activity to affect the
56 host phenotype or to function as part of an engineered pathway. Unlike obstacles to heterologous
57 protein expression, the obstacles that prevent heterologous protein activity are rarely studied¹.
58 One enzyme class that is particularly prone to inactivity in heterologous hosts is iron-sulphur
59 cluster (Fe-S) enzymes²⁻⁵. Fe-S enzymes use bound iron and sulphur metalclusters to catalyse
60 a variety of essential reactions, including many that are relevant to biotechnology such as nitrogen
61 reduction, dehydration, sulphur insertion, and methyl transfer⁶. Identifying obstacles to functional
62 heterologous expression of Fe-S proteins is therefore an important challenge.

63 Fe-S enzymes are matured by dedicated Fe-S biogenesis pathways and proteins
64 **(Figure 1A)**. If a heterologous Fe-S enzyme cannot obtain Fe-S clusters from its host, enzyme
65 activity is impaired. How Fe-S enzymes obtain clusters from Fe-S biogenesis proteins continues
66 to be intensively investigated in both prokaryotic and eukaryotic models^{7,8}. Two Fe-S biogenesis
67 pathways, ISC and SUF, catalyse assembly and delivery of Fe-S clusters to many apo-target
68 proteins in prokaryotes (e.g. more than 150 in *Escherichia coli*). Common to both ISC and SUF is
69 the use of cysteine desulphurase enzymes (IscS or SufSE) to obtain sulphur from L-cysteine and
70 scaffold proteins to assemble Fe-S clusters (IscU or SufBCD). Fe-S cluster carrier proteins of the
71 A-type (IscA, SufA, ErpA) Nfu-type (NfuA) or ApbC-type (Mrp) subsequently deliver clusters to

72 apo-targets^{9,10}. Moreover, some Fe-S require additional factors to sustain catalytic turnover,
73 particularly electron carrier proteins¹¹. This requirement may present an additional barrier to
74 heterologous expression.

75 The full biotechnological potential of Fe-S enzymes cannot be harnessed until we
76 understand why many of them lose activity when heterologously expressed. A lack of systematic
77 data prevents any meaningful prediction on whether a heterologous Fe-S enzyme will remain
78 active within a host. Defining the limits of horizontal transfer of Fe-S enzymes between species is
79 therefore needed to identify the specific factors that impede functional heterologous expression.
80 Here, we combine bioinformatics with an *in vivo* platform to map the transferability of several
81 biotechnologically-interesting Fe-S enzymes between prokaryotes. We also characterize at a
82 molecular level the requirements for sustaining the heterologous activity of an Fe-S enzyme from
83 the radical S-adenosylmethionine (rSAM) methyltransferase family.

84 Results

85 Heterologous Fe-S enzymes are less likely to be functionally expressed than non-Fe-S 86 enzymes

87 We performed a preliminary screen to evaluate the ability of Fe-S enzymes to retain activity
88 when expressed in a heterologous host. We constructed four strains of *E. coli* MG1655 that each
89 lack a conditionally-essential Fe-S enzyme: NadA (quinolinate synthase), IspG (4-hydroxy-3-
90 methylbut-2-enyl-diphosphate synthase), BioB (biotin synthase), and ThiC (4-amino-5-
91 hydroxymethyl-2-methylpyrimidine phosphate synthase). These proteins were chosen because
92 they are involved in metabolic pathways for valuable compounds (e.g. vitamins, biofuels, and
93 fragrances). Orthologs of these enzymes were identified in genomes belonging to a diverse range
94 of bacterial phyla (e.g. *Actinobacteria*, *Firmicutes*, *Proteobacteria*, *Cyanobacteria*) for a
95 preliminary survey (**Table S1**). Each orthologous gene was codon-optimized, cloned into a low-
96 copy expression vector, and transformed into the corresponding *E. coli* knockout strain. The
97 activities of the heterologous enzymes were tested using a complementation assay: growth in
98 selective media indicates that the heterologous enzyme is functional when expressed in *E. coli*
99 (**Figure 1B, 1C**). The majority of Fe-S enzyme orthologs (115 out of 135) failed to complement
100 growth (**Figure 1D, Table S1**). For comparison, we tested orthologs of a non-Fe-S enzyme (IspD,
101 2-C-methyl-D-erythritol 4-phosphate cytidyltransferase) for functional expression in *E. coli*. In
102 contrast with the Fe-S enzymes tested, the majority of IspD orthologs (38 out of 45) complemented
103 growth of *E. coli* Δ *ispD*, including many orthologs from species whose IspG orthologs failed to
104 complement growth of *E. coli* Δ *ispG* (**Table S1**). Of note, average sequence identity percentage
105 of the heterologous IspG and IspD orthologs with the *E. coli* orthologs were 45% and 36%,
106 respectively. Altogether, these observations suggest that Fe-S enzymes are far more likely to lose
107 activity within a heterologous host as compared to enzymes whose activities do not require Fe-S
108 cluster or other post-translational modifications.

109 **Correlating compatibility of heterologous enzymes with phylogenetic distance from host**
110 **ortholog**

111 We next sought to explore in detail the factors that determine whether a Fe-S enzyme
112 retains its activity within a heterologous host. We hypothesized that heterologous Fe-S enzymes
113 whose sequences are less similar to the host *E. coli* ortholog might be less likely to complement
114 growth. To test this hypothesis, we focused on orthologs of NadA and IspG. We first built a protein
115 database that is representative of known prokaryotic diversity, comprising 248 prokaryotic
116 (Archaea and Bacteria) proteomes corresponding to 48 phyla, 96 classes, and 218 orders (**Table**
117 **S2**). This database contains 65 proteomes that were used for preliminary screen. We identified
118 192 and 196 orthologs of NadA and IspG, respectively (**Table S3**), which were used to infer a
119 Maximum Likelihood phylogeny (**Figure 2A, D**). Next, we used a clustering approach based on
120 phylogenetic distances to select representative sequences to be experimentally tested, which
121 yielded 47 sequences for NadA and IspG each (**Tables S3, S4**). The distribution of patristic
122 distances of the selected orthologs from the *E. coli* NadA and IspG sequences closely matched
123 with the patristic distance distribution obtained from all enzyme homologs, indicating that the
124 selected orthologs are representative of the entire database of known orthologs (**Figure 2B, E**).
125 Furthermore, a two-sided Wilcoxon test between full and sampled distances from *E. coli* rejected
126 the hypothesis of significant differences between the selected set of orthologs (47 both) and the
127 set of all known orthologs (191 for NadA and 195 for IspG, p -value = 0.4714 for NadA and p -value
128 = 0.4739 for IspG). Altogether, these results indicate that the set of orthologs tested experimentally
129 is a representative sample of the phylogenetic diversity of NadA and IspG sequences across
130 prokaryotes.

131 Sequences from our 47 phylogenetically-representative sets of IspG and NadA were each
132 then codon-optimized for expression in *E. coli* and tested using our complementation assay.
133 Among the NadA orthologs representative of phylogenetic diversity, 14 out of 47 orthologs

134 complemented growth of *E. coli* Δ nadA, indicating functional expression (**Table 1**). Consistent with
135 our hypothesis, NadA orthologs that were functionally expressed in *E. coli* exhibited a low patristic
136 distance from the *E. coli* NadA ortholog (**Figure 2C**). The high correlation factor (η^2) between
137 heterologous expression and patristic distance ($\eta^2 = 0.596$) suggests that phylogenetic proximity
138 is a useful predictor of whether a NadA ortholog can be functionally expressed in *E. coli*.

139 Among the 47 IspG orthologs selected from a representative set, only 8 complemented
140 growth of *E. coli* Δ ispG (**Table 2**). The sequences of functionally-expressed IspG orthologs did not
141 group together in the IspG phylogeny but were instead patchily dispersed across the entire
142 phylogenetic distribution of sequences (**Figure 2F**). Accordingly, no correlation between functional
143 expression and patristic distance from *E. coli* was observed ($\eta^2 = 0.003$). Altogether, these results
144 indicate that the ability of heterologous IspG orthologs to maintain activity in *E. coli* does not
145 correlate with phylogenetic proximity, in contrast with NadA orthologs.

146 We sought to confirm that the inactive NadA and IspG orthologs used for the representative
147 distribution were expressed in our knockout strains. Because even some of the active orthologs
148 could not be detected by SDS-PAGE (**Table S5**), we used a mass spectroscopy-based shotgun
149 proteomics approach to determine whether 7 inactive NadA orthologs and the 16 inactive IspG
150 orthologs could be detected when expressed in permissive media. We detected all 7 of the inactive
151 NadA orthologs and 10 of the 16 inactive IspG orthologs (**Table S5**). Given that *E. coli* IspG is
152 largely insoluble when overexpressed¹², and that insoluble proteins are often difficult to detect
153 using mass spectroscopy, we suspect that insolubility may contribute to inactivation of
154 heterologous IspG orthologs and potentially other Fe-S enzymes. Nevertheless, it is very unlikely
155 that a correlation between phylogenetic proximity and heterologous activity is obscured by poor
156 expression of IspG orthologs.

157 **Activity of NadA orthologs depends upon Fe-S biogenesis pathway components**

158 We next tested approaches for restoring function of the inactive NadA orthologs. Our
159 finding that the activity of heterologous NadA orthologs correlates with phylogenetic proximity to
160 *E. coli* NadA suggests that co-expression of a heterologous Fe-S pathway might re-activate the
161 heterologous NadA orthologs. We therefore cloned a heterologous Fe-S pathway (the *Bacillus*
162 *subtilis* SUF operon) into a separate vector (pBbA5a), which was introduced into *E. coli* Δ nadA
163 carrying *B. subtilis* NadA (**Table S6**). Satisfyingly, co-expression of the *B. subtilis* SUF system
164 together with *B. subtilis* NadA fully recovered growth of *E. coli* Δ nadA (**Figure S1, Table 1**).
165 Expression of *B. subtilis* SUF also activated 8 out of 33 additional NadA orthologs (**Table 1**). By
166 comparison, overexpression of the *E. coli* suf operon activated only 3 of the 8 NadA orthologs
167 activated by *B. subtilis* SUF (*B. subtilis*, *Mycolicibacterium smegmatis*, and *Novosphingobium*
168 *stygium*) (**Table 1**). Thus, use of a heterologous Fe-S pathway may often broaden host
169 compatibility with certain heterologous Fe-S enzymes.

170 **Compatibility of IspG orthologs is limited by need for taxa-specific electron transfer** 171 **proteins**

172 While co-expression of either *B. subtilis* or *E. coli* SUF pathways did not recover any
173 inactive IspG ortholog tested, several heterologous IspG orthologs that did not complement *E. coli*
174 Δ ispG were surprisingly able to obtain Fe-S clusters *in vitro* when incubated with the *E. coli* Fe-S
175 transfer protein ErpA (**Figure S2**). These observations suggest that the inactivity of IspG orthologs
176 may be unrelated to Fe-S transfer. The Fe-S cluster of *E. coli* IspG requires electron inputs from
177 a electron carrier protein, flavodoxin 1 (FldA), to sustain catalysis¹¹. Although many enzymes that
178 require electron carrier proteins are able to obtain electrons from a variety of carrier proteins^{13,14},
179 the activity of *E. coli* IspG cannot be supported by *E. coli* flavodoxin 2 (FldB)¹⁵. Therefore we
180 suspected that the inactivity of heterologous IspG orthologs reflects a requirement for a specific
181 electron carrier protein from the original hosts. Specific electron carriers that support activity of
182 *Synechocystis* and *B. subtilis* IspG orthologs have been previously identified. For example, the

183 IspG ortholog from the cyanobacterium *Synechocystis* is known to be reduced by the ferredoxin
184 PetF¹⁶. We therefore cloned PetF into a pBbA5a vector and co-transformed into *E. coli* Δ IspG
185 together with a vector bearing *Synechocystis* IspG. Co-expression of PetF with *Synechocystis*
186 IspG indeed complemented growth of *E. coli* Δ IspG (**Table 2**). Similarly, co-expression of *B.*
187 *subtilis* flavodoxins observed to activate *B. subtilis* IspG, YkuP and YkuN¹⁷, allowed
188 complementation of *E. coli* Δ IspG.

189 The specific electron carriers needed to activate the remaining IspG orthologs tested are
190 not known. Furthermore, while prokaryotic genomes usually encode multiple electron carrier
191 proteins, no reliable bioinformatics method for identifying the carrier proteins able to support IspG
192 activity currently exists. Therefore, we first simultaneously co-expressed multiple electron carrier
193 proteins taken from 27 additional microorganisms together with their cognate IspG orthologs. Of
194 the 27 inactive IspG orthologs tested, 6 were re-activated by cognate electron carriers identified
195 within their original host (**Table S7**). We individually tested the ability of electron carrier proteins
196 from *Rhodobacter sphaeroides*, *Synechocystis*, and *Streptomyces cattleya* to support the activity
197 of their corresponding IspG ortholog. Strikingly, only 1 of the 5 electron carriers tested from *R.*
198 *sphaeroides* could activate *R. sphaeroides* IspG. IspG orthologs from *Synechocystis* and *S.*
199 *cattleya* IspG were also selective for specific carriers taken from their cognate hosts (**Table S6**).

200 We next explored the cross-species compatibility of IspG orthologs with heterologous
201 electron transfer proteins obtained from different species (**Table 2 and S7**). *Synechocystis* PetF
202 activated nearly every *Cyanobacteria* IspG ortholog tested but only a small number of IspG
203 orthologs outside the *Cyanobacteria* phylum. The individual ferredoxins from *S. cattleya* and *R.*
204 *sphaeroides* further activated additional IspG orthologs from the *Actinobacteria* and
205 *Proteobacteria*, respectively, but very few IspG orthologs outside their respective taxa. Altogether,
206 19 out of 53 IspG orthologs tested were re-activated by co-expression of either cognate or

207 heterologous electron transfer proteins. Collectively, these results indicate that the requirement
208 for specific electron transfer proteins can prevent functional expression of heterologous IspG.

209 **A specific double-cluster ferredoxin enhances the activity of the Fe-S-dependent** 210 **methyltransferase TsrM**

211 The molecular basis underlying the selectivity of Fe-S enzymes for specific electron
212 transfer proteins is poorly understood. We therefore explored the electron transfer protein
213 selectivity of a biotechnologically-relevant Fe-S enzyme¹⁸. rSAM methyltransferase enzymes rely
214 upon Fe-S clusters to catalyse biosynthesis of valuable natural products, including antibiotics^{19,20}.
215 The cobalamin-dependent rSAM methyltransferase TsrM, an enzyme from the thiostrepton
216 synthesis pathway of *Streptomyces laurentii*, uses its cobalamin cofactor as an intermediate
217 carrier to transfer a methyl group to C2 of tryptophan²¹. Although its catalytic cycle does not require
218 an electron donor for each turnover event, TsrM requires one electron to reduce its cobalamin
219 cofactor to cob(I)alamin, which reacts with SAM to form MeCbl (**Figure 3A**). TsrM can also be
220 inactivated by adventitious oxidation of the bound cobalamin factor from cob(I)alamin to
221 cob(II)alamin. Recovering the catalytically-active cob(I)alamin form requires reduction by an
222 electron transfer protein²².

223 *E. coli* NCM3722 co-expressing *S. laurentii* TsrM together with the *E. coli* cobalamin import
224 pathway (*btu* operon expressed from a plasmid²³) produced 3 ±2 mg/L 2-methyltryptophan (**Figure**
225 **3B**). We sought to determine whether *in vivo* TsrM activity could be improved using *Streptomyces*
226 [4Fe-4S] ferredoxins, as [4Fe-4S] ferredoxins exhibit low reducing potentials (-700 – -300 mV)²⁴
227 that are expected to match the low reduction potential of cob(II)alamin/cob(I)alamin (< -500 mV
228 vs. NHE)^{25,26}. We used three [4Fe-4S] ferredoxins (here designated Fdx1, Fdx2, and Fdx3), two
229 of which (Fdx1 and Fdx3) were able to activate *S. cattleya* IspG (**Table S6**). Simultaneous co-
230 expression of all three [4Fe-4S] ferredoxins together with TsrM increased 2-methyltryptophan
231 titers to 36 ±16 mg/L. By co-expressing each ferredoxin individually with TsrM, we determined that

232 Fdx1 alone improves activity of TsrM (45 ±7 mg/L 2-methyltryptophan), while expression of Fdx2
233 and Fdx3 do not increase 2-methyltryptophan titers (**Figure 3B**).

234 We further investigated the molecular basis of ferredoxin selectivity by TsrM using *in vitro*
235 experiments with purified proteins. Consistent with previous observations, purified TsrM contains
236 0.9 cobalamin and 3.5 Fe per monomer (**Figure S3**). The iron content of the ferredoxins after
237 anaerobic purification or reconstitution with iron and sulphur salts (7.4, 2.9, and 2.8 Fe atoms *per*
238 monomer for Fdx1, Fdx2, and Fdx3 respectively) is consistent with their corresponding
239 annotations as double-cluster and single-cluster ferredoxins. We first used bio-layer interferometry
240 (BLI) to determine whether TsrM associates preferentially with Fdx1. While we were able to detect
241 association between Fdx1 and TsrM (dissociation constant $K_D = 40 \mu\text{M}$), we found that Fdx2 and
242 Fdx3 did not measurably associate with TsrM (**Figure 3C**). We also compared the redox potentials
243 of each ferredoxin against TsrM using voltammetry. The reduction potentials determined for the
244 [4Fe-4S] cluster of TsrM matches with cob(II)alamin/cob(I)alamin (-730 mV vs Ag/AgCl), in
245 agreement with previous reports¹⁵. The voltammogram of Fdx1 indicated two redox transitions
246 associated with each [4Fe-4S] cluster (-750 mV and -650 mV vs Ag/AgCl), the lower of which
247 supports reduction of TsrM metallic cofactors ([4Fe-4S]²⁺ to [4Fe-4S]¹⁺ cluster) and
248 (cob(II)alamin to cob(I)alamin). The redox potentials of 4Fe-4S Fdx2 and [4Fe-4S] Fdx3 are each
249 too high for favourable reduction of TsrM metallic cofactors (-620 mV and -585 mV respectively
250 vs Ag/AgCl) (**Figure 3D**). Therefore, both binding specificity and reduction potential determine
251 productive interactions between electron transport proteins and TsrM, which thus constrains its
252 functional heterologous expression. We note that the TsrM-Fdx1 dissociation constant is
253 consistent with a transient interaction characteristic of an electron transfer process.

254

255 Discussion

256 Our study has used Fe-S enzymes to map the interspecies versatility of Fe-S biogenesis
257 pathways. In doing so, we have distinguished specific barriers that limit functional transfer of
258 heterologous Fe-S enzymes. Most strikingly, we found that Fe-S pathways are highly compatible
259 with certain heterologous Fe-S enzymes, except at large phylogenetic distances from the host.
260 This is demonstrated by the activity of nearly one third of the heterologous NadA orthologs tested
261 and is also supported by our observation that many IspG orthologs and the rSAM enzyme TsrM
262 require only a specific electron transfer protein to achieve activity in *E. coli*. The versatility of Fe-
263 S pathways is also illustrated by our finding that many of the NadA orthologs taken from organisms
264 that lack an ISC pathway are activated even by the endogenous *E. coli* ISC expressed in our
265 conditions. This versatility, which is likely a consequence of the need for Fe-S pathways to provide
266 Fe-S clusters to many diverse Fe-S enzymes present in the cell, also appears to facilitate
267 functional gene transfer across moderate phylogenetic distances.

268 While approximately one-third of the NadA orthologs complemented *E. coli* Δ nadA and
269 were thus compatible with *E. coli* ISC, several NadA orthologs exhibited activity only when co-
270 expressed with either *E. coli* or *B. subtilis* SUF pathways. Interestingly, several of the NadA
271 orthologs were selectively activated by *B. subtilis* SUF. This may be due to structural differences
272 between components of the *E. coli* and *B. subtilis* SUF systems that effectively restrict Fe-S cluster
273 transfer to orthologs that are closely-related or otherwise similar²⁷. In particular, *B. subtilis* SUF
274 includes two proteins that have been proposed as possible Fe-S scaffolds, SufB and SufU,
275 whereas the *E. coli* SUF system includes only SufB. However, co-expressing *B. subtilis* SufU with
276 *E. coli* SUF failed to activate additional orthologs (data not shown). Interestingly, all NadA
277 orthologs that successfully complemented (with or without SUF expression) are phylogenetically
278 grouped (**Figure 2A**), implying that specific structural features might distinguish NadA orthologs
279 that can be functionally expressed within *E. coli* from NadA orthologs inactive in *E. coli*. The NadA

280 phylogeny is not congruent with known taxonomy (*i.e.* a mix of *Proteobacteria*, *Actinobacteria*,
281 *Cyanobacteria* and *Archaea*, implying that the sources of NadA orthologs that are compatible with
282 *E. coli* have a complex history of horizontal gene transfer.

283 We were surprised to find that electron carrier specificity may present a more stringent
284 barrier to functional expression of IspG orthologs than the requirement for Fe-S cluster transfer.
285 Although other redox-dependent enzyme classes (*e.g.* cytochrome P450 enzymes) are known to
286 accept a limited range of electron carriers, many redox-dependent enzymes exhibit activity with
287 multiple electron transfer proteins, and vice-versa. For instance, the *E. coli* flavodoxin FldA
288 delivers electrons to BioB, IspG, and IspH²⁸. Three *B. subtilis* electron transfer proteins (including
289 YkuN and YkuP) are able to support activity of *B. subtilis* acyl lipid desaturase¹³ and five
290 *Thermotoga maritima* ferredoxins are capable of supporting the rSAM enzyme MiaB¹⁴. Despite
291 the fact that some redox enzymes might use electron transfer proteins interchangeably,
292 prokaryotic genomes retain multiple electron transfer proteins that often show an extremely high
293 level of sequence divergence, structural variation, and reduction potential²⁹. The retention of
294 multiple electron carriers indicates that each carrier is not compatible with all redox partners,
295 despite some interchangeability. We observed this selectivity as well: only one electron transfer
296 protein from *R. sphaeroides* out of five tested activated *R. sphaeroides* IspG, and only two out of
297 five *S. cattleya* electron transfer proteins activated *S. cattleya* IspG (**Table S6**). Retaining an array
298 of carriers enables species to support enzymes such as IspG or TsrM that specifically require low-
299 potential electrons. In these cases, it is the need for low-potential reducing equivalents, rather than
300 possession of an Fe-S cluster, that establishes a stringent barrier to functional gene transfer.

301 What do our results mean for using Fe-S enzymes in the context of engineered
302 biosynthetic pathways? Overexpression of the host ISC and SUF pathways has been shown to
303 improve the yield of properly-matured Fe-S enzymes³⁰, likely because overexpression of Fe-S
304 enzymes overtaxes the ability of the host Fe-S biogenesis systems to provide Fe-S clusters.

305 However, we find that this approach may not reliably activate heterologous Fe-S enzymes.
306 Instead, some may require co-expression with a heterologous Fe-S biogenesis pathway, as we
307 observed for several NadA orthologs. If the Fe-S enzyme requires electron transfer, a specific
308 electron transfer protein from the original host may be required. Given that prokaryotic genomes
309 possess multiple electron carriers, identifying suitable electron transfer proteins may prove
310 difficult. Suitable candidates may be identified based upon genome context of the Fe-S enzyme
311 or by knowledge of its biochemical properties. In the case of TsrM, our awareness of the low
312 reducing potential of cobalamin guided our selection to [4Fe-4S]-type ferredoxins. Some Fe-S
313 enzymes require protein cofactors in addition to electron carriers. For instance, the rSAM enzyme
314 lipoyl synthase (LipA) requires the *E. coli* Fe-S carrier protein NfuA to sustain catalytic turnovers³¹.
315 In such cases, additional proteins from the original host may also be required to sustain
316 heterologous Fe-S enzyme activity.

317

318 **Acknowledgements**

319 We thank members of the Bokinsky Lab and the Barras Unit for discussions throughout
320 the work, and Mohamed Atta and Julien Pérard for their assistance with biochemistry experiments.
321 This project (IRONPLUGNPLAY) has received funding from the European Union's Horizon 2020
322 research and innovation programme under grant 722361 and was supported by grants from
323 Agence Nationale Recherche (ANR) ERA-Institut Pasteur and the ANR-10-LABX-62-IBEID. The
324 work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of
325 Science User Facility, is supported under Contract No. DE-AC02-05CH11231. SL was supported
326 by the China Scholarship Council.

327 **Author Contributions**

328 FD, EF, HS, and SL performed all cloning and complementation experiments. PG
329 performed bioinformatics analysis. EF and RR engineered 2-methyltryptophan production. NvdB
330 quantified 2-methyltryptophan by LCMS. F. Büke, CR, and M. Pabst measured expression of
331 proteins by untargeted proteomics. M. Pelosse performed biochemical assays. ND performed
332 electrochemistry experiments. BP, SG, SO-C, F. Barras, and GB supervised the project. All
333 authors contributed to the final manuscript.

334 **Materials and Methods**

335 **Selection of sequences and bioinformatics analysis**

336 Orthologs of IspG, NadA, BioB, and ThiC selected in the preliminary screen were found in
337 a variety of sequenced prokaryotic genomes using homology searches guided by *E. coli*
338 sequences. To assemble the phylogenetically-representative distribution of NadA and IspG
339 orthologs that was assembled following the preliminary screen, a systematic approach was used.
340 First, a protein database was built from 248 representative prokaryotic proteomes (**Table S2**)

341 (including 65 proteomes that were used for preliminary screen) gathered from NCBI FTP
342 (<ftp://ftp.ncbi.nlm.nih.gov/genomes/all>), selecting around one genome *per* order (27 Archaea and
343 221 Bacteria) by using NCBI taxonomy and privileging complete proteomes. Homologues of NadA
344 and IspG were first identified by BLASTp v2.8.1+³², using *E. coli* sequences as seeds
345 (AAC73837.1 and AAC75568.1, respectively). Sequences associated to an *E*-value lower than
346 10^{-4} were aligned with MAFFT v7.419³³ and used to build HMM profiles using hmmbuild from the
347 HMMER v3.2.1 package³⁴. These HMM profiles were then used to request the database using
348 hmmsearch, and the sequences presenting an *E*-value lower than 10^{-2} were retrieved and aligned.
349 Alignments were manually curated using Aliview v1.25³⁵ and trimmed using BMGE v1.1³⁶ with the
350 substitution matrix BLOSUM30.

351 The best suited model for each alignment was selected using ModelFinder implemented
352 in IQ-TREE v1.6.10³⁷, according to the Bayesian Information Criteria (BIC) (LG+I+G4 for both
353 NadA and IspG). Maximum Likelihood trees were inferred using PhyML v3.1³⁸ and 100 bootstrap
354 replicates performed to assess the robustness of the branches by using the transfer bootstrap
355 value using BOOSTER v1.0³⁹. For both NadA and IspG, the number of copies *per* genome was in
356 the extreme majority equal to one. All homologues are thus considered as orthologues.

357 In order to select representative NadA and IspG sequences to be experimentally tested,
358 the ML trees were split into clusters using TreeCluster v1.0⁴⁰ with a threshold of 1.75, leading to
359 15 sequence clusters for NadA and 17 for IspG. Then, each cluster was further split into sub-
360 clusters using hierarchical clustering implemented in scipy.cluster.hierarchy python library⁴¹. The
361 number of sub-clusters for each cluster was set to 20% of the sequences within each cluster. One
362 sequence per sub-cluster was finally selected, privileging reference strains and already tested
363 sequences, leading to 47 sequences for NadA and 47 for IspG. The representativeness was
364 checked by comparing the shape of distribution of patristic distances between the entire NadA
365 and IspG trees and the selected sequences. The correlation between complementation data and

366 phylogenetic distance was calculated by the correlation ratio η^2 . Data for phylogenetic analysis
367 (containing protein database of the 248 prokaryotes, NadA/ IspG alignments and phylogenies and
368 python script used to select sequences by clustering approach) is available at
369 figshare.com/articles/dataset/Dangelo_et_al_Supplementary_data_zip/13664927.

370 **Bacterial strains, media and chemicals**

371 The plasmids, oligonucleotides, and *E. coli* strains used in this study are listed in **Table S8**. *E. coli*
372 Δ *ispG* was constructed by first integrating genes encoding the lower half of the mevalonate
373 pathway to enable isoprenoid synthesis from exogenously-provided mevalonate. Genes encoding
374 mevalonate kinase (MK), phosphomevalonate kinase (PMK) and mevalonate diphosphate
375 decarboxylase (PMD) from *Saccharomyces cerevisiae* control by the P_{trc} promoter was amplified
376 from pJBEI-2999⁴² and assembled by PCR to a FRT-flanked kanamycin resistance (KanR)
377 cassette from pKD13 and integrated into *intA* using homologous recombination⁴³. The KanR
378 cassette was subsequently removed using the plasmid pCP20. *nadA*, *bioB*, and *thiC* were
379 removed using homologous recombination in *E. coli* MG1655. Bacterial strains were routinely
380 grown in aeration at 37°C in Luria-Bertani broth (LB) or in M9 medium (M9) supplemented with
381 glucose (0.4%), CaCl₂ (100 μM) and MgSO₄ (1 mM). Solid media contained 1.5% agar.
382 Anhydrotetracycline (aTc) 100 ng/mL, isopropyl β-D-1-thiogalactopyranoside (IPTG) 250 μM,
383 mevalonate (MVA) 0.5 mM and nicotinic acid (NA) 10 μg/mL were added when indicated. When
384 required, antibiotics were added at the concentrations of 50 μg/ml kanamycin (Kan) and 100
385 μg/mL ampicillin (Amp). Mevalolactone (MVL) and nicotinic acid (NA) were purchased from
386 Sigma-Aldrich and resuspended in water at final concentrations of 1 M and 10 mg/mL,
387 respectively. To prepare MVA, an equal volume of 1 M KOH was added to 1 M MVL and incubated
388 at 37°C for 30 minutes.

389 **Plasmid construction for complementation assays**

390 All the genes encoding heterologous orthologs and electron carrier proteins (with the exception of
391 genes from *B. subtilis*) were codon-optimized for expression in *E. coli*, designed with strong
392 ribosome binding sites, synthesized, and cloned into the expression vector pBbS2k⁴⁴ by the Joint
393 Genome Institute (JGI) directly, or were obtained as gene fragments by Twist Bioscience (TWB)
394 (**Table S1 and S4**), assembled, and cloned into pBbS2k. DH5 α was used for all cloning steps.
395 Genes encoding electron carrier proteins were codon-optimized for expression in *E. coli*, designed
396 with strong ribosome binding sites, and assembled into the expression vector pBbA5a by JGI.

397

398 **Complementation assays**

399 *E. coli* knockout strains were transformed with pBbS2k plasmid encoding corresponding orthologs
400 and selected on permissive LB medium containing kanamycin. Plates supporting growth of *E. coli*
401 Δ *ispG* strains additionally contained 0.5 mM MVA. For liquid culture-based complementation
402 experiments, two individual colonies for each strain were inoculated into LB medium, grown for 7
403 hours at 37°C, and subsequently diluted 1:500 in M9 and added to M9 medium containing Kan
404 and aTc. For reactivation experiments of inactive NadA orthologs, an *E. coli* Δ *nadA* strain bearing
405 a plasmid encoding the *B. subtilis* *SUF* operon (pBbA5a-*sufCDSUB*) or the plasmid encoding the
406 *E. coli* *SUF* operon (pBbA5a-*sufABCDSE*) was transformed with plasmids encoding orthologs and
407 grown as described with the addition of 250 μ M IPTG to induce expression of *SUF* (Table S6, S8).
408 For reactivation experiments of inactive NadA orthologs, an *E. coli* Δ *ispG* strain bearing a plasmid
409 encoding *IspG* orthologs was transformed with pBbA5a plasmids encoding electron transfer
410 proteins. All growth curves have been performed in 96-wells microplates and cell density (OD₆₀₀)
411 was recorded by using an automated Spark 10M luminometer-spectrophotometer (Tecan) for 18
412 hours (NadA experiments) or 12 hours (*IspG* experiments), every 30 minutes at 37°C in shaking
413 condition. To test functionality of ThiC and BioB orthologs, *E. coli* Δ *bioB* and Δ *thiC* strains were
414 transformed with pBbS2k plasmid encoding corresponding orthologs. Kanamycin resistant

415 colonies were streaked on LB selective plates. The complementation test was performed by
416 streaking on solid M9 glucose M9 medium supplemented with kanamycin and aTc. Plates were
417 incubated at 37°C and growth was scored after 16 and 40 hours.

418

419 **SDS-PAGE test for protein expression**

420 *ΔnadA* strains containing the pBbS2k plasmids and carrying NadA orthologs were grown in LB,
421 Kan and aTc at 37°C to OD₆₀₀ of 1.2 – 1.5. *ΔispG* strains containing the pBbS2k plasmids carrying
422 the IspG orthologs were grown in LB, Kan, MVA and aTc at 37°C to OD₆₀₀ of 1.2 – 1.5. In both
423 cases, the cells were centrifuged at 12,000 rpm for 2 min at room temperature and after removing
424 the supernatant, the pellets were resuspended in PBS 1X buffer and then loaded on 12% SDS-
425 PAGE gel after denaturing with Leammli loading dye at 100°C for 15 min. The gel was stained
426 with Coomassie brilliant blue G-250 and decolorized using a solution of 60% water, 30%
427 isopropanol, 10% acetic acid.

428

429 **Mass spectrometry detection of orthologs**

430 **Sample preparation.** *ΔispG* and *ΔnadA* strains bearing corresponding inactive orthologs were
431 inoculated in 10 mL permissive media (*ΔispG* in LB with 0.5 mM MVA, 100 μM IPTG, 25 ug/mL
432 kanamycin; or *ΔnadA* in LB with 25 ug/mL kanamycin). When cultures reached early exponential
433 phase (OD = 0.1), ortholog expression was maximally induced by 10 ng/mL doxycycline (added
434 at early exponential phase before OD = 0.1). The cultures were subsequently grown to mid-
435 logarithmic phase (OD = 0.4-0.6), at which point 2 mL solution of 10% (w/vol) trichloroacetic acid
436 was added. Quenched samples were incubated on ice for at least 10 minutes before centrifugation
437 (20,000 *g*, 4°C, 10 min), after which the supernatant was removed and pellet stored at -80°C
438 before analysis.

439 **Cell lysis and protein extraction.** Briefly, biomass amount equivalent to 1 mL of 1 OD *E. coli*
440 were collected in an Eppendorf tube and solubilised in a suspension solution consisting of 200µL
441 B-PER reagent (Thermo Scientific) and 200µL TEAB buffer (50 mM TEAB, 1% (w/w) NaDOC,
442 adjusted to pH 8.0) including 0.2µL protease inhibitor (P8215, Sigma Aldrich). Further, 0.1 g of
443 glass beads (acid, washed, approx. 100 µm diameter) were added and cells were disrupted using
444 3 cycles of bead beating on a vortex for 30 seconds followed by cooling on ice for 30 seconds in-
445 between cycles. In the following, a freeze/thaw step was performed by freezing the suspension at
446 -80°C for 15 minutes and thawing under shaking at elevated temperature using an incubator. The
447 cell debris was further pelleted by centrifugation using a bench top centrifuge at max speed, under
448 cooling for 10 minutes. The supernatant was transferred to a new Eppendorf tube and kept at 4°C
449 until further processed. Protein was precipitated by adding 1 volume of TCA (trichloroacetic acid,
450 Sigma Aldrich) to 4 volumes of supernatant. The solution was incubated at 4°C for 10 minutes and
451 further pelleted at 14,000 g for 10 minutes. The obtained protein precipitate was washed twice
452 using 250µL ice cold acetone.

453 **Proteolytic digestion.** The protein pellet was dissolved to approx. 100 µg/100 µL of 200 mM
454 ammonium bicarbonate containing 6 M Urea to a final concentration of approximately 100 µg/µL.
455 To 100 µL protein solution, 30 µL of a 10 mM DTT solution were added and incubated at 37°C for
456 1 hour. In the following, 30 µL of a freshly prepared 20 mM IAA solution was added and incubated
457 in the dark for 30 minutes. The solution was diluted to below 1 M urea using 200 mM bicarbonate
458 buffer and an aliquot of approximately 25 µg protein were digested using sequencing grade
459 Trypsin at 37°C overnight (Trypsin/Protein 1:50). Finally, protein digests were then further desalted
460 using an Oasis HLB 96 well plate (Waters) according to the manufacturer protocols. The purified
461 peptide eluate was dried using a speed vac concentrator.

462 **One dimensional shot-gun proteomics approach.** Briefly, the samples were analysed using a
463 nano-liquid-chromatography system consisting of an ESAY nano LC 1200, equipped with an

464 Acclaim PepMap RSLC RP C18 separation column (50 μ m x 150 mm, 2 μ m), and an QE plus
465 Orbitrap mass spectrometer (Thermo). The flow rate was maintained at 350 nL/min over a linear
466 gradient from 4% to 30% solvent B over 32.5 minutes, and finally to 70% B over 12.5 minutes.
467 Solvent A was H₂O containing 0.1% formic acid, and solvent B consisted of 80% acetonitrile in
468 H₂O and 0.1% formic acid. The Orbitrap was operated in data depended acquisition mode
469 acquiring peptide signals form 385-1250 m/z at 70K resolution with a max IT of 100 ms and an
470 AGC target of 3e6. The top 10 signals were isolated at a window of 2.0 m/z and fragmented using
471 a NCE of 28. Fragments were acquired at 17K resolution with a max IT of 75 ms and an AGC
472 target of 2e5. Singly charged, 6x and higher charged mass peaks were excluded from selection.
473 Data were acquired from 0 to 60 min.

474 **Database search.** Data were analysed against the proteome database from *E. coli* K-12
475 (UniprotKB, TaxID 83333), including amino acid sequences for heterologous orthologs, using
476 PEAKS Studio 10.0 (Bioinformatics Solutions Inc) allowing for 20 ppm parent ion and 0.02m/z
477 fragment ion mass error, 2 missed cleavages, carbamidomethylation as fixed and methionine
478 oxidation and N/Q deamidation as variable modifications. Peptide spectrum matches were filtered
479 against 1% false discovery rate (FDR) and protein identifications with ≥ 2 unique peptides were
480 accepted as significant.

481 **2-methyl tryptophan biosynthesis**

482 The mature protein coding sequence of the tryptophan 2-C-methyltransferase (TsrM) from *S.*
483 *laurentii* (GenBank: FJ652572.1) was codon-optimized for *E. coli* expression and synthesized by
484 Twist Bioscience. The optimized TsrM sequence was inserted into a pBbA5k BglBrick backbone
485 between the BglIII and BamHI sites to obtain the vector pTsrM. Codon-optimized genes encoding
486 electron transfer proteins were cloned between the BamHI and XhoI sites of the pTsrM vector
487 **(Table S8)**. The vector for cobalamin importer overexpression was constructed similarly to a
488 previously-published report²³. The genes for proteins BtuB (AYG21241.1, GenBank), BtuC

489 (AYG19236.1, GenBank), BtuD (AYG19238.1, GenBank), BtuE (AYG19237.1, GenBank) and
490 BtuF (AYG20683.1, GenBank) were amplified from *E. coli* MG1655 genome and cloned into the
491 backbone pBbS2c BglBrick backbone to create the vector pBtu.

492 2-methyl tryptophan was produced using *E. coli* NCM3722 pBtu and with plasmids pTsrM,
493 pTsrM_Scatt3Fd, pTsrM_Fdx1, pTsrM_Fdx2, or pTsrM_Fdx3. Individual colonies were picked and
494 inoculated for pre-cultures grown overnight in 1 mL of SAM MOPS minimal medium (MOPS
495 minimal medium⁴⁵ supplemented with 1% w/vol glucose, 0.25% Casamino acids, and 7 μ M
496 hydroxocobalamin), at 37°C. The cultures (triplicates) were obtained by inoculating 50 μ L of pre-
497 cultures in 3 mL of SAM MOPS minimal medium. The cultures were grown at 37°C to OD₆₀₀ of
498 0.05-0.1 for pBtu induction (83.3 ng/mL aTc) and continued to grow in the same conditions. At
499 OD₆₀₀ 0.5, TsrM and ferredoxin expression was induced (0.25 mM IPTG) and the medium
500 supplemented with 0.15 mM cysteine and 32.5 μ M FeCl₃. Cultures were incubated at room
501 temperature. Samples were taken after 24 hours (1 ml and the equivalent of 1 ml sample with 0.5
502 OD₆₀₀ samples). Samples were centrifuged at 15,000 rpm for 2 min, and supernatants were
503 transferred to new tubes. Pellets were quenched with NMM (methanol, acetonitrile and water, ratio
504 of 5:3:1) + 0.1% formic acid and resuspended. Pellet samples were dried in speedvac at 40°C and
505 resuspended with NMM for LC-MS measurement.

506 **2-methyltryptophan quantification**

507 2-methyl tryptophan was quantified by an LC-MS system (Agilent) using an XBridge BEH Amide
508 2.5 μ m (Waters, Bridge Columns) with a precolumn and equipped with a standard ESI source
509 mass spectrometer (sample injection volume of 5 μ L). The mobile phase was composed by two
510 solvents, A (20 mM ammonium formate in 10% acetonitrile) and B (20 mM ammonium formate in
511 80% acetonitrile). After 6 min at 100% solvent B, the metabolites were separated by a gradient
512 from 100% to 70% of solvent B for 6 min (flow rate 0.4 mL/min), followed by a gradient from 70%
513 to 100% for 50 sec (same flow rate) and held at 100% solvent B for 3 min 10 sec. The 2-methyl

514 tryptophan precursor ion (positive polarity, 219.1) was fragmented into product ions (144.1 and
515 128) using an ESI ionization in MRM mode. 2-methyl tryptophan concentration was estimated
516 using a calibration curve constructed with standard samples.

517 **Expression and purification of heterologous IspGs**

518 Coding sequences of all tested IspG proteins were ordered codon optimized and cloned in a
519 pET28a(+) plasmid by Genscript. A Tobacco etch virus (TEV) protease cleavage site was inserted
520 between the N-Terminal His-tag and the downstream IspG sequences. Each IspG was expressed
521 and purified under aerobic conditions. Expression was carried out in BL21DE3 cells for 3 hours
522 under 1 mM IPTG induction at 37°C. Harvested cells were resuspended in lysis buffer (50 mM
523 Tris pH 7.5, 250 mM NaCl, 10 mM imidazole), sonicated on ice (5 min, 10/30 sec ON/OFF, power
524 70%) and resulting lysate was clarified by ultracentrifugation (40,000 g, 20 min, 4°C). Supernatant
525 was loaded onto a 5 mL Ni-NTA column. Column was washed using lysis buffer + 20 mM imidazole
526 and proteins were eluted with lysis buffer + 300 mM imidazole. Fractions containing IspG were
527 pooled together and incubated with TEV protease harbouring a His-tag (molar ratio: 1/100) o/n at
528 4°C for dialysis against storage buffer (50 mM Tris pH 7.5, 250 mM NaCl, 1 mM DTT). Dialysis
529 bag content was reloaded onto 5 mL Ni-NTA column and flow-through (FT) containing IspG
530 (without His-tag) were collected. Pure apo-IspG were concentrated and stored at -80°C. To
531 counter instability of *E. coli* IspG, glycerol (50 % final) was added to concentrated purified protein
532 which was subsequently stored at -20°C.

533 **Expression, purification and reconstitution of *E. coli* ErpA**

534 *E. coli* ErpA was obtained as previously reported⁴⁶.

535 **Fe-S cluster transfer between *E. coli* ErpA and heterologous IspGs**

536 Fe-S cluster transfer experiment from *E. coli* ErpA to purified IspG orthologs was carried out in an
537 anaerobic chamber ($O_2 < 2$ ppm). For each transfer, one equivalent of apo-IspG (52 nmoles) from
538 *E. coli*, *B. subtilis*, *S. cattleya*, or *Synechocystis sp. CACIAM05* was incubated in transfer buffer
539 (50 mM Tris pH 7.5, 250 mM NaCl, 1 mM DTT) for 5 min before addition of 2.2 equivalents of [Fe-
540 S]-ErpA from *E. coli* (115 nmoles). After 1 hour incubation, the mixture was loaded onto a 5 mL
541 Ni-NTA column. Fractions containing [Fe-S]-IspG were collected in the FT whereas ErpA was
542 eluted using elution buffer (50 mM Tris pH 7.5, 250 mM NaCl, 300 mM imidazole). Fractions were
543 analysed by SDS-PAGE. UV-Vis spectra of ErpA and IspG were recorded before and after
544 incubation and Fe-S transfer.

545 **TsrM overexpression and purification for *in vitro* studies**

546 The coding sequence of TsrM from *S. laurentii* preceded by a N-terminal Tobacco etch virus (TEV)
547 protease cleavable His-tag, was ordered codon optimized and sub-cloned into pET28a(+) vectors
548 (Genscript). TsrM was co-expressed together with the Btu operon in BL21(DE3) cultured in LB
549 medium supplemented with $FeCl_3$ (50 μ M), L-Cysteine (150 μ M) and hydroxy-Cobalamin (2 μ M).
550 Btu operon expression was induced at $OD_{600} = 0.3$ using aTc (100 ng/mL) and TsrM expression
551 was induced using IPTG (1 mM) at an $OD_{600} = 0.7$ for 18 hours at 18°C. Cells were harvested
552 (6000 rpm, 20 min, 4°C), washed with NaCl 0.9 % and stored at -80°C. TsrM was purified
553 purification under anaerobic conditions ($O_2 < 2$ ppm) in a glove box (Jacomex). Cells expressing
554 TsrM were resuspended in buffer containing 50 mM Tris pH 7.5, 250 mM NaCl, 10 mM imidazole,
555 10 % glycerol, 0.1 % Tween 20 and sonicated for 7 min (10/30 sec ON/OFF, power 50%). After
556 ultracentrifugation (40,000 g, 20 min, 4°C), the soluble fraction was loaded onto a 5 mL Ni-NTA
557 column equilibrated with buffer containing 50 mM Tris pH 7.5, 250 mM NaCl. After an extensive
558 washing with buffer containing 50 mM Tris pH 7.5, 250 mM NaCl, 20 mM imidazole buffer, TsrM
559 was eluted with buffer containing 50 mM Tris pH 7.5, 250 mM NaCl, 300 mM imidazole. Imidazole
560 was removed using a HiPrep 26/10 desalting column equilibrated with buffer containing 50 mM

561 Tris pH 7.5, 250 mM NaCl, 1 mM DTT. TsrM was then concentrated using an amicon cell until 37
562 mg/mL. As-isolated TsrM was then buffer exchanged into 25 mM Hepes pH 7.5, 300 mM KCl, 5%
563 glycerol using micro Bio-spin 6 desalting column and used for electrochemistry (cyclic
564 voltammetry) under anaerobic conditions or aliquoted for storage in liquid N₂.

565 **Expression and purification of *S. cattleya* ferredoxins**

566 Coding sequences of the three ferredoxins from *S. cattleya* (AEW96347.1 (Fdx1), AEW92689.1
567 (Fdx2), AEW97532.1 (Fdx3)), all preceded by a N-terminal Tobacco etch virus (TEV) protease
568 cleavable His-tag, were ordered codon optimized for expression in *E. coli* and sub-cloned into
569 pET28a(+) vectors (Genscript). Each ferredoxin was expressed in BL21(DE3) cells cultured in LB
570 medium. Protein expression was induced using IPTG (1 mM) for 3 hours at 37°C. Cells were
571 harvested (6000 rpm, 20 min, 4°C) and washed with NaCl 0.9 % before storage at -80°C. All
572 ferredoxins were purified under anaerobic conditions (O₂ < 2 ppm) in a glove box. Cells expressing
573 ferredoxins were resuspended in buffer containing 50 mM Tris pH 7.5, 250 mM NaCl, 10 mM
574 imidazole, 10 % glycerol, 0.1 % Tween 20 and sonicated for 7 min (10/30 sec ON/OFF, power
575 50%). After ultracentrifugation (40,000 g, 20 min, 4°C), the soluble fractions were loaded onto 3
576 separate 5 mL Ni-NTA columns equilibrated with 50 mM Tris pH 7.5, 250 mM NaCl. Columns were
577 washed using buffer containing 50 mM Tris pH 7.5, 250 mM NaCl, 20 mM imidazole and proteins
578 were eluted with buffer containing 50 mM Tris pH 7.5, 250 mM NaCl, 300 mM imidazole. Imidazole
579 was removed using a HiPrep 26/10 desalting column, equilibrated with a buffer containing 50 mM
580 Tris pH 7.5, 250 mM NaCl, 1 mM DTT. Ferredoxins cleared of imidazole were incubated overnight
581 in presence of TEV (molar ratio: 1/100). The mixtures were reloaded onto 5 different 5 mL Ni-NTA
582 columns and flow-through (FT), containing ferredoxins, were collected.

583 **Fe-S reconstitution of *S. cattleya* Fdx1**

584 Fdx2 and Fdx3 from *S. cattleya* were purified anaerobically with their Fe-S clusters in contrast to
585 Fdx1 which was obtained as an apo-form. Fdx1 was therefore reconstituted within an anaerobic

586 chamber. Fdx1 was incubated in buffer containing 50 mM Tris pH 7.5, 250 mM NaCl, 1 mM DTT
587 with 10-fold equivalents of Fe²⁺ (Mohr salt) and 10-fold equivalents of S²⁻ (Na₂S). The reaction
588 occurred for 3 hours at RT and after centrifugation (15 min at 16,000 g) the mixture was loaded
589 onto a Superdex-75 10/300 column. Fractions containing reconstituted Fdx1 were pooled and
590 concentrated using concentrators for microfuge. Subsequent analysis of Fe content and
591 concentration determination were performed as described below.

592 **Biochemical analyses of TsrM**

593 Concentration of proteins (TsrM and ferredoxins) was determined using Rose Bengal with BSA
594 as standard⁴⁷ and the Fe content was determined using the Fish method⁴⁸. The cobalamin content
595 of as-isolated TsrM was determined by UV-Vis absorption spectroscopy through its conversion to
596 dicyanocobalamin ($\epsilon_{367} = 30,800 \text{ M}^{-1}\text{cm}^{-1}$) using a treatment with 0.1 M potassium cyanide
597 following a procedure previously reported²¹.

598 **Protein-Film Electrochemistry**

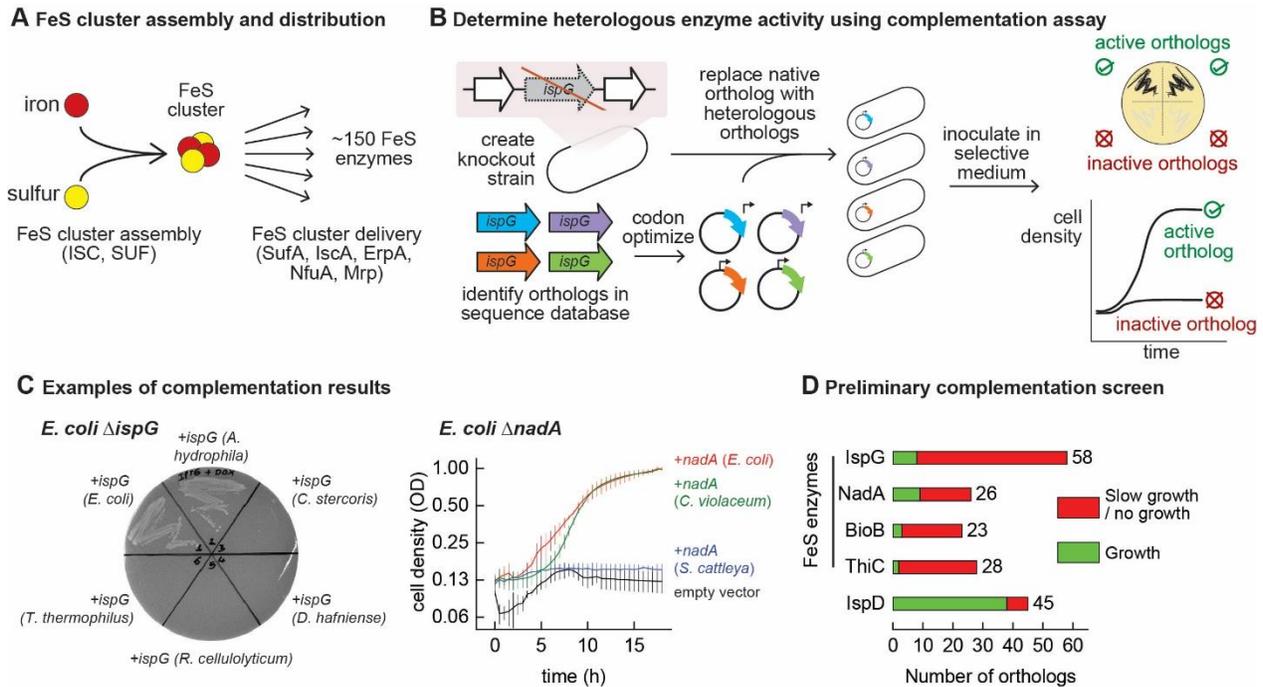
599 Protein-film electrochemistry experiments were performed anaerobically inside an anaerobic
600 chamber (O₂ < 2 ppm) using freshly purified TsrM, [Fe-S]-containing ferredoxins and a potentiostat
601 (Biologic). A three-electrode configuration was used in a small volume analytical cell (Biologic)
602 with a platinum wire and an Ag/AgCl electrode as counter and reference electrode respectively.
603 When analysing TsrM, a pyrolytic graphite edge (PGE) electrode was used to collect
604 electrochemical measurements. The electrode was polished with sand paper 1200 followed by 1
605 μm alumina, and then baseline measurements were collected by placing the PGE electrode into
606 the buffer cell solution (10 mM MES, 10 mM CHES, 10 mM TAPS, 10 mM HEPES pH 8, 200 mM
607 NaCl). Then, 3 μL of 350 μM TsrM were applied to the polished PGE electrode for 5 min before
608 being rinsed with 200 μL of the buffer cell solution. Next, the electrode was immediately placed
609 back in the buffer cell solution for measurements. For ferredoxins, a glassy carbon (GC) electrode

610 polished using 1 μm alumina was used within a setup combining a small volume analytical cell
611 and the sample holder of the SVC-2 kit (Biologic). Ferredoxin solutions were at 50 μM (in 10 mM
612 MES, 10 mM CHES, 10 mM TAPS, 10 mM HEPES pH 8, 200 mM NaCl buffer). Cyclic
613 voltammograms were collected at room temperature with a scan rate of 100 mV/sec. Redox
614 potentials of TsrM and ferredoxins were determined through square wave voltammetry (SWV)
615 under the same conditions as previously described for cyclic voltammetry. The SWV input signal
616 consisted of a staircase ramp from -1 to -0.3 V vs. Ag/AgCl, with 2 mV increments, 50 mV stair
617 amplitude and 5 Hz frequency. Electrochemical signals were analysed by correction of the non-
618 Faradaic component from the raw data using the QSoas package⁴⁹.

619 **TsrM-ferredoxin affinity measurements**

620 Affinity interactions between TsrM and *S. cattleya* ferredoxins were measured by BioLayer
621 Interferometry inside an anaerobic chamber ($\text{O}_2 < 2$ ppm) equipped with a BLItz system (FortéBio).
622 All proteins were buffer exchanged into 25 mM Hepes pH 7.5, 300 mM KCl, 5% glycerol using
623 micro Bio-spin6 desalting columns. Ferredoxins were biotinylated through incubation with 1
624 equivalent of NHS-PEG₄-Biotin. After 30 min, excess NHS-PEG₄-Biotin was removed using a
625 micro Bio-spin 6 desalting column. Biotinylated ferredoxins (10 $\mu\text{g}/\text{mL}$) were bound for 60 sec to
626 streptavidin biosensors equilibrated in BLI buffer (25mM Hepes pH 7.5, 300 mM KCl, 5% glycerol)
627 containing Tween 20 and BSA to inhibit non-specific binding, following manufacturer
628 recommendations. After having been plunged for 20 sec for equilibration in BLI buffer, the loaded-
629 biosensor was plunged into TsrM solution for 80 sec, followed by 60 sec into BLI buffer for
630 association constant and dissociation constant measurements, respectively. This procedure was
631 repeated for various concentrations of TsrM. Affinity constants were extracted from fitted raw data
632 using the BLItz software (PALL/FortéBio).

633 Figures and Tables



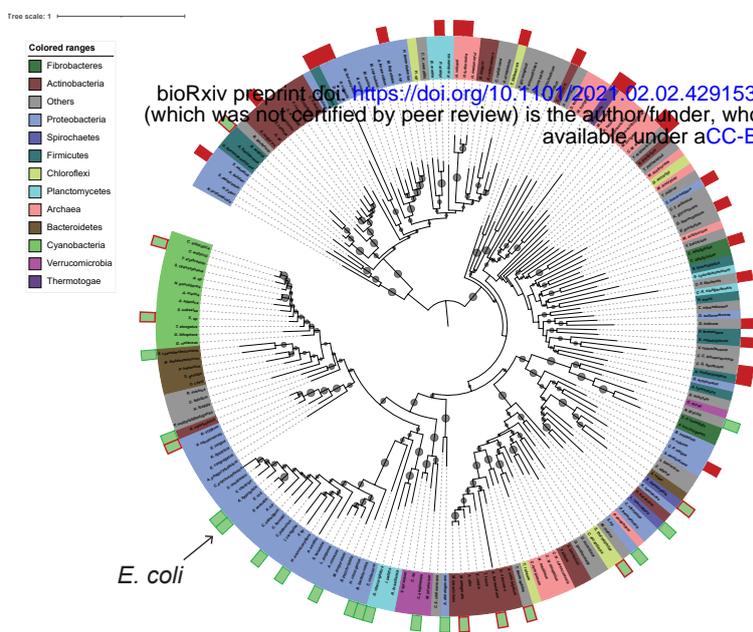
634

635 **Figure 1.** Mapping functional expression of heterologous Fe-S enzymes using complementation
636 experiments. **A.** In prokaryotes, two pathways that assemble and deliver Fe-S clusters have been
637 identified (ISC, SUF), as have several Fe-S delivery proteins (SufA, IscA, ErpA, NfuA in *E. coli*).
638 **B.** The complementation assays used to determine activity of heterologous orthologs expressed
639 by *E. coli*. Orthologs of conditionally-essential enzymes are identified in sequence databases,
640 codon-optimized, and cloned in a low-copy plasmid. Expression is controlled by an inducible
641 promoter (P_{Tet}). The plasmids are then transformed into *E. coli* strains lacking the corresponding
642 *E. coli* ortholog. Each strain is inoculated in selective media and expression of the heterologous
643 ortholog is induced. Growth in selective media indicates that the heterologous ortholog retains
644 sufficient activity to support growth of *E. coli*. **C.** Exemplary results of complementation assay
645 obtained using either solid or liquid selective medium. Error bars represent standard error of
646 duplicate cultures inoculated from separate colonies. **D.** Results of preliminary complementation

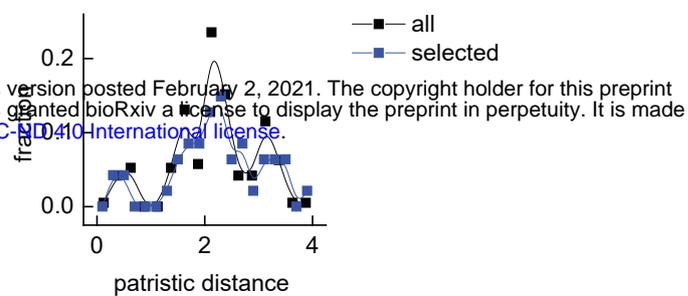
647 tests comparing activity of Fe-S enzymes with a non-Fe-S enzyme. Orthologs tested for the
648 preliminary screen are listed in Table S1.

649

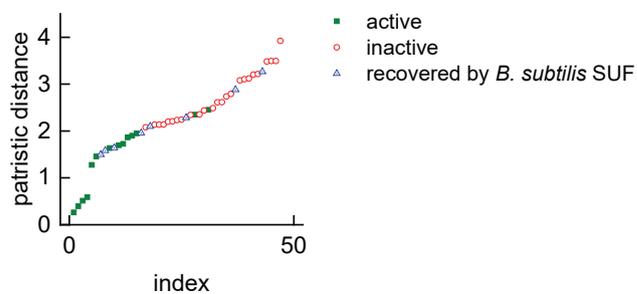
A Phylogenetic tree of NadA orthologs



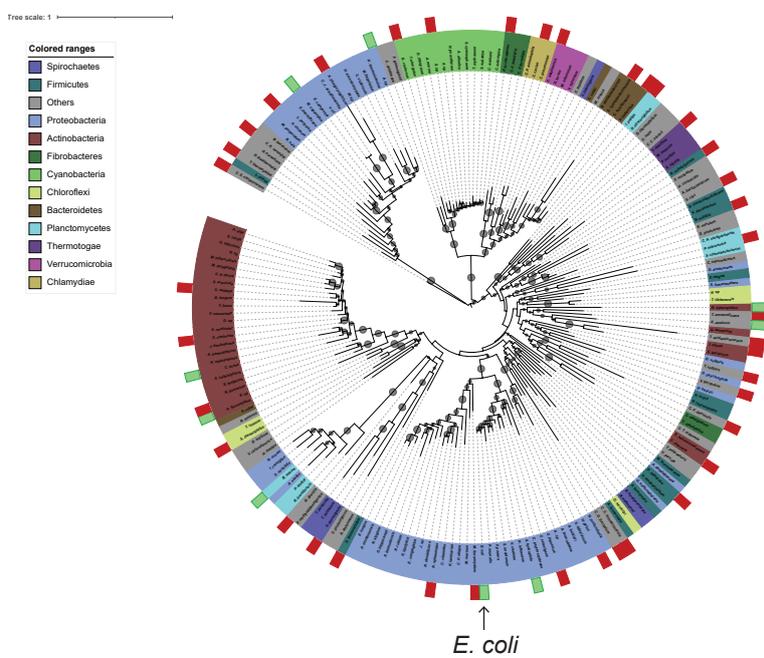
B Phylogenetic distributions of NadA orthologs



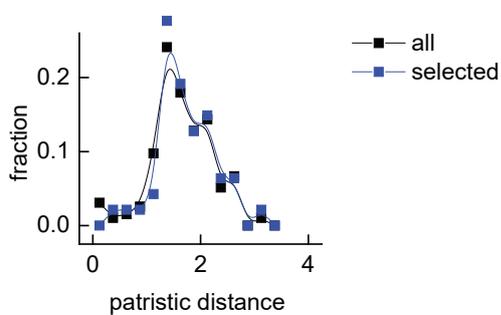
C Patristic distance and NadA activity in *E. coli*



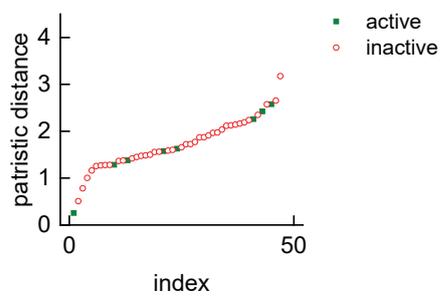
D Phylogenetic tree of IspG orthologs



E Phylogenetic distributions of IspG orthologs

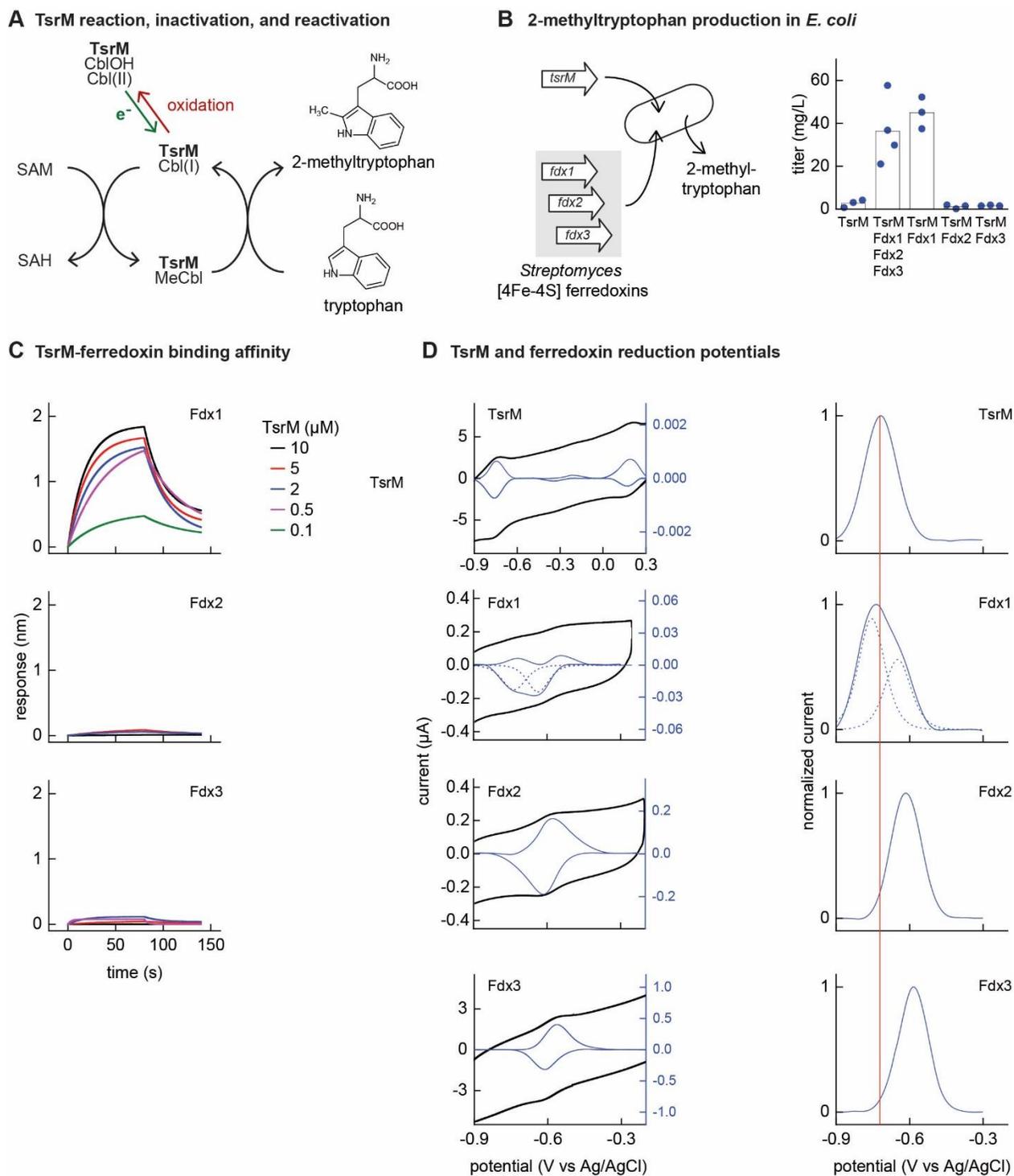


F Patristic distance and IspG activity in *E. coli*



652 **Figure 2.** Phylogenetic maps contrasting the compatibility of Fe-S enzymes NadA (**A-C**) and IspG
653 (**D-F**) across a selection of orthologs representative of known prokaryotic sequence diversity. **A,**
654 **D:** Phylogeny of NadA orthologs (**A**) and IspG orthologs (**D**). In each diagram, the position of *E.*
655 *coli* ortholog is highlighted by an arrow. Orthologs selected for our complementation assay are
656 indicated by color rectangles (green: successful complementation; green surrounded by red:
657 complementation by co-expression of *B. subtilis* SUF; red: no complementation) **B, E:**
658 Distributions of patristic distances from the corresponding *E. coli* ortholog of all sequences (black
659 curves) and the sequences of orthologs selected for testing (blue curves). **C, F:** Complementation
660 and reactivation results correlated with patristic distance from the *E. coli* ortholog.

661



662

663 **Figure 3.** Identifying the barrier to optimal activity of an Fe-S-dependent rSAM enzyme. **A.** The
 664 enzyme TsrM synthesizes 2-methyltryptophan as part of the synthesis pathway of the antibiotic
 665 thiostrepton. A one-electron reduction (in green) converts the bound cobalamin to cob(I)alamin,

666 which is converted to MeCbl via reaction with SAM. TsrM can be inactivated during catalysis by
667 adventitious oxidation of the bound cob(I)alamin to cob(II)alamin (in red). Recovering the
668 catalytically-active cob(I)alamin form requires a one-electron reduction (green). **B.** Titrers of 2-
669 methyltryptophan in a TsrM-expressing *E. coli* strain increase substantially upon co-expression of
670 a specific *S. cattleya* ferredoxin (Fdx1) whereas co-expression of alternative *S. cattleya* [4Fe-4S]
671 ferredoxins, Fdx2 and Fdx3, do not increase titers. **C, D.** Identifying the molecular basis enabling
672 Fdx1 to recover TsrM activity. **C.** Bio-layer interferometry (BLI) of TsrM with Fdx1, Fdx2, and Fdx3
673 indicates that TsrM interacts measurably with Fdx1 alone. **D.** Cyclic voltammetry of TsrM and the
674 three ferredoxins indicates that the reducing potentials of one of the [Fe-S] clusters of Fdx1 is
675 sufficiently low-potential to reduce TsrM metal cofactors (cob(II)alamin to cob(I)alamin and [4Fe-
676 4S]⁺² to [4Fe-4S]⁺¹) and thus restore activity after adventitious oxidation.

677

Phylum	Species/Strain	Accession number	Complementation		
			NadA	NadA + <i>B. subtilis</i> SUF	NadA + <i>E. coli</i> SUF
Actinobacteria	<i>Mycolicibacterium smegmatis</i> MC2 155	ABK69796.1	-	+	+
	<i>Rubrobacter xylanophilus</i> DSM 9941	ABG03084.1	+	N.D.	N.D.
	<i>Collinsella stercoris</i> DSM 13279	EEA90923.1	-	-	-
	<i>Acidimicrobium ferrooxidans</i> DSM 10331	ACU53899.1	-	+	-
	<i>Streptomyces cattleya</i> NRRL 8057 = DSM 46488	AEW93658.1	-	-	-
Bacteroidetes	<i>Salinivirga cyanobacteriivorans</i>	ALO15705.1	+	N.D.	N.D.
Candidatus Fermentibacteria	<i>Candidatus Fermentibacter daniensis</i>	KZD19396.1	-	-	-
Candidatus Melainabacteria	<i>Candidatus Gastranaerophilales bacterium HUM_7</i>	DAA95005.1	-	-	-
Chloroflexi	<i>Sphaerobacter thermophilus</i> DSM 20745	ACZ38300.1	-	+	-
	<i>Thermogemmatispora tikiterensis</i>	RAQ94418.1	-	-	-
Chrysiogenetes	<i>Desulfurispirillum indicum</i> S5	ADU65000.1	-	-	-
Cyanobacteria	<i>Synechocystis</i> sp. PCC 6803	BAA18685.1	-	+	-
	<i>Crocospaera watsonii</i> WH 8501	WP_007304059.1	-	+	-
Deinococcus-Thermus	<i>Thermus thermophilus</i> HB27	AAS80968.1	-	+	-
Dictyoglomi	<i>Dictyoglomus thermophilum</i> H-6-12	ACI18692.1	-	-	-
Euryarchaeota	<i>Methanopyrus kandleri</i> AV19	AAM01290.1	-	-	-
	<i>Haloferax volcanii</i> DS2	ADE02976.1	-	-	-
	<i>Haloterrigena turkmenica</i> DSM 5511	ADB59428.1	-	-	-
	<i>Candidatus Methanomethylophilus alvus</i> Mx1201	AGI85733.1	-	-	-
Fibrobacteres	<i>Chitinivibrio alkaliphilus</i> ACh1	ERP31080.1	-	-	-
Firmicutes	<i>Bacillus subtilis</i> subsp. <i>subtilis</i> str. 168	CAB14745.1	-	+	+
	<i>Desulfitobacterium hafniense</i> DCB-2	ACL18980.1	-	-	-
	<i>Ruminiclostridium cellulolyticum</i> H10	ACL77768.1	-	-	-
	<i>Helio bacterium modesticaldum</i> Ice1	ABZ83886.1	-	-	-
Nitrospinae	<i>Nitrospina gracilis</i>	CCQ89850.1	+	N.D.	N.D.
Planctomycetes	<i>Gemmata obscuriglobus</i>	WP_010044270.1	+	N.D.	N.D.
	<i>Pirellula staleyi</i> DSM 6068	ADB17048.1	-	-	-
Proteobacteria	<i>Aeromonas hydrophila</i>	ABK37682.1	+	N.D.	N.D.
	<i>Desulfovibrio vulgaris</i> str. <i>Hildenborough</i>	AAS96285.1	-	-	-
	<i>Allochromatium vinosum</i>	ADC62035.1	+	N.D.	N.D.
	<i>Arcobacter butzleri</i>	WP_020848486.1	-	-	-
	<i>Novosphingobium stygium</i>	SCY25206.1	-	+	+
	<i>Hermiimonas arsenicoxydans</i>	CAL61268.1	+	N.D.	N.D.
	<i>Chromobacterium violaceum</i> ATCC 12472	AAQ61340.1	+	N.D.	N.D.
	<i>Bdellovibrio bacteriovorus</i> HD100	CAE81099.1	+	N.D.	N.D.
	<i>Methylobacillus flagellatus</i> KT	ABE50646.1	-	-	-
<i>Anaeromyxobacter dehalogenans</i> 2CP-C	ABC83113.1	+	N.D.	N.D.	

	<i>Candidatus Pelagibacter ubique</i> HTCC1002	EAS84697.1	-	-	-
	<i>Syntrophobacter fumaroxidans</i> MPOB	ABK17287.1	-	-	-
	<i>Cellvibrio japonicus</i> Ueda107	ACE83018.1	+	N.D.	N.D.
	<i>Escherichia coli</i> str. K-12 substr. MG1655	AAC73837.1	+	N.D.	N.D.
	<i>Desulfarculus baarsii</i> DSM 2075	ADK83407.1	-	-	-
	<i>Salinisphaera</i> sp. LB1	AWN14904.1	+	N.D.	N.D.
<i>Spirochaetes</i>	<i>Leptospira interrogans</i> serovar Lai str. 56601	AAN47817.2	+	N.D.	N.D.
	<i>Spirochaeta thermophila</i> DSM 6578	AEJ62405.1	-	+	-
<i>Synergistetes</i>	<i>Thermanaerovibrio acidaminovorans</i> DSM 6589	ACZ19971.1	-	-	-
<i>Thermotogae</i>	<i>Thermotoga maritima</i> MSB8	AAD36711.1	-	-	-
<i>Verrucomicrobia</i>	<i>Coraliomargarita akajimensis</i> DSM 45221	ADE55989.1	+	N.D.	N.D.

678

679 **Table 1.** Results of NadA complementation of *E. coli* Δ nadA. Minus (-): negative complementation; plus (+): positive complementation;

680 N.D.: not determined.

681

Phylum	Species/Strain	Accession number	Complementation				
			IspG	IspG + <i>Synechocystis</i> PetF	IspG + <i>B. subtilis</i> YkuP/YkuN	IspG + <i>S. cattleya</i> Fdx	IspG + <i>R. sphaeroides</i> Fdx
<i>Acidobacteria</i>	<i>Candidatus Koribacter versatilis</i> Ellin345	ABF40425.1	-	N.D.	N.D.	N.D.	N.D.
<i>Actinobacteria</i>	<i>Thermoleophilum album</i>	SEH14556.1	-	N.D.	N.D.	N.D.	N.D.
	<i>Streptomyces coelicolor</i> A3	CAB91132.1	-	-	-	-	N.D.
	<i>Rubrobacter xylanophilus</i> DSM 9941	ABG04368.1	+	N.D.	N.D.	N.D.	N.D.
	<i>Cutibacterium acnes</i> KPA171202	AAT83253.1	+	N.D.	N.D.	N.D.	N.D.
	<i>Collinsella stercoris</i> DSM 13279	EEA90248.1	-	-	-	-	N.D.
	<i>Acidimicrobium ferrooxidans</i> DSM 10331	ACU53587.1	+	N.D.	N.D.	N.D.	N.D.
	<i>Streptomyces cattleya</i> NRRL 8057 = DSM 46488	AEW97312.1	-	-	-	+	N.D.
	<i>Egibacter</i> sp.	PWFH01000026.1_54	-	N.D.	N.D.	N.D.	N.D.
<i>Aquificae</i>	<i>Aquifex aeolicus</i> VF5	AAC07467.1	+	N.D.	N.D.	N.D.	N.D.
	<i>Thermovibrio ammonificans</i> HB-1	ADU96577.1	-	N.D.	N.D.	N.D.	N.D.
<i>Bacteroidetes</i>	<i>Pedobacter heparinus</i> DSM 2366	ACU05649.1	-	N.D.	N.D.	N.D.	N.D.
	<i>Salinivirga cyanobacteriivorans</i>	ALO15331.1	-	N.D.	N.D.	N.D.	N.D.
<i>Candidatus Fermentibacteria</i>	<i>Candidatus Fermentibacter daniensis</i>	KZD16972.1	-	N.D.	N.D.	N.D.	N.D.
<i>Candidatus Melainabacteria</i>	<i>Candidatus Caenarcanum bioreactoricola</i>	MOYB01000001.1_3	-	N.D.	N.D.	N.D.	N.D.
	<i>Candidatus Gastranaerophilales bacterium HUM_7</i>	DAA93923.1	-	N.D.	N.D.	N.D.	N.D.
<i>Candidatus Sumerlaeota</i>	<i>Candidatus Sumerlaea chitinovorans</i>	AXA34785.1	-	N.D.	N.D.	N.D.	N.D.
<i>Chlamydiae</i>	<i>Chlamydia caviae</i> GPIC	AAP05169.1	-	-	-	-	N.D.
<i>Chloroflexi</i>	<i>Sphaerobacter thermophilus</i> DSM 20745	ACZ37901.1	-	N.D.	N.D.	N.D.	N.D.
<i>Cyanobacteria</i>	<i>Gloeobacter violaceus</i> PCC 7421	BAC91563.1	-	-	N.D.	N.D.	N.D.
	<i>Synechocystis</i> sp. PCC 6803	BAA17717.1	-	+	-	-	N.D.
<i>Deinococcus-Thermus</i>	<i>Thermus thermophilus</i> HB27	AAS82019.1	-	-	-	-	N.D.
<i>Fibrobacteres</i>	<i>Fibrobacter succinogenes</i> subsp. <i>succinogenes</i> S85	ACX75977.1	-	N.D.	N.D.	N.D.	N.D.
<i>Firmicutes</i>	<i>Bacillus subtilis</i> subsp. <i>subtilis</i> str. 168	CAB14437.1	-	+	+	+	N.D.
	<i>Desulfitobacterium hafniense</i> DCB-2	ACL21716.1	-	-	-	-	N.D.
	<i>Symbiobacterium thermophilum</i> IAM 14863	BAD40486.1	-	N.D.	-	N.D.	N.D.
	<i>Ruminiclostridium cellulolyticum</i> H10	ACL74833.1	-	-	-	-	N.D.
<i>Lentisphaerae</i>	<i>Victivallales bacterium</i> CCUG 44730	AVM43904.1	-	N.D.	N.D.	N.D.	N.D.
<i>Nitrospinae</i>	<i>Nitrospina gracilis</i>	CCQ91393.1	-	-	N.D.	N.D.	N.D.
<i>Planctomycetes</i>	<i>Blastopirellula marina</i> DSM 3645	EAQ77076.1	+/-	+	+	-	N.D.
	<i>Isosphaera pallida</i> ATCC 43644	ADV63232.1	-	-	-	-	N.D.
	<i>Phycisphaera mikurensis</i> NBRC 102666	BAM02738.1	-	N.D.	N.D.	N.D.	N.D.

Proteobacteria	<i>Aeromonas hydrophila</i>	ABK39186.1	+	N.D.	N.D.	N.D.	N.D.
	<i>Allochromatium vinosum</i>	ADC62931.1	-	N.D.	N.D.	N.D.	N.D.
	<i>Azoarcus sp. BH72</i>	CAL93544.1	+	N.D.	N.D.	N.D.	N.D.
	<i>Helicobacter pylori J99</i>	WP_000892444.1	-	-	-	N.D.	N.D.
	<i>Methylococcus capsulatus str. Bath</i>	AAU91397.1	+	N.D.	N.D.	N.D.	N.D.
	<i>Rhodobacter sphaeroides 2.4.1</i>	ABA79143.1	-	-	-	-	+
	<i>Mariiprofundus ferrooxydans</i>	EAU54713.1	-	N.D.	N.D.	N.D.	N.D.
	<i>Cellvibrio japonicus Ueda107</i>	ACE82628.1	-	-	-	-	N.D.
	<i>Escherichia coli str. K-12 substr. MG1655</i>	AAC75568.1	+	N.D.	N.D.	N.D.	N.D.
	<i>Anaplasma phagocytophilum str. CRT38</i>	EOA62006.1	-	N.D.	N.D.	N.D.	-
<i>Spirochaetes</i>	<i>Spirochaeta thermophila DSM 6578</i>	WP_014624852.1	-	N.D.	N.D.	-	N.D.
<i>Synergistetes</i>	<i>Thermanaerovibrio acidaminovorans DSM 6589</i>	ACZ19163.1	-	N.D.	N.D.	N.D.	N.D.
<i>Tenericutes</i>	<i>Spiroplasma citri</i>	APE74616.1	-	N.D.	N.D.	N.D.	N.D.
<i>Thermodesulfobacteria</i>	<i>Thermodesulfatator indicus DSM 15286</i>	AEH44631.1	-	N.D.	N.D.	N.D.	N.D.
<i>Thermotogae</i>	<i>Thermotoga maritima MSB8</i>	AAD35972.1	-	-	-	-	N.D.
<i>Verrucomicrobia</i>	<i>Coralimargarita akajimensis DSM 45221</i>	ADE56011.1	-	-	-	-	N.D.

682

683 **Table 2.** Results of the IspG complementation of *E. coli* Δ ispG. Minus (-): negative complementation; plus (+): positive complementation;

684 N.D.: not determined.

685

686 **References**

- 687 1. Porse, A., Schou, T. S., Munck, C., Ellabaan, M. M. H. & Sommer, M. O. A. Biochemical
688 mechanisms determine the functional compatibility of heterologous genes. *Nat. Commun.*
689 **9**, (2018).
- 690 2. Carlsen, S. *et al.* Heterologous expression and characterization of bacterial 2-C-methyl-D-
691 erythritol-4-phosphate pathway in *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.*
692 **97**, 5753–5769 (2013).
- 693 3. Kirby, J. *et al.* Engineering a functional 1-deoxy-D-xylulose 5-phosphate (DXP) pathway in
694 *Saccharomyces cerevisiae*. *Metab. Eng.* **38**, 494–503 (2016).
- 695 4. Ajikumar, P. K. *et al.* Isoprenoid pathway optimization for Taxol precursor overproduction
696 in *Escherichia coli*. *Science (80-.)*. **330**, 70–74 (2010).
- 697 5. Eseverri, Á. *et al.* Use of synthetic biology tools to optimize the production of active
698 nitrogenase Fe protein in chloroplasts of tobacco leaf cells. *Plant Biotechnol. J.* **18**, 1882–
699 1896 (2020).
- 700 6. Johnson, D. C., Dean, D. R., Smith, A. D. & Johnson, M. K. Structure, Function, and
701 Formation of Biological Iron-Sulfur Clusters. *Annu. Rev. Biochem.* **74**, 247–281 (2005).
- 702 7. Roche, B. *et al.* Iron/sulfur proteins biogenesis in prokaryotes: Formation, regulation and
703 diversity. *Biochim. Biophys. Acta - Bioenerg.* **1827**, 923–937 (2013).
- 704 8. Braymer, J. J. & Lill, R. Iron–sulfur cluster biogenesis and trafficking in mitochondria.
705 *Journal of Biological Chemistry* vol. 292 12754–12763 (2017).
- 706 9. Py, B. *et al.* The ErpA/NfuA complex builds an oxidation-resistant Fe-S cluster delivery
707 pathway. *J. Biol. Chem.* **293**, 7689–7702 (2018).

- 708 10. Burschel, S. *et al.* Iron-sulfur cluster carrier proteins involved in the assembly of
709 Escherichia coli NADH:ubiquinone oxidoreductase (complex I). *Mol. Microbiol.* **111**, 31–45
710 (2019).
- 711 11. Puan, K. J., Wang, H., Dairi, T., Kuzuyama, T. & Morita, C. T. fldA is an essential gene
712 required in the 2-C-methyl-D-erythritol 4-phosphate pathway for isoprenoid biosynthesis.
713 *FEBS Lett.* **579**, 3802–3806 (2005).
- 714 12. Zhou, K., Zou, R. Y., Stephanopoulos, G. & Too, H. P. Enhancing solubility of
715 deoxyxylulose phosphate pathway enzymes for microbial isoprenoid production. *Microb.*
716 *Cell Fact.* **11**, (2012).
- 717 13. Chazarreta-Cifre, L., Martiarena, L., de Mendoza, D. & Altabe, S. G. Role of ferredoxin
718 and flavodoxins in Bacillus subtilis fatty acid desaturation. *J. Bacteriol.* **193**, 4043–4048
719 (2011).
- 720 14. Arcinas, A. J., Maiocco, S. J., Elliott, S. J., Silakov, A. & Booker, S. J. Ferredoxins as
721 interchangeable redox components in support of MiaB, a radical S-adenosylmethionine
722 methylthiotransferase. *Protein Sci.* **28**, 267–282 (2019).
- 723 15. Gaudu, P. & Weiss, B. Flavodoxin mutants of Escherichia coli K-12. *J. Bacteriol.* **182**,
724 1788–1793 (2000).
- 725 16. Okada, K. & Hase, T. Cyanobacterial non-mevalonate pathway: (E)-4-hydroxy-3-
726 methylbut-2-enyl diphosphate synthase interacts with ferredoxin in
727 Thermosynechococcus elongatus BP-1. *J. Biol. Chem.* **280**, 20672–20679 (2005).
- 728 17. Kirby, J. *et al.* Engineering a functional 1-deoxy-D-xylulose 5-phosphate (DXP) pathway in
729 Saccharomyces cerevisiae. *Metab. Eng.* (2016) doi:10.1016/j.ymben.2016.10.017.
- 730 18. Broderick, J. B., Duffus, B. R., Duschene, K. S. & Shepard, E. M. Radical S-

- 731 adenosylmethionine enzymes. *Chem. Rev.* **114**, 4229–4317 (2014).
- 732 19. Mehta, A. P. *et al.* Radical S-adenosylmethionine (SAM) enzymes in cofactor
733 biosynthesis: A treasure trove of complex organic radical rearrangement reactions. *J. Biol.*
734 *Chem.* **290**, 3980–3986 (2015).
- 735 20. Mahanta, N., Hudson, G. A. & Mitchell, D. A. Radical S-Adenosylmethionine Enzymes
736 Involved in RiPP Biosynthesis. *Biochemistry* **56**, 5229–5244 (2017).
- 737 21. Blaszczyk, A. J. *et al.* Spectroscopic and Electrochemical Characterization of the Iron-
738 Sulfur and Cobalamin Cofactors of TsrM, an Unusual Radical S-Adenosylmethionine
739 Methylase. *J. Am. Chem. Soc.* **138**, 3416–3426 (2016).
- 740 22. Blaszczyk, A. J., Knox, H. L. & Booker, S. J. Understanding the role of electron donors in
741 the reaction catalyzed by TsrM, a cobalamin-dependent radical S-adenosylmethionine
742 methylase. *J. Biol. Inorg. Chem.* **24**, 831–839 (2019).
- 743 23. Lanz, N. D. *et al.* Enhanced Solubilization of Class B Radical S-Adenosylmethionine
744 Methylases by Improved Cobalamin Uptake in *Escherichia coli*. *Biochemistry* **57**, 1475–
745 1490 (2018).
- 746 24. Cammack, R. Iron—sulfur clusters in enzymes: Themes and variations. *Adv. Inorg.*
747 *Chem.* **38**, 281–322 (1992).
- 748 25. Lexa, D. & Saveant, J. M. The Electrochemistry of Vitamin B12. *Acc. Chem. Res.* **16**,
749 235–243 (1983).
- 750 26. Jarrett, J. T., Choi, C. Y. & Matthews, R. G. Changes in protonation associated with
751 substrate binding and cob(I)alamin formation in cobalamin-dependent methionine
752 synthase. *Biochemistry* **36**, 15739–15748 (1997).

- 753 27. Zheng, C. & Dos Santos, P. C. Metallocluster transactions: Dynamic protein interactions
754 guide the biosynthesis of Fe–S clusters in bacteria. *Biochemical Society Transactions* vol.
755 46 1593–1603 (2018).
- 756 28. Wolff, M. *et al.* Isoprenoid biosynthesis via the methylerythritol phosphate pathway: The
757 (E)-4-hydroxy-3-methylbut-2-enyl diphosphate reductase (LytB/IspH) from *Escherichia*
758 *coli* is a [4Fe-4S] protein. *FEBS Lett.* **541**, 115–120 (2003).
- 759 29. Atkinson, J. T., Campbell, I., Bennett, G. N. & Silberg, J. J. Cellular Assays for
760 Ferredoxins: A Strategy for Understanding Electron Flow through Protein Carriers That
761 Link Metabolic Pathways. *Biochemistry* **55**, 7047–7064 (2016).
- 762 30. Nakamura, M., Saeki, K. & Takahashi, Y. Hyperproduction of recombinant ferredoxins in
763 *Escherichia coli* by coexpression of the ORF1-ORF2-iscS-iscU-iscA-hscB-hscA-fdx-ORF3
764 gene cluster. *J. Biochem.* **126**, 10–18 (1999).
- 765 31. McCarthy, E. L. & Booker, S. J. Destruction and reformation of an iron-sulfur cluster
766 during catalysis by lipoyl synthase. *Science (80-.).* **358**, 373–377 (2017).
- 767 32. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: A new generation of protein
768 database search programs. *Nucleic Acids Research* vol. 25 3389–3402 (1997).
- 769 33. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7:
770 Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
- 771 34. McClure, M. A., Smith, C. & Elton, P. Parameterization studies for the SAM and HMMER
772 methods of hidden Markov model generation. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **4**,
773 155–164 (1996).
- 774 35. Larsson, A. AliView: A fast and lightweight alignment viewer and editor for large datasets.
775 *Bioinformatics* **30**, 3276–3278 (2014).

- 776 36. Criscuolo, A. & Gribaldo, S. BMGE (Block Mapping and Gathering with Entropy): A new
777 software for selection of phylogenetic informative regions from multiple sequence
778 alignments. *BMC Evol. Biol.* **10**, (2010).
- 779 37. Nguyen, L. T., Schmidt, H. A., Von Haeseler, A. & Minh, B. Q. IQ-TREE: A fast and
780 effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol.*
781 *Evol.* **32**, 268–274 (2015).
- 782 38. Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood
783 phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
- 784 39. Lemoine, F. *et al.* Renewing Felsenstein’s phylogenetic bootstrap in the era of big data.
785 *Nature* **556**, 452–456 (2018).
- 786 40. Balaban, M., Moshiri, N., Mai, U., Jia, X. & Mirarab, S. TreeCluster: Clustering biological
787 sequences using phylogenetic trees. *PLoS One* **14**, (2019).
- 788 41. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python.
789 *Nat. Methods* **17**, 261–272 (2020).
- 790 42. Peralta-Yahya, P. P. *et al.* Identification and microbial production of a terpene-based
791 advanced biofuel. *Nat. Commun.* **2**, (2011).
- 792 43. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in
793 *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 6640–6645
794 (2000).
- 795 44. Lee, T. *et al.* BglBrick vectors and datasheets: A synthetic biology platform for gene
796 expression. *J. Biol. Eng.* **5**, 12 (2011).
- 797 45. Neidhardt, F. C., Bloch, P. L. & Smith, D. F. Culture Medium for Enterobacteria. *J.*

- 798 *Bacteriol.* **119**, 736–747 (1974).
- 799 46. Loiseau, L. *et al.* ErpA, an iron-sulfur (Fe-S) protein of the A-type essential for respiratory
800 metabolism in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 13626–13631 (2007).
- 801 47. Elliott, J. I. & Brewer, J. M. The inactivation of yeast enolase by 2,3-butanedione. *Arch.*
802 *Biochem. Biophys.* **190**, 351–357 (1978).
- 803 48. Fish, W. W. Rapid Colorimetric Micromethod for the Quantitation of Complexed Iron in
804 Biological Samples. *Methods Enzymol.* **158**, 357–364 (1988).
- 805 49. Fourmond, V. *et al.* SOAS: A free program to analyze electrochemical data and other
806 one-dimensional signals. *Bioelectrochemistry* **76**, 141–147 (2009).
- 807