Ubiquinone biosynthesis over the entire O_2 range: characterization of a conserved, O_2-independent pathway

Ludovic Pelosi[1¶], Chau-Duy-Tam Vo[2¶], Sophie Abby[1], Laurent Loiseau[3], Bérengère Rascalou[1], Mahmoud Hajj Chehade[1], Bruno Faivre[2], Mathieu Goussé[1], Clothilde Chenal[1], Nadia Touati[4], Laurent Binet[4], David Cornu[5], Cameron David Fyfe[2], Marc Fontecave[2], Frédéric Barras[3, 6], Murielle Lombard[2]*, Fabien Pierrel[1]*

[1] Univ. Grenoble Alpes, CNRS, CHU Grenoble Alpes, Grenoble INP, TIMC-IMAG, F-38000 Grenoble, France

[2] Laboratoire de Chimie des Processus Biologiques, Collège de France, Université Pierre et Marie Curie, CNRS UMR 8229, PSL Research University, 11 place Marcelin Berthelot, 75005 Paris, France.

[3] Aix Marseille Université, CNRS, Laboratoire Chimie Bactérienne, Institut Microbiologie de la Méditerranée, 31 Chemin Joseph Aiguier, Marseille 13009, France

[4] ENSCP-Chimie ParisTech, Institut de Recherche de Chimie Paris, CNRS UMR 8247, 11 rue Pierre et Marie Curie, 75005 Paris, France

[5] Plateforme SICaPS, Institut de Biologie Intégrative de la cellule (I2BC), Bâtiment 21, Avenue de la Terrasse, 91 190 Gif-sur-Yvette, France

[6] SAMe Unit, Department de Microbiologie, Institut Pasteur, 25 Rue du Dr Roux, 75015 Paris, France

[¶] These authors contributed equally to the work

* Correspondence: fabien.pierrel@univ-grenoble-alpes.fr, murielle.lombard@college-de-france.fr

Keywords:

Ubiquinone, iron-sulfur, anaerobic metabolism, hydroxylation, bioenergetics, proteobacteria, oxygen, peptidase U32, microaerobic respiration, facultative anaerobe

SUMMARY

The majority of cellular ATP is produced by respiratory chains that shuttle electrons from reduced substrates to terminal electron acceptors, via quinone molecules like ubiquinone. Dioxygen (O_2) is a critical parameter of bioenergetics since it is a major terminal electron acceptor and a co-substrate in the biosynthesis of ubiquinone. Here, we characterize a novel, O_2-independent pathway for the biosynthesis of ubiquinone. This pathway relies on three proteins UbiT (YhbT), UbiU (YhbU) and UbiV (YhbV). UbiT contains an SCP2 lipid-binding domain and is likely an accessory factor of the biosynthetic pathway, while UbiU-UbiV are involved in hydroxylation reactions and represent a novel class of O_2-independent hydroxylases. We demonstrate that UbiU-UbiV form a heterodimer, wherein each protein binds a 4Fe-4S cluster via conserved cysteines that are essential for activity. The UbiT, -U, -V proteins are found in α-, β-, γ-proteobacterial clades including several human pathogens, supporting the widespread distribution of a previously-unrecognized capacity to synthesize ubiquinone in the absence of O_2. Together, the O_2-dependent and O_2-independent ubiquinone biosynthesis pathways contribute to optimize bacterial metabolism over the entire O_2 range.

IMPORTANCE

In order to colonize environments with large O_2 gradients or fluctuating O_2 levels, bacteria have developed metabolic responses that remain incompletely understood. Such adaptations have been recently linked to antibiotic resistance, virulence and the capacity to develop in complex ecosystems like the microbiota. Here, we identify a novel pathway for the biosynthesis of ubiquinone, a molecule with a key role in cellular bioenergetics. We link three uncharacterized genes of *Escherichia coli* to this pathway and show that the pathway functions independently from O_2. In contrast, the long-described pathway for ubiquinone biosynthesis requires O_2 as substrate. In fact, we find that many proteobacteria are equipped with the O_2-dependent and O_2-independent pathways, supporting that they are able to synthesize ubiquinone over the entire O_2 range. Overall, we propose that the novel O_2-independent pathway is part of the metabolic plasticity developed by proteobacteria to face varying environmental O_2 levels.

57 **INTRODUCTION**

58 Since the oxygenation of the Earth's atmosphere some 2.3 billion years ago, many organisms
59 adopted dioxygen (O_2) as a terminal electron acceptor of their energy-producing respiratory chains [1].
60 Indeed, oxygenic (aerobic) respiration has a superior energetic output compared to anaerobic respiration
61 or fermentation, which are both O_2-independent processes [1]. In fact, several microorganisms, including
62 many important human pathogens, are facultative anaerobes that are able to adopt either an aerobic or
63 an anaerobic lifestyle depending on the environmental conditions [2,3]. In the laboratory, bacteria are
64 usually cultured and studied under fully aerobic or completely anaerobic conditions (absence of O_2) [4],
65 whereas natural habitats cover the entire range of O_2 concentrations [5]. For instance, large O_2 gradients
66 are typically encountered in the human large intestine, in biofilms or in transition zones between oxic and
67 anoxic environments [5]. Moreover, bacteria can experience rapid transitions between environments with
68 vastly different O_2 contents, as during the infection process of enteric pathogens that progress along the
69 gastrointestinal tract [3].

70 To maximize their bioenergetic capacities according to the various levels of O_2 encountered in
71 their environment, bacteria modulate the composition of their respiratory chains, notably the quinone
72 species and the terminal reductases [3,4,6]. Quinones are lipophilic redox molecules that fuel electrons to
73 terminal reductases, which reduce O_2 whenever available, or alternative electron acceptors for instance
74 nitrate, dimethyl sulfoxide (DMSO), trimethylamine N-oxide [7]. Naphtoquinones (menaquinone (MK) and
75 demethylmenaquinone (DMK)) and ubiquinone (UQ) are the two main groups of bacterial quinones.
76 (D)MK and UQ differ by the nature of their head group and the value of their redox midpoint potential [8].
77 (D)MK are considered "anaerobic quinones" since they function primarily in anaerobic respiration whereas
78 UQ is considered an "aerobic quinone" because its supplies electrons mostly to the reductases that reduce
79 O_2 [1,8,9]. Accordingly, UQ is the main quinone of the facultative anaerobe *Escherichia coli* in aerobic
80 conditions, whereas the naphtoquinones are predominant in the absence of O_2 [10,11], UQ being
81 nevertheless present.

82 The biosynthesis of UQ requires a total of eight reactions to modify the aromatic ring of the
83 precursor, 4-hydroxybenzoic acid (4-HB): one prenylation, one decarboxylation, three hydroxylation and
84 three methylation reactions (Figure 1A) [12]. In addition to the enzymes that catalyze the various steps,
85 three accessory factors - UbiB, UbiJ and UbiK – are also needed. UbiB has an ATPase activity [13] and we
86 showed that UbiJ and UbiK [14,15] belong to a multiprotein UQ biosynthesis complex, in which the SCP2
87 domain (sterol carrier protein 2) of UbiJ binds the hydrophobic UQ biosynthetic intermediates [16]. This

88 UQ biosynthetic pathway is under the dependence of O_2 since all three hydroxylases - UbiI, UbiH and UbiF
89 – use O_2 as a cosubstrate (Figure 1A) [17,18]. We showed recently that other hydroxylases, UbiL, UbiM
90 and Coq7, replace UbiI, UbiH and UbiF in some proteobacteria [19]. The six hydroxylases have in common
91 their dependence to O_2 and thus function in UQ biosynthesis only under aerobic conditions. Interestingly,
92 Alexander and Young established 40 years ago that *E. coli* was able to synthesize UQ anaerobically [20],
93 suggesting the existence of an O_2-independent biosynthesis pathway, which is still uncharacterized.

94 In this study, we describe the O_2-independent UQ biosynthetic pathway in *E. coli* and identify three
95 essential components, the UbiT, UbiU and UbiV proteins, formerly called YhbT, YhbU and YhbV. We show
96 that the O_2-independent UQ biosynthetic pathway is widely conserved in proteobacteria. UbiT likely
97 functions as an accessory factor in the O_2-independent UQ biosynthetic pathway and we show that UbiU
98 and UbiV are involved in at least one O_2-independent hydroxylation reaction. Moreover, we demonstrate
99 that both UbiU and UbiV bind a [4Fe-4S] cluster essential for activity, which identifies these proteins as
100 prototypes of a new class of O_2-independent hydroxylases. Our results highlight that many proteobacterial
101 species use two different and complementary molecular pathways to produce UQ over the entire
102 continuum of environmental 0_2.

103

104 # RESULTS

105 **4-HB is the precursor of UQ synthesized in anaerobic conditions:**

106 4-HB is the precursor of UQ synthesized in aerobic conditions [21]. Accordingly, an *E. coli* Δ*ubiC*
107 mutant impaired in 4-HB biosynthesis [22], is deficient in UQ_8 and is complemented by addition of 4-HB to
108 the growth medium [23]. In order to evaluate whether 4-HB is also the precursor of the O_2-independent
109 UQ biosynthetic pathway, we grew a Δ*ubiC* strain anaerobically. The Δ*ubiC* strain showed a diminished
110 level of UQ_8, which was partially recovered by supplementation with 4-HB (Figure 1B). Furthermore, we
111 grew the Δ*ubiC* strain in a medium supplemented with $^{13}C_7$-4HB and we analyzed the labelling of
112 biosynthesized UQ_8 by HPLC-mass spectrometry. In cells grown in aerobic or anaerobic conditions, the
113 labelled form of UQ ($^{13}C_6$-UQ_8, m/z= 750.4 for the NH_4^+ adduct) represented 98.3% and 97.3% (±0.2%) of
114 the total UQ_8 pool, respectively (Figure 1C, 1D). As expected, Δ*ubiC* cells grown anaerobically with
115 unlabeled 4-HB didn't show any $^{13}C_6$-UQ_8 (Figure 1E). Altogether, these results establish that 4-HB is the
116 precursor of the O_2-independent UQ biosynthetic pathway.

117 **Ubi enzymes, except hydroxylases, are common to the aerobic and anaerobic UQ biosynthesis**
118 **pathways:**

119 The above result suggests that the UQ biosynthetic pathway decorates 4-HB with the same
120 chemical groups irrespective of the presence of environmental O_2. Thus, we evaluated whether the
121 enzymes of the aerobic pathway are also involved in the O_2-independent pathway by measuring the UQ_8
122 content of knock-out (KO) strains grown in aerobic and anaerobic conditions. Deletion of *ubiA*, *ubiE* or
123 *ubiG* abrogated UQ_8 biosynthesis in both conditions whereas Δ*ubiB*, Δ*ubiD* or Δ*ubiX* strains synthesized a
124 limited amount of UQ_8 but only in aerobic conditions (Figure 1F). In contrast, *ubiJ* and *ubiK* had no effect
125 on UQ biosynthesis under anaerobic conditions (Figure 1F).

126 In aerobic conditions, the hydroxylation reactions are catalyzed by the flavin monooxygenases
127 UbiF, UbiH and UbiI that use dioxygen as a co-substrate [17,18,24] (Figure 1A). We previously reported
128 that cells deleted for a single O_2-dependent hydroxylase (Δ*ubiF*, Δ*ubiI* or Δ*ubiH*) were deficient in UQ, when
129 cultured in the presence of air, but synthesized UQ in anaerobic conditions [17], consistent with the
130 existence of an alternative hydroxylation system in the O_2-independent pathway [20]. Indeed, we
131 confirmed that all three FMOs are dispensable for the O_2-independent UQ biosynthetic pathway since a
132 triple mutant Δ*ubiF* Δ*ubiI* Δ*ubiH* was deficient for UQ when grown in air, but synthesized wild-type (WT)
133 level of UQ_8 in anaerobic conditions (Figure 1G). Together, our results demonstrate that the O_2-dependent

134 and O_2-independent UQ biosynthetic pathways share the enzymes involved in the prenylation (UbiA),
135 decarboxylation (UbiX and UbiD) and methylation reactions (UbiE and UbiG), but differ by their
136 hydroxylases and the accessory factors UbiJ and UbiK (Figure 1H).

137 **Identification of three genes required for UQ biosynthesis in anaerobic conditions:**

138 To identify genes involved in the O_2-independent UQ biosynthetic pathway, we cultivated
139 anaerobically a collection of ~ 200 *E. coli* strains that contained deletions covering multiple ORFs [25,26]
140 and we analyzed their UQ_8 content by HPLC-electrochemical detection (HPLC-ECD). We found a complete
141 absence of UQ_8 in strains ME4561, ME5034 and ME4746 that carry deletions encompassing, *ubiE-ubiJ-*
142 *ubiB-ubiD*, *ubiG,* and *ubiX*, respectively (Table S1). Several other strains had a low UQ_8 content and a poor
143 growth in synthetic medium supplemented with glycerol and nitrate (SMGN). However, those strains
144 showed better growth and higher UQ_8 content in LB medium (Table S1). Thus we did not investigate them
145 further, as a genetic defect affecting directly the O_2-independent UQ pathway was unlikely. In contrast,
146 ME4641 showed a profound UQ_8 deficiency and robust anaerobic growth in LB and SMGN media (Table
147 S1). Importantly, ME4641 had a WT UQ_8 level when grown aerobically, suggesting that only the O_2-
148 independent pathway was altered (Figure 2A). ME4641 contains a deletion named OCL30-2 that covers 9
149 genes, 5 of them lacking an identified function (Figure 2B). To find the candidate gene involved in the
150 anaerobic biosynthesis of UQ, we obtained 8 single gene KO strains from the Keio collection [27] and
151 analyzed their quinone content after anaerobic growth (Figure 2C). The Δ*yhbT* and Δ*yhbU* strains were
152 strongly deficient in UQ_8. We then transduced the Δ*yhbT* and Δ*yhbU* mutations from the Keio strains into
153 a MG1655 genetic background and also constructed the Δ*yhbV* strain, which was not available in the Keio
154 collection. We found that all three strains had very low levels of UQ_8 when grown in anaerobic conditions
155 but showed normal levels after aerobic growth (Figure 2D, Table 1). In addition, the mutant strains showed
156 a two-fold decrease in MK_8 and a two-fold increase in DMK_8 after anaerobic growth (Table 1). This effect
157 might indirectly result from the UQ_8 deficiency as it was also observed in the Δ*ubiG* strain (data not shown).

158

culture condition	strain	UQ_8 (pmoles/mg cells)	DMK_8 (247nm peak area/mg cells)	MK_8 (247nm peak area/mg cells)
Air	WT	208 ± 9,4	0,94 ± 0,26	0,15 ± 0,09
Air	Δ*yhbT*	193,9 ± 12,4	1,34 ± 0,05	0,13 ± 0,03
Air	Δ*yhbU*	233,8 ± 16,2	0,9 ± 0,26	0,16 ± 0,09
Air	Δ*yhbV*	227,9 ± 25,1	0,93 ± 0,15	0,13 ± 0,04
anaerobic	WT	114,2 ± 19,1	1,28 ± 0,4	0,41 ± 0,15
anaerobic	Δ*yhbT*	3,3 ± 0,4	2,88 ± 0,47	0,17 ± 0,01
anaerobic	Δ*yhbU*	6,1 ± 3,3	2,55 ± 0,55	0,19 ± 0,01
anaerobic	Δ*yhbV*	7,3 ± 4,7	2,61 ± 0,49	0,21 ± 0,01
anaerobic	Δ*yhbT*+pK	2 ± 0,4	2,46 ± 0,33	0,16 ± 0,04
anaerobic	Δ*yhbT*+pK-yhbT	64,6 ± 16,1	1,77 ± 0,29	0,3 ± 0,04
anaerobic	Δ*yhbU*+pBAD	1 ± 0,6	2,29 ± 0,58	0,18 ± 0,04
anaerobic	Δ*yhbU*+pBAD-yhbU	82 ± 15,9	1,2 ± 0,18	0,27 ± 0,04
anaerobic	Δ*yhbV*+pBAD	2,2 ± 3,5	2,14 ± 0,15	0,23 ± 0,01
anaerobic	Δ*yhbV*+pBAD-yhbV	120,4 ± 18,5	1,87 ± 0,19	0,42 ± 0,06

Table 1: Quinone content of WT and Δ*yhb* strains cultured in SMGN either aerobically or anaerobically, n=3-5. For strains containing pK or pBAD vectors, SMGN was supplemented with 0.02% arabinose.

Deletion of *yhbT*, *yhbU* or *yhbV* causes UQ_8 deficiency specifically in anaerobic conditions:

We then transformed the mutant strains with an empty vector or a vector carrying a WT allele of the studied gene. In *yhb* KO strains expressing the corresponding gene from the plasmid, we observed a complementation of the UQ_8 deficiency (Table 1) and a normalization of the levels of octaprenylphenol (OPP), an early UQ_8 biosynthetic intermediate (Figure 2E). The ~3 fold elevation of OPP in the *yhbT, -U, -V* KO mutants suggested that the O_2-independent UQ biosynthetic pathway was blocked downstream of OPP in these strains. In these experiments, no cross-complementation was observed - for example the plasmid with *yhbV* had no effect in Δ*yhbT* or Δ*yhbU* strains - suggesting the absence of redundancy in the function of each gene. We also measured a profound UQ_8 deficiency when the Δ*yhbT,* Δ*yhbU* and Δ*yhbV* strains were grown anaerobically in various media (glycerol + DMSO, lactate + KNO_3) (Figure 2F), showing that the UQ_8 biosynthetic defect is not linked to a particular carbon source or electron acceptor. Altogether, our results demonstrate that the *yhbT*, *yhbU* and *yhbV* genes are part of the O_2-independent UQ biosynthetic pathway, so we propose to rename them *ubiT*, *ubiU* and *ubiV*, respectively.

177 ***ubiT,-U,-V* strongly co-occur, exclusively in genomes with potential for UQ biosynthesis**

178 We investigated the distribution of *ubiT*, -*U* and -*V* in a large genome dataset gathering 5750
179 genomes of bacteria and archaea. We found no evidence of genomes harboring matches for more than
180 one of the three genes of interest outside of Alphaproteobacteria, Betaproteobacteria and
181 Gammaproteobacteria (α-, β-, γ-proteobacteria) and three genomes of Acidithiobacillia (Table S2A).
182 Interestingly, the three former classes are the only known so far to be able to produce UQ [1,8], consistent
183 with a specific link between UbiT, -U, -V and UQ biosynthesis.

184 This was confirmed by analyzing in more details the distribution of *ubiT*, -*U* and -*V* and of three
185 marker genes of the UQ biosynthetic pathway (*ubiA*, *ubiE* and *ubiG)* in 611 representative and reference
186 genomes of α-, β-, γ-proteobacteria (Table S2B). We chose *ubiA, -G, -E* because they have a higher
187 conservation compared to other genes like *ubiC, -D, -X* [28] and because they are part of the O_2-dependent
188 and O_2-independent pathways (Figure 1H). 575 genomes had positive matches for *ubiA, -G, -E* and 589 for
189 at least two of them. Regarding the distribution of *ubiT*, -*U* and -*V*, 221 genomes had matches for at least
190 one of the three genes, and in 210 cases (95%), the three genes were present. Importantly, all genomes
191 with *ubiT*, -*U* and -*V* harbored at least two of the three marker genes for the UQ biosynthetic pathway
192 (Table S2B). In addition, we found that 22 out of the 29 proteobacterial orders analyzed had up to 50%
193 genomes harboring a complete set of the *ubiT*, -*U* and -*V* genes (Figure 3A), demonstrating a wide
194 taxonomic distribution. Overall, our analysis indicates a strong pattern of co-occurrence of *ubiT*, -*U* and -
195 *V*, and demonstrates that they are uniquely found in genomes showing signs of UQ production.

196 We found *ubiA, -G, -E* and O_2-dependent hydroxylase genes in 570 genomes, of which 204 also
197 contained *ubiT*, -*U* and –*V* (Table S2B and Figure 3B). Only 3 species, *Phaeospirillum fulvum*,
198 *Magnetococcus marinus* and *Oxalobacter formigenes*, seem to rely exclusively on the O_2-independent
199 pathway for UQ production as their genomes contain the *ubiT*, -*U* and -*V* genes but no O_2-dependent
200 hydroxylases (Figure 3B). We noticed that the *ubiT*, -*U* and -*V* genes are found next to other *ubi* genes in
201 *M. marinus* (Figure S1). Interestingly, *P. fulvum* and *M. marinus* were described to synthesize UQ in
202 microaerobic conditions [29-31] and *O. formigenes* has been documented as an obligate anaerobe [32].
203 Together, our results show that the O_2-independent UQ biosynthesis pathway is widespread in α-, β-, γ-
204 proteobacterial orders and co-exists with the O_2-dependent pathway in 98% of the cases.

205 We then looked into the relative positioning of *ubiT*, -*U* and -*V* in the 210 genomes harboring the
206 three genes. We found 106 cases, covering α-, β-, γ-proteobacterial orders, where they were encoded next

207 to each other ("3-genes loci") and 82 cases of a 2-genes locus with the third gene being encoded elsewhere
208 in the genome (Figure 3C). Evaluation of the genetic architecture of *ubiT*, *-U* and *–V* revealed that *ubiU* and
209 *ubiV* were found exactly next to each other in 69% of the loci (all 3-genes loci and in 39/82 of the 2-genes
210 loci) and that the three genes were encoded in three separate parts of the genome only in 21 cases (Figure
211 3C). Interestingly, as an additional support for their involvement in a same function, we found an example
212 of a gene fusion between *ubiT* and *ubiU* in two genomes from *Zymomonas mobilis* strains (α-
213 proteobacteria), which also contain an *ubiV* gene directly downstream of the fused gene (Figure 3C).

214 **UbiT is an SCP2 protein and UbiU-V are required for the O_2-independent hydroxylation of DMQ_8**

215 We then analyzed the sequences of the UbiT, UbiU and UbiV proteins. The major part of UbiT
216 (amino acids 45-133, on a total of 174 aa in the *E. coli* protein) corresponds to a SCP2 domain (PFAM:
217 PF02036, Figure S2). SCP2 domains typically form a hydrophobic cavity that binds various lipids [33] and
218 our sequence alignment indeed showed the conservation of hydrophobic amino acids at several positions
219 in the SCP2 domain of UbiT (Figure S2). Recently, we reported that UbiJ binds UQ biosynthetic
220 intermediates in its SCP2 domain and organizes a multiprotein complex composed of several Ubi enzymes
221 [16]. We propose that UbiT and its SCP2 domain may fulfill similar functions in the O_2-independent UQ
222 pathway, as UbiJ is required exclusively for the O_2-dependent biosynthesis of UQ (Figure 1F).

223 UbiU and UbiV have ~330 and 300 aa respectively, and contain an uncharacterized motif called
224 peptidase U32 (PF01136) (Figure S3 and S4). Since only the hydroxylation reactions are uncharacterized in
225 the O_2-independent pathway (Figure 1H), we hypothesized that UbiU and UbiV might function in these
226 steps. To test our hypothesis, we developed an *in vivo* assay based on the O_2-independent conversion of
227 labeled DMQ_8 into labeled UQ_8. This assay monitors the C6-hydroxylation and the subsequent O6-
228 methylation (Figure 4A). Δ*ubiC* Δ*ubiF* cells grown aerobically with $^{13}C_7$-4HB synthesized DMQ_8, 73% of
229 which was labelled with $^{13}C_6$. Upon transfer to anaerobic conditions, the cells gradually converted a
230 significant part of ($^{13}C_6$)-DMQ_8 into ($^{13}C_6$)-UQ_8 (Figure 4B). Inactivation of either *ubiU* or *ubiV* in Δ*ubiC* Δ*ubiF*
231 cells did not perturb the accumulation of ($^{13}C_6$)-DMQ_8 but prevented its conversion into ($^{13}C_6$)-UQ_8 (Figure
232 4C-D). This result demonstrates that UbiU and UbiV are essential for the C6-hydroxylation reaction of the
233 O_2-independent UQ biosynthetic pathway.

234 **UbiV contains a [4Fe-4S] cluster**

235 To get insights into the potential presence of cofactors in UbiU and UbiV, we attempted to
236 characterize them biochemically. UbiU was not soluble but we purified $UbiV_{6His}$, which behaved as a

monomer in solution (Figure S5A-B). UbiV was slightly pink-colored and had a UV-visible absorption spectrum with features in the 350-550 nm region (Figure 5A, dotted line), suggesting the presence of iron-sulfur (Fe-S) species [34,35]. We indeed detected substoichiometric amounts of iron and sulfur (0.2 Fe and 0.2 S/monomer), indicating a potential oxidative degradation of the Fe-S cluster during aerobic purification, as already observed with many other Fe-S proteins [36,37]. Consistent with this hypothesis, anaerobic reconstitution of the Fe-S cluster yielded a UbiV protein with 3.9 iron and 3.3 sulfur/monomer (Table 2) and with a UV-Vis spectrum characteristic of a $[4Fe\text{-}4S]^{2+}$ cluster [38] (Figure 5A, solid line), that was affected by exposure to air (Figure S5C). The electron paramagnetic resonance (EPR) spectrum of the cluster reduced anaerobically displayed features characteristic of a $[4Fe\text{-}4S]^{1+}$ cluster in the S = 1/2 state (Figure 5B) [37,39]. Overall, we conclude that, under anaerobic conditions, UbiV is able to bind one air-sensitive, redox-active [4Fe-4S] cluster.

Fe-S clusters are typically coordinated by cysteine residues [40,41] and we obtained evidence that the [4Fe-4S] cluster in UbiV is coordinated by four conserved cysteines arranged in a $CX_nCX_{12}CX_3C$ motif (Figure S4). Indeed, combinatorial elimination of C39, C180, C193, and C197 in double, triple and quadruple mutants resulted in proteins incapable of binding [Fe-S] clusters *in vivo*, as shown by the absence of absorption bands in the 350-550nm region of their UV-vis spectra (Figure S5D). Furthermore, after anaerobic reconstitution of the cluster, the Fe and S contents, and the absorbance at 410 nm were largely decreased in the double and triple mutants, and were undetectable in the quadruple mutant (Table 2 and Figure 5C). At last, we found that the mutation of C180 or C193 altered the function of UbiV *in vivo* (Figure 5D), suggesting that the [4Fe-4S] cluster is important for activity.

	A_{280}/A_{410}	Iron content (nmoles/nmole of protein)	Sulfur content (nmoles/nmole of protein)
UbiV WT	5.8	3.9 ± 0.13	3.3 ± 0.18
UbiV C193A C197A	8.4	2.4 ± 0.02	2.2 ± 0.13
UbiV C39A C193A C197A	15.5	1.9 ± 0.01	1.5 ± 0.4
UbiV C180A C193A C197A	26.7	0.7 ± 0.05	0.6 ± 0.1
UbiV C39A C180A C193A C197A	57.4	0.1 ± 0.04	0
UbiU WT – UbiV WT	4.3	7.9 ± 0.05	7.5 ± 0.06
UbiU C169A C176A – UbiV WT	7.3	4.7 ± 0.03	5.8 ± 0.1
UbiU C193A C232A – UbiV WT	8	3.9 ± 0.05	5.8 ± 0.1

Table 2: Characterization of UbiV proteins and UbiU-UbiV heterodimeric complexes (wild-type and mutants): metal content, and UV-Vis properties after attempts of anaerobic reconstitution of their Fe-S centers.

267 **UbiU contains a [4Fe-4S] cluster and forms a complex with UbiV**

268 We succeeded purifying UbiU after coexpressing it with $UbiV_{6His}$ (Figure S6A). The two proteins
269 copurified in the form of a heterodimer (Figure S6A-B) that showed traces of Fe-S clusters (Figure 6A,
270 dotted line), with substoichiometric amounts of iron and sulfide (0.4 Fe and 0.4 S/heterodimer).
271 Reconstitution with iron and sulfide yielded a heterodimer with about 8 iron and 8 sulfur (Table 2) and a
272 UV-visible spectrum characteristic of $[4Fe\text{-}4S]^{2+}$ clusters (Figure 6A, solid line). The EPR spectrum of
273 reduced UbiU-UbiV was also consistent with the presence of 2 different [4Fe-4S] clusters since it showed
274 a composite signal, which reflects the presence of two different S=1/2 species (Figure 6B).

275 Four strictly conserved cysteines are also found in UbiU (Figure S3) and we hypothesized that they
276 might bind the [4Fe-4S]. We eliminated these cysteines in pairs and purified heterodimers composed of
277 WT $UbiV_{6His}$ and mutant UbiU (Figure S6C). After reconstitution, UbiU C169A C176A-UbiV and UbiU C193A
278 C232A-UbiV had about twice less iron than the WT heterodimer (Table 2), and their A_{280}/A_{410} ratio were
279 also diminished about two-fold (Figure 6C and Table 2). Altogether, our data clearly demonstrate that each
280 protein of the heterodimeric UbiU-UbiV complex binds one [4Fe-4S] cluster and that the iron-chelating
281 cysteines in UbiU are C169, C176, C193 and C232. Finally, an *in vivo* complementation assay demonstrated
282 that C176 was important for the function of UbiU (Figure 6D).

283 **Many U32 proteases display motifs of four conserved cysteines**

284 The presence of Fe-S clusters in UbiU and UbiV, two U32 proteases family members, led us to
285 evaluate the presence of conserved Cys motifs in other U32 proteins. Kimura *et al.* reported a phylogenetic
286 tree of 3521 peptidase U32 domains which formed 12 groups, belonging to 10 protein families [42]. We
287 extracted and aligned the sequences of the 10 protein families and found highly conserved 4-cysteines
288 clusters (97-100% conservation) in eight of them (Figure 7), suggesting an important functional role for
289 these residues. Only families PepU32#5 and PepU32#6 had respectively none, and two mildly (60-80%)
290 plus three poorly conserved cysteines in their sequences (40-65%). The cysteine motifs for each of the
291 eight families showed a high degree of conservation, and strikingly, most of them could even be aligned
292 with each other, the $CX_6CX_{15}CX_{3/4}C$ patterns appearing recurrently (Figure 7). To be noted, UbiV had a
293 slightly distinct motif, with the first cysteine occurring much upstream of the three others and outside of
294 the U32 domain. Overall, our data suggest that most members of the U32 peptidase family may contain a
295 4Fe-4S cluster coordinated by conserved cysteines, similar to what we demonstrated for UbiU and UbiV.

296

297 # DISCUSSION

298 **The O_2-independent UQ biosynthesis pathway is widespread in proteobacteria:**

299 The evidence for an O_2-independent synthesis of UQ by *E. coli* was reported more than forty years
300 ago [20], yet this pathway remained uncharacterized until now. Circumstantial evidence had been
301 obtained that a few species – limited, to our knowledge, to *E. coli* [20], *Rhodobacter sphaeroides* [43],
302 *Paracoccus denitrificans* [44], *Halorhodospira halophila* [45] – were able to synthesize UQ in anaerobic
303 conditions, as demonstrated by biochemical measurements of the quinone content of cells grown
304 anaerobically. Here, we demonstrate that the O_2-dependent and O_2-independent UQ biosynthesis
305 pathways differ only by three hydroxylation steps (Figure 1), and we identify three genes, *ubiT,-U,-V* that
306 are essential for the O_2-independent biosynthesis of UQ in *E. coli* (Figure 2). The facts that the UbiT, -U, -V
307 proteins are widespread in α, β, γ–proteobacterial clades (Figure 3) and co-occur with UbiA, -E, -G enzymes
308 (Table S2B), reveal UbiT, -U, -V as key elements of the broadly-distributed, O_2-independent UQ pathway.
309 Overall, our data support that many proteobacteria have the previously-unrecognized capacity to
310 synthesize UQ independently from O_2.

311 **Physiological possibilities offered by UQ biosynthesis over the entire O_2 range:**

312 In our set of reference genomes, only three species (*P. fulvum*, *M. marinus* and *O. formigenes*)
313 seem to rely exclusively on the O_2-independent pathway for UQ production (Table S2B). Indeed, the vast
314 majority of proteobacteria with the O_2-independent UQ biosynthesis pathway also possess the O_2-
315 dependent hydroxylases of the aerobic pathway (207 out of 210) (Figure 3B). This result supports that both
316 pathways confer physiological advantages, allowing production of UQ over the entire spectrum of O_2 levels
317 encountered by facultative aerobes.

318 So-called microaerobes, able to respire O_2 in microaerobic conditions, are very abundant in Nature
319 [5] and *E. coli* is known to respire nanomolar O_2 concentrations [46]. To sustain O_2 respiration in the
320 microaerobic range, these organisms are equipped with high affinity O_2 reductases [5,46,47]. These
321 enzymes reduce efficiently the low levels of environmental O_2 present at the cell's membrane, leaving the
322 cytoplasm devoid of any O_2 [48]. Under such conditions, UQ - which is the main electron donor for the high
323 affinity O_2 reductases *bd*I and *bd*II of *E. coli* [9] - must therefore be synthesized via the O_2-independent
324 pathway. Overall, we believe that the O_2-independent UQ biosynthesis pathway operates not only in
325 anaerobic conditions but also in microaerobic conditions, in which UQ is likely crucial for bacterial
326 physiology.

327 The O_2-independent UQ biosynthesis pathway may also confer a significant advantage to
328 facultative bacteria in case of a rapid transition from an anaerobic to an aerobic environment. Indeed,
329 anaerobic biosynthesis of UQ will result in cellular membranes containing UQ at the time of the transition,
330 allowing an immediate switch to the energetically favorable metabolism of O_2 respiration. Our
331 identification of the anaerobic UQ pathway provides the unique opportunity to selectively disrupt UQ
332 biosynthesis depending on O_2 levels and should foster new research on bacterial physiology in the
333 microaerobic range. Indeed, apart from *E. coli*, which was thoroughly studied over the microaerobic range
334 [4,48,49], details on bacterial physiology in microaerobiosis are scarce.

335 ***ubiT,U,V* mutants and pathogenicity**:

336 In addition to bioenergetics per say, anaerobic and microaerobic respirations are thought to be
337 important for pathogenicity [3,50]. Interestingly, homologs of UbiT, -U, -V have been linked to
338 pathogenicity in several bacterial models. Indeed, the inactivation of *ubiU-V* homologs in *Proteus mirabilis*
339 leads to a decreased infection of the urinary tract of mice [51] and to a diminished virulence of *Yersinia*
340 *ruckeri* [52], a pathogen that develops in the gut of fish, an environment with a notoriously low O_2 content.
341 Furthermore, inactivation of PA3911 [53] (*ubiT*) and PA3912-PA3913 [54] (*ubiU-V*) in *Pseudomonas*
342 *aeruginosa*, abolished nitrate respiration, the main anaerobic metabolism used by the bacterium in the
343 lungs of cystic fibrosis patients [55,56]. Based on our results showing that the deletion of *ubiT*, *ubiU* or
344 *ubiV* abrogates the O_2-independent biosynthesis of UQ in *E. coli*, we suggest that the attenuation of the
345 mutants discussed above results from their UQ deficiency in microaerobic / anaerobic conditions.

346 **Proposed roles for UbiT, UbiU and UbiV:**

347 UbiT possesses an SCP2 domain, similar to UbiJ that we recently demonstrated to be an accessory
348 factor that binds the hydrophobic UQ biosynthetic intermediates and structures a multiprotein Ubi
349 complex [16]. Since UbiJ functions exclusively in the O_2-dependent pathway, whereas UbiT is important
350 only for the O_2-independent pathway, we propose that UbiT may fulfill, in anaerobiosis, the same functions
351 than UbiJ in aerobiosis. Whether UbiT is part of a complex and is able to bind UQ biosynthetic
352 intermediates, will be addressed in future studies. Interestingly, PA3911, the homolog of UbiT in *P.*
353 *aeruginosa*, was recently shown to bind phosphatidic acid [53], demonstrating an affinity of UbiT for lipid
354 molecules.

355 UbiU and UbiV form a tight heterodimer suggesting that the proteins function together, as further
356 supported by the fact that the *ubiU* and *ubiV* genes co-occur in genomes in 99% of the cases and that they

357 are mostly found next to each other (Figure 3). We demonstrated that UbiU and UbiV are both required
358 for the O_2-independent C6-hydroxylation of DMQ and the accumulation of OPP in Δ*ubiU* or Δ*ubiV* mutants
359 suggests that the two proteins may also function in C5-hydroxylation. We want to emphasize our recent
360 demonstration that a single hydroxylase catalyzes all three hydroxylation steps in the O_2-dependent UQ
361 pathway of *Neisseria meningitidis* [19]. This result showed that three different enzymes are not necessarily
362 required and opens the possibility that UbiU-V may in fact catalyze all three hydroxylation reactions of the
363 O_2-independent UQ biosynthesis pathway. Establishing the hydroxylase activity and the regioselectivity of
364 UbiU-V will require the development of an *in vitro* assay, a challenging task given that the oxygen donor
365 of the reaction is currently unknown and that the substrates are not commercial and highly hydrophobic.
366 Of note, such assay as not yet been reported for the long-discovered O_2-dependent hydroxylases.

367 **Fe-S clusters in UbiU-V and other members of the U32 peptidase family:**

368 Up to now, members of the peptidase family U32 were not shown to bind Fe-S clusters or to
369 contain any of the ~30 cysteine motifs found in well-characterized iron-sulfur proteins [57]. The
370 expression, purification and spectroscopic characterization of UbiV and of the UbiU-UbiV heterodimeric
371 complex clearly showed that each protein contains one [4Fe-4S] cluster (Figure 5 and Figure 6). Mutation
372 of the candidate cysteine ligands, arranged in a $CX_6CX_{16}CX_{38}C$ motif in UbiU and in a $CX_nCX_{12}CX_3C$ motif in
373 UbiV disrupted Fe-S binding and abolished *in vivo* complementation, suggesting a crucial function of the
374 [4Fe-4S] clusters in these proteins. The conservation of a $CX_6CX_{15}CX_{3/4}C$ motif in other U32 proteases
375 supports that these proteins likely bind Fe-S clusters. This hypothesis should guide and stimulate
376 investigations of U32 members, very few of which have currently an established molecular function ([58],
377 http://www.ebi.ac.uk/merops/). Interestingly, RlhA, a member of the U32 proteases family involved in the
378 C-hydroxylation of a cytidine on *E. coli* 23S rRNA, was recently shown to be connected to iron metabolism
379 [42], corroborating our suggestion that RlhA is also an Fe-S protein.

380 In biological systems, Fe-S clusters are mainly known to be involved in electron transfer reactions,
381 but also in substrate binding and activation, in transcription regulation, in iron storage, and as a sulfur
382 donor [40,59,60]. The role of the Fe-S clusters in UbiU and UbiV is unknown at this stage. Our current
383 working hypothesis is a role as an electron transfer chain between the substrate (the UQ biosynthetic
384 intermediate to be hydroxylated) and an unidentified electron acceptor required for the activation of the
385 substrate. Clearly, the Fe-S clusters of UbiU-V are distinct from the molybdenum cofactor present in
386 molybdenum-containing hydroxylases, the only family currently known to catalyze O_2-independent
387 hydroxylation reactions [61]. Together, our results identify UbiU and UbiV as prototypes of a novel class

388 of O_2-independent hydroxylases and extend the framework of the chemically-fascinating O_2-independent
389 hydroxylation reactions.

390

391 ## MATERIALS and METHODS

392 **Culture conditions:**

393 *E. coli* strains were grown at 37 °C in lysogeny broth (LB) medium or in synthetic medium (SM) containing
394 either 0.4% glycerol (w/v), 0.4% lactate (w/v) or 0.2% glucose (w/v) as carbon sources. Autoclaved SM
395 medium was supplemented with 0.5% casaminoacids (w/v) and with 1/100 volume of a filter-sterile
396 solution of 1mM $CaCl_2$, 200 mM $MgCl_2$, 1% thiamine (w/v) [62]. Ampicillin (50 mg/L), kanamycin (25 mg/L),
397 and chloramphenicol (25 mg/L) were added from stocks (1000X solution sterilized through 0.22 µm filters
398 and stored at -20°C), when needed. External electron acceptors like KNO_3 (100 mM) or dimethylsulfoxide
399 (DMSO, 50 mM) were added to SM medium for anaerobic cultures, which were performed in Hungate
400 tubes containing 12 mL medium deoxygenated by argon bubbling (O_2<0.1 ppm) before autoclave. Hungate
401 tubes were inoculated through the septum with 100 µL of overnight precultures taken with disposable
402 syringe and needles from closed Eppendorf tubes filled to the top. Aerobic cultures were performed in
403 Erlenmeyer flasks filled to 1/10 of the maximal volume and shaken at 180 rpm.

404 For the initial screen, we grew ME strains anaerobically in SMGN (SM medium supplemented with glycerol
405 and nitrate). Strains that presented a severe growth defect or a low UQ_8 content were subsequently grown
406 anaerobically in LB medium.

407 **Strain construction:**

408 Strains used in this study are listed in table S3. We obtained the collection of *E. coli* strains containing large
409 and medium deletions from the National BioResource Project, National Institute of Genetics, Japan
410 (http://www.shigen.nig.ac.jp/ecoli/pec/).

411 The Δ*ubiA*::*cat*, Δ*ubiD*::*cat*, Δ*ubiT*::*cat* and Δ*ubiV*::*cat* mutations were constructed in a one-step
412 inactivation of *ubi* genes as described [63]. A DNA fragment containing the *cat* gene flanked with a 5' and
413 3' region bordering the *E.coli ubi* was amplified by PCR using pKD3 as a template and oligonucleotides
414 5wanner and 3wanner (Table S3). Strain BW25113, carrying the pKD46 plasmid, was transformed by
415 electroporation with the amplified fragment and cat^R colonies were selected. The replacement of
416 chromosomal *ubi* by *cat* gene was checked by PCR amplification in the cat^R clones. All mutations were
417 introduced into MG1655 strains by P1 *vir* transduction [64], selecting for the appropriate antibiotic
418 resistance. The antibiotic resistance cassettes were eliminated when needed using plasmid pCP20 as
419 described [65].

420 **Plasmid construction:**

421 All plasmids generated in this study were verified by DNA sequencing. The *yhbU, yhbT and yhbV* inserts
422 (UniProtKB: P45527, P64599, P45475) were obtained by PCR amplification using *E. coli* MG1655 as
423 template and the oligonucleotides pairs *yhbU5*/*yhbU3, yhbT5*/*yhbT3* and *yhbV5*/*yhbV3*, respectively. *Yhb*
424 inserts were EcoRI-SalI digested and inserted into EcoRI-SalI-digested pBAD24 plasmids, yielding the pBAD-
425 *yhbU*, pK-*yhbT* or pBAD-*yhbV* plasmids, respectively.

426 To create a plasmid expressing the *ubiV* (*yhbV*) ORF as C-terminally His-tagged protein, the *ubiV* gene was
427 amplified using pET-22-UbiV-FW (introducing the NdeI site) and pET-22-UbiV-RV (introducing the XhoI site)
428 as primers and pBAD-*yhbV* as template. The NdeI and XhoI digested amplicon was ligated to NdeI and XhoI
429 digested pET-22b(+) plasmid to obtain pET-22-UbiV.

430 The plasmid pETDUET-UbiUV containing UbiU in MCS1 and UbiV in MCS2 was obtained as follows. *ubiU*
431 was amplified from pBAD-*yhbU* using pETDUET-UbiU-FW (introducing the NcoI site) and pETDUET-UbiU-
432 RV (introducing the EcoRI site) as primers. The NcoI and EcoRI digested amplicon was ligated to NcoI and
433 EcoRI digested pETDUET-1- plasmid to obtain pETDUET-UbiU. The *ubiV* gene was then cloned from pET-
434 22-UbiV into the MSC2 of pETDUET-UbiU by PCR amplification with pET-22-UbiV-FW and pETDUET-UbiV-
435 RV (introducing the C-terminal His_6-tag) as primers. The NdeI and XhoI digested amplicon was ligated to
436 NdeI and XhoI digested pETDUET-UbiU to obtain pETDUET-UbiUV.

437 A hexahistidine-tag was fused at the N-terminal extremity of UbiV to create pBAD-$UbiV_{6His}$. The $ubiV_{6His}$
438 gene was obtained by PCR amplification (Phusion High-Fidelity DNA Polymerase) using pBAD-UbiV as a
439 template and 6HisV5 (introducing the NcoI site) and 6HisV3 (introducing the HindIII site and the DNA
440 sequence of 6His-tag) as primers. The NcoI/HindIII digested amplicon was cloned into the NcoI/HindII-
441 digested pBAD plasmid.

442 Variants of UbiV and UbiU were constructed using the Q5 Site-Directed Mutagenesis Kit (New England
443 Biolabs) according to the manufacturer's specifications. The plasmids (pET-22b-UbiV, pETDUET-UbiUV,
444 pBAD-*yhbU* or pBAD-*yhbV)* were used as templates in conjunction with the appropriate primers for each
445 respective amino acid substitution.

446 ***Lipid extraction and quinone analysis*:**

447 Cultures (5-10 mL) were cooled down on ice before centrifugation at 3200 g, 4°C, 10 min. Cell pellets were
448 washed in 1 mL ice-cold PBS and transferred to pre-weighted 1.5 mL Eppendorf tubes. After centrifugation

449 at 12000 g, 4°C, 1 min and elimination of supernatant, the cells wet weight was determined (~10-20 mg)
450 and pellets were stored at -20°C. Quinone extraction from cell pellets was performed as previously
451 described [17]. Lipid extracts corresponding to 1 mg of cells wet weight were analyzed by HPLC-
452 electrochemical detection-mass spectrometry (ECD-MS) with a BetaBasic-18 column at a flow rate of 1
453 mL/min with mobile phases composed of methanol, ethanol, acetonitrile and a mix of 90% isopropanol,
454 10% ammonium acetate (1 M), 0.1% TFA: mobile phase 1 (50% methanol, 40% ethanol and 10% mix),
455 mobile phase 2 (40% acetonitrile, 40% ethanol, 20% mix). Mobile phase 1 was used in MS detection on a
456 MSQ spectrometer (Thermo Scientific) with electrospray ionization in positive mode (probe temperature
457 400°C, cone voltage 80V). Single ion monitoring (SIM) detected the following compounds: OPP ($M+NH_4^+$),
458 m/z 656.0-656.8, 5-10 min, scan time 0.2 s; DMQ_8 ($M+Na^+$), m/z 719-720, 6-10 min, scan time 0.2 s; $^{13}C_6$-
459 DMQ_8 ($M+Na^+$), m/z 725-726, 6-10 min, scan time 0.2 s; UQ_8 ($M+Na^+$), m/z 749-750, 6-10 min, scan time
460 0.2 s; $^{13}C_6$-UQ_8 ($M+Na^+$), m/z 755.0-756, 6-10 min, scan time 0.2 s; UQ_{10} ($M+NH_4^+$), m/z 880.2-881.2, 10-17
461 min, scan time 0.2 s. MS spectra were recorded between m/z 600 and 900 with a scan time of 0.3 s. UV
462 detection at 247 nm was used to quantify DMK_8 and MK_8. ECD, MS and UV peak areas were corrected for
463 sample loss during extraction on the basis of the recovery of the UQ_{10} internal standard and were then
464 normalized to cell's wet weight. The peak of UQ_8 obtained with electrochemical detection was quantified
465 with a standard curve of UQ_{10} [17].

466 ***Anaerobic $^{13}C_6$-UQ_8 biosynthesis activity assay*:**

467 Δ*ubiC* Δ*ubiF* cells containing or not the additional Δ*ubiU or* Δ*ubiV* deletions were grown overnight in MS
468 medium supplemented with 0.2% glucose. This preculture was used to inoculated at OD_{600}=0.1, 100 mL of
469 fresh medium supplemented with 10 µM $^{13}C_7$-4HB. The culture was grown at 37°C, 180 rpm until OD_{600}=1,
470 at which point 100 µM 4HB was added. The cells were pelleted by centrifugation at 3200 g, 4°C, 10 min
471 and suspended in 100 mL MSGN medium. A 10 mL aliquot was taken for quinone extraction (aerobiosis,
472 Figure 4B-D) and the rest of the culture was placed at 37°C in an anaerobic bottle with a two-port cap
473 fitted with plastic tubing used to inject argon to create and maintain anaerobiosis. After 5 min bubbling, a
474 10 mL sample was taken corresponding to 0 min anaerobiosis, then samples were taken every 30 minutes
475 and analyzed for quinone content.

476 ***Overexpression and purification of proteins:***

477 *Overexpression and purification E. coli wild-type UbiV and variants.*

478 The pET-22b(+) plasmid, encoding wild-type UbiV or variants, were co- transformed with pGro7 plasmid
479 (Takara Bio Inc.) into *E. coli* BL21 (DE3) competent cells. Single colonies obtained from transformation were
480 grown overnight at 37°C in LB medium supplemented with ampicillin (50 µg/mL) and chloramphenicol
481 (12.5 µg/mL). 10 mL of preculture was used to inoculate 1 L of LB medium with the same antibiotics, and
482 the bacteria were cultured further at 37 °C with shaking (200 rpm). At an OD_{600} of 1.2, D-arabinose was
483 added to the cultures at a final concentration of 2 mg/mL. At an OD_{600} of 1.8, the culture was cooled in an
484 ice-water bath, and Isopropyl 1-thio-β-D-galactopyranoside (IPTG) was added at a final concentration of
485 0.1 mM. Cells were then allowed to grow further at 16 °C overnight. All subsequent operations were
486 carried out at 4°C. Cells were harvested in an Avanti® J-26XP High-Performance centrifuge from Beckman
487 Coulter with a JLA-8.1000 rotor at 5,000 × *g* for 10 minutes. The cell pellets were resuspended in 5 volumes
488 of buffer A (50 mM Tris-HCl, pH 8.5, 150 mM NaCl, 15% (v/v) glycerol, 1 mM DTT) containing Complete™
489 Protease Inhibitor Cocktail (one tablet per 50 mL) (Roche) and disrupted by sonication (Branson Digital
490 Sonifier, amplitude 40% for 10 min). Cells debris were removed by ultracentrifugation in an Optima™ XPN-
491 80 ultracentrifuge from Beckman Coulter with a 50.2 Ti rotor at 35,000 x *g* for 60 min. The resulting
492 supernatant was loaded onto a HisTrap FF Crude column (GE Healthcare) pre-equilibrated with buffer A.
493 The column was washed with 10 column volumes of buffer B (50 mM Tris-HCl pH 8.5, 150 mM NaCl, 15%
494 (v/v) glycerol, 1 mM DTT, 10 mM imidazole) to remove non-specifically bound *E. coli* proteins then eluted
495 with a linear gradient of 10 column volumes of buffer B containing 500 mM imidazole . Fractions containing
496 WT UbiV or variants were pooled and phenylmethylsulfonyl fluoride was added at a final concentration of
497 1mM. The proteins were then loaded on a HiLoad 16/600 Superdex 75 pg (GE Healthcare) pre-equilibrated
498 in buffer C (50 mM Tris-HCl pH 8.5, 25 mM NaCl, 15% (v/v) glycerol, 1 mM DTT). The purified proteins were
499 concentrated to 30-40 mg/mL using Amicon concentrators (30-kDa cutoff; Millipore), aliquoted, frozen in
500 liquid nitrogen and stored at -80 °C. Overall, a high yield of 150 mg UbiV/L culture was obtained.

501 *Overexpression and purification of UbiU/V complex and variants.*

502 The overexpression of wild-type UbiU/V complex or variants in *E. coli* BL21 (DE3) competent cells were
503 performed following the same protocol as for UbiV. The expression of these proteins was induced by
504 addition of IPTG to a final concentration of 0.05 mM. Wild-type UbiU/V complex or variants were purified
505 with the same procedure as for UbiV, with the exception that the proteins were loaded on the HiLoad
506 16/600 Superdex 75 pg with buffer A.

507 **[Fe-S] cluster reconstitution**

508 The [Fe-S] cluster(s) reconstitution of holo-UbiV and holo-UbiU/V were conducted under strict anaerobic
509 conditions in an Mbraun LabStar glove box containing less than 0.5 ppm O_2. Classically, a solution
510 containing 100 µM of as-purified UbiV or UbiU/V complex, was treated with 5 mM DTT for 15 min at 20°C
511 and then incubated for 1 hour with a 5-fold molar excess of both ferrous ammonium sulfate and L-cysteine.
512 The reaction was initiated by the addition of a catalytic amount of the *E.coli* cysteine desulfurase CsdA (1-
513 2% molar equivalent) and monitored UV-visible absorption spectroscopy. The holo-UbiV or holo-UbiU/V
514 complexes were then loaded onto a Superdex 75 Increase 10/300 GL column (GE Healthcare) pre-
515 equilibrated with buffer C or A, respectively, to remove all excess of iron and L-cysteine. The fractions
516 containing the holo-proteins were pooled and concentrated to 20-30mg/mL on a Vivaspin concentrator
517 (30-kDa cutoff).

518 **Quantification Methods**

519 Protein concentrations were determined using the method of Bradford (Bio-Rad) with bovine serum
520 albumin as the standard. The iron and acid-labile sulfide were determined according to the method of Fish
521 [66] and Beinert [67], respectively.

522 **UV-Vis spectroscopy**

523 UV-Visible spectra were recorded in 1 cm optic path quartz cuvette under aerobic conditions on a Cary
524 100 UV-Vis spectrophotometer (Agilent) and under anaerobic conditions in a glove box on a XL-100 Uvikon
525 spectrophotometer equipped with optical fibers.

526 **EPR spectroscopy**

527 EPR spectra of frozen solutions were recorded on a Bruker Continuous-Waves (CW) X-Band ELEXSYS E500
528 spectrometer operating at 9.39 GHZ, equipped with an SHQE cavity cooled by an helium flow cryostat ESR
529 900 Oxford Instruments under non-saturating conditions and using the following parameters: a microwave
530 power in the range 2 to 10 mW and a modulation of the magnetic field at 100 kHz with a modulation
531 amplitude of 0.6 mT. Holo UbiV or holo UbiU-UbiV complex were treated with 10-fold molar excess of
532 dithionite to reduce the Fe-S cluster. Each solution was introduced into EPR quartz tubes in a glove box
533 and frozen with liquid nitrogen before the EPR measurements.

534 **Genome datasets**

535 The protein sequences from 5750 complete genomes ("extended dataset") were downloaded from the
536 NCBI Refseq database (bacteria and archaea, last accessed in November 2016, Table S2A). A representative
537 set of genomes from a monophyletic group of bacteria that potentially harbor the ubiquinone biosynthesis
538 pathway was also created: "Reference" and "Representative" genomes were downloaded from the NCBI
539 Refseq database for 204 Alphaproteobacteria, 103 Betaproteobacteria, 303 Gammaproteobacteria (last
540 accessed in November 2018). In addition to these 610 genomes, the genome of *Phaeospirillum fulvum* was
541 included (Table S2B).

542 **HMM protein profiles creation**

543 An initial set of protein sequences ("curated set") was retrieved from genomes manually and from a
544 publication [68], to cover the diversity of UQ-producing organisms. The curated set included 48 pairs of
545 YhbU and YhbV from 10 alpha, 19 beta and 19 gamma- proteobacteria, 17 sequences for YhbT, 64, 181,
546 69 and 189 sequences for UbiA, MenA, UbiG and UbiE respectively (Table S4). Then, for each gene family
547 these sequences were aligned with Mafft (v7.313, "linsi") [69], and each alignment trimmed at its N-ter
548 and C-ter extremities based on the filtering results of BMGE (BLOSUM 30) [70]. The core of the alignments
549 were kept as is, and Hidden Markov Model (HMM) profiles were created directly from the trimmed
550 alignments using the Hmmbuild program (Hmmer suite, version 3.1b2) [71].

551 To ensure a more sensitive search, and good delineation between homologs, phylogenetic curation was
552 used, as YhbU and YhbV are known to be part of the larger U32 protease gene family [42]. A search using
553 the YhbU and YhbV HMM profiles was performed with Hmmsearch (Hmmer suite) on the extended 5750
554 genomes dataset, and sequences with an i-evalue ("independent e-value") lower than 10E-20 and a
555 coverage of the profiles higher than 90% were selected. The 4212 sequences obtained were de-replicated
556 using Uclust from the Usearch program suite (80% identity level) [72]. A phylogenetic tree was built by
557 maximum-likelihood with the IQ-Tree program (best evolutionary model) based on the alignment (Mafft:
558 "linsi", BMGE with BLOSUM30) of the 480 selected sequences including all curated YhbU and YhbV
559 sequences [69,70,73]. YhbU and YhbV proteobacterial sequences formed two separate monophyletic
560 groups, giving credit to our curated set (100% and 80% UF-Boot support respectively). The other sequences
561 that formed a large monophyletic group of bacterial sequences were categorized as "U32 proteases" (98%
562 UF-Boot, https://doi.org/10.6084/m9.figshare.7800614.v1). The 98 sequences from this U32 proteases
563 group (Table S4) were used to re-create an HMM profile as described above, and served as an outgroup
564 for the profile search.

565 For YhbT, the first profile obtained from the curated set of sequences was used altogether with the YhbU,
566 YhbV and U32 proteases profiles for a search in the 611 proteobacteria genomes dataset. A second profile
567 was created from YhbT sequences (YhbT2, Table S4) that were co-localizing with YhbU and YhbV hits (10E-
568 20 i-evalue and 80% profile coverage). The two YhbT profiles matched complementary sets of sequences
569 and therefore were both used for annotating YhbT in genomes.

570 A similar approach was taken in order to identify the six known aerobic hydroxylases. 51 Coq7, 73 UbiF,
571 80 UbiH, 58 UbiI, 24 UbiL and 32 UbiM sequences were extracted manually and from publications [68]
572 [19](Table S4, "version 1") to serve as a reference, annotated set of sequences. Profiles were created as
573 described above. To ensure their specificity, we ran the HMM profiles against our 5570 genomes dataset
574 and selected the sequences that had an i-evalue lower than 10E-20 and a coverage of the profiles higher
575 than 90%. We built two phylogenetic trees as described above: one for Coq7, and another one for
576 UbiFHILM, which are known to be part of the large FMO protein family [19]. In the latter case, we de-
577 replicated the 1619 sequences obtained for the FMO protein family before performing the alignment,
578 alignment filtering, and tree reconstruction steps (using Uclust at the 60% identity level). The Coq7 tree
579 obtained showed our reference Coq7 sequences covered the whole diversity of retrieved sequences,
580 suggesting that they all could be *bona fide* Coq7 (https://doi.org/10.6084/m9.figshare.7800680). The FMO
581 tree showed a monophyletic group containing all reference FMO ubiquinone hydroxylases, forming sub-
582 groups for the different homologs (UbiFHILM) in Proteobacteria
583 (https://doi.org/10.6084/m9.figshare.7800620). Further, a large set of sequences formed an outgroup
584 consisting of sequences from various clades of bacteria, a lot being found outside of Proteobacteria,
585 robustly separated from the ubiquinone hydroxylases. We split this large clade into four sub-trees, and
586 extracted the corresponding sequences to obtain four new HMM profiles (as described above), to be used
587 for precise discrimination between ubiquinone hydroxylases and other members of the FMO family
588 ("AlloFMO_1" to "AlloFMO_4" in Table S4, "version 2"). FMO ubiquinone hydroxylases sub-trees were also
589 used to re-design improved HMM profiles for UbiFHILM (36, 168, 198, 139, and 65 sequences respectively,
590 see Table S4 "version 2").

591 **Evaluation of genomic distributions with HMMER and MacSyFinder**

592 Two MacSyFinder models were created to i) search sequences of interest in genomes using Hmmer, and
593 ii) investigate their genetic architecture [74]. A first model was created to focus only on YhbTUV-related
594 genes. In this model, the YhbTUV components were defined as “mandatory”, and U32 protease as
595 “accessory”. A second, more comprehensive model “Ubi+YhbTUV” was designed to list the families

596 corresponding to the 9 profiles obtained (UbiA, MenA, UbiE, UbiG, 2 YhbT, YhbU, YhbV and U32 proteases).
597 For both models, the two YhbT profiles were set as "exchangeable". The parameter
598 "inter_gene_max_space" was set to 5 in the "YhbTUV" model, and 10 in the "Ubi+YhbTUV" model.
599 MacSyFinder was set to run HMMER with the options "--i-evalue-select 10E-20" and "--coverage-profile
600 0.8". Independently of their genetic context, sequences corresponding to selected HMMER hits were listed
601 for all profiles in all genomes analyzed in order to establish the genomic distribution for each of the protein
602 families of interest. When several profiles matched a sequence, only the best hit (best i-evalue) was
603 considered.

604 **U32 proteases sequence analysis**

605 We retrieved from the UniProt-KB database 3460 protein sequences of U32 proteases that were
606 categorized in 12 families by Kimura *et al.* [42], and created a FASTA for each of these families. 50
607 sequences from different families could not be retrieved as they had been deleted from the Uniprot-KB
608 database (46 "obsolete") or were not found based on the published accession numbers. As the "RlhA1"
609 and "RlhA2" families mostly corresponded to two domains from the same protein sequences that had
610 been split, we put whole sequences all together into a single fasta file for sequence analysis of the overall
611 "RlhA" family. For each of the 10 family, sequences were de-replicated at the 80% identity level with
612 Uclust, in order to limit any potential taxonomic sampling bias, and sequences were aligned (Mafft, linsi).
613 The alignments were visualized in Jalview [75] and used to create the logo sequences. Images of
614 alignments were created using the ESPript webserver (http://espript.ibcp.fr/ESPript/ESPript/) [76].

615 ## ACKNOWLEDGEMENTS

616 This work was supported by the Agence Nationale de la Recherche (ANR), ANR Blanc (An)aeroUbi ANR-15-
617 CE11-0001-02 to FP. We thank Amélie Amblard for technical assistance, Louis Givelet for preliminary
618 bioinformatic analyses, Barbara Schoepp-Cothenet for providing accession numbers to sequences of UbiA,
619 -G, -E and for critical reading of the manuscript, and the GEM team at TIMC for discussions and suggestions.
620 CDTV, ML and MF acknowledge support from the French National Research Agency (Labex program
621 DYNAMO, ANR-11-LABX-0011). We thank the National Bioresource Project, National Institute of Genetics
622 for providing ME strains from the medium and large deletions *E. coli* collection.

623

624 ## COMPETING INTERESTS

625 The authors declare that they have no competing interests.

626

627 **REFERENCES**

628

629 1. Schoepp-Cothenet B, van Lis R, Atteia A, Baymann F, Capowiez L, Ducluzeau AL, et al. On the universal
630 core of bioenergetics. Biochim Biophys Acta. 2013; 1827: 79-93.
631 2. Marteyn B, Scorza FB, Sansonetti PJ, Tang C. Breathing life into pathogens: the influence of oxygen on
632 bacterial virulence and host responses in the gastrointestinal tract. Cell Microbiol. 2011; 13: 171-
633 176.
634 3. Wallace N, Zani A, Abrams E, Sun Y (2016) The Impact of Oxygen on Bacterial Enteric Pathogens. In:
635 Sariaslani S, Gadd GM, editors. Advances in Applied Microbiology, Vol 95. San Diego: Elsevier
636 Academic Press Inc. pp. 179-204.
637 4. Bettenbrock K, Bai H, Ederer M, Green J, Hellingwerf KJ, Holcombe M, et al. (2014) Towards a Systems
638 Level Understanding of the Oxygen Response of Escherichia coli. In: Poole RK, editor. Advances in
639 Microbial Systems Biology. London: Academic Press Ltd-Elsevier Science Ltd. pp. 65-114.
640 5. Morris RL, Schmidt TM. Shallow breathing: bacterial life at low O-2. Nature Reviews Microbiology.
641 2013; 11: 205-212.
642 6. Marreiros BC, Calisto F, Castro PJ, Duarte AM, Sena FV, Silva AF, et al. Exploring membrane respiratory
643 chains. Biochimica Et Biophysica Acta-Bioenergetics. 2016; 1857: 1039-1067.
644 7. Lecomte SM, Achouak W, Abrouk D, Heulin T, Nesme X, Haichar FE. Diversifying Anaerobic Respiration
645 Strategies to Compete in the Rhizosphere. Frontiers in Environmental Science. 2018; 6: 16.
646 8. Nowicka B, Kruk J. Occurrence, biosynthesis and function of isoprenoid quinones. Biochim Biophys
647 Acta. 2010; 1797: 1587-1605.
648 9. Sharma P, Teixeira de Mattos MJ, Hellingwerf KJ, Bekker M. On the function of the various quinone
649 species in Escherichia coli. FEBS J. 2012; 279: 3364-3373.
650 10. Shestopalov AI, Bogachev AV, Murtazina RA, Viryasov MB, Skulachev VP. Aeration-dependent
651 changes in composition of the quinone pool in Escherichia coli - Evidence of post-transcriptional
652 regulation of the quinone biosynthesis. Febs Letters. 1997; 404: 272-274.
653 11. Nitzschke A, Bettenbrock K. All three quinone species play distinct roles in ensuring optimal growth
654 under aerobic and fermentative conditions in E. coli K12. PLoS One. 2018; 13: e0194699.
655 12. Aussel L, Pierrel F, Loiseau L, Lombard M, Fontecave M, Barras F. Biosynthesis and physiology of
656 coenzyme Q in bacteria. Biochim Biophys Acta. 2014; 1837: 1004-1011.
657 13. Reidenbach AG, Kemmerer ZA, Aydin D, Jochem A, McDevitt MT, Hutchins PD, et al. Conserved Lipid
658 and Small-Molecule Modulation of COQ8 Reveals Regulation of the Ancient Kinase-like UbiB
659 Family. Cell Chem Biol. 2018; 25: 154-165 e111.
660 14. Loiseau L, Fyfe C, Aussel L, Hajj Chehade M, Hernandez SB, Faivre B, et al. The UbiK protein is an
661 accessory factor necessary for bacterial ubiquinone (UQ) biosynthesis and forms a complex with
662 the UQ biogenesis factor UbiJ. J Biol Chem. 2017; 292: 11937-11950.
663 15. Aussel L, Loiseau L, Hajj Chehade M, Pocachard B, Fontecave M, Pierrel F, et al. ubiJ, a New Gene
664 Required for Aerobic Growth and Proliferation in Macrophage, Is Involved in Coenzyme Q
665 Biosynthesis in Escherichia coli and Salmonella enterica Serovar Typhimurium. J Bacteriol. 2014;
666 196: 70-79.
667 16. Hajj Chehade M, Pelosi L, Fyfe C, Loiseau L, Rascalou B, Brugiere S, et al. A soluble metabolon
668 synthesizes the isoprenoid lipid Ubiquinone. Cell Chemical Biology (in press). 2019.

669 17. Hajj Chehade M, Loiseau L, Lombard M, Pecqueur L, Ismail A, Smadja M, et al. ubiI, a New Gene in
670 Escherichia coli Coenzyme Q Biosynthesis, Is Involved in Aerobic C5-hydroxylation. J Biol Chem.
671 2013; 288: 20085-20092.
672 18. Alexander K, Young IG. Three hydroxylations incorporating molecular oxygen in the aerobic
673 biosynthesis of ubiquinone in Escherichia coli. Biochemistry. 1978; 17: 4745-4750.
674 19. Pelosi L, Ducluzeau AL, Loiseau L, Barras F, Schneider D, Junier I, et al. Evolution of Ubiquinone
675 Biosynthesis: Multiple Proteobacterial Enzymes with Various Regioselectivities To Catalyze Three
676 Contiguous Aromatic Hydroxylation Reactions. mSystems. 2016; 1: e00091-00016.
677 20. Alexander K, Young IG. Alternative hydroxylases for the aerobic and anaerobic biosynthesis of
678 ubiquinone in Escherichia coli. Biochemistry. 1978; 17: 4750-4755.
679 21. Cox GB, Gibson F. Biosynthesis of Vitamin K and Ubiquinone. Relation to the Shikimic Acid Pathway in
680 Escherichia Coli. Biochim Biophys Acta. 1964; 93: 204-206.
681 22. Siebert M, Severin K, Heide L. Formation of 4-hydroxybenzoate in Escherichia coli: characterization of
682 the ubiC gene and its encoded enzyme chorismate pyruvate-lyase. Microbiology. 1994; 140 (Pt
683 4): 897-904.
684 23. Payet LA, Leroux M, Willison JC, Kihara A, Pelosi L, Pierrel F. Mechanistic Details of Early Steps in
685 Coenzyme Q Biosynthesis Pathway in Yeast. Cell Chem Biol. 2016; 23: 1241-1250.
686 24. Kwon O, Kotsakis A, Meganathan R. Ubiquinone (coenzyme Q) biosynthesis in Escherichia coli:
687 identification of the ubiF gene. FEMS Microbiol Lett. 2000; 186: 157-161.
688 25. Hashimoto M, Ichimura T, Mizoguchi H, Tanaka K, Fujimitsu K, Keyamura K, et al. Cell size and
689 nucleoid organization of engineered Escherichia coli cells with a reduced genome. Mol Microbiol.
690 2005; 55: 137-149.
691 26. Mizoguchi H, Sawano Y, Kato J, Mori H. Superpositioning of deletions promotes growth of Escherichia
692 coli with a reduced genome. DNA Res. 2008; 15: 277-284.
693 27. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, et al. Construction of Escherichia coli K-12
694 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol. 2006; 2: 2006 0008.
695 28. Esposti MD. A journey across genomes uncovers the origin of ubiquinone in cyanobacteria. Genome
696 Biol Evol. 2017.
697 29. Hiraishi A, Hoshino Y, Kitamura H. Isoprenoid quinone composition in the classification of
698 rhodospirillaceae. Journal of General and Applied Microbiology. 1984; 30: 197-210.
699 30. Imhoff JF, Petri R, Suling J. Reclassification of species of the spiral-shaped phototrophic purple non-
700 sulfur bacteria of the alpha-Proteobacteria: description of the new genera Phaeospirillum gen.
701 nov., Rhodovibrio gen. nov., Rhodothalassium gen. nov. and Roseospira gen. nov. as well as
702 transfer of Rhodospirillum fulvum to Phaeospirillum fulvum comb. nov., of Rhodospirillum
703 molischianum to Phaeospirillum molischianum comb. nov., of Rhodospirillum salinarum to
704 Rhodovibrio salinarum comb. nov., of Rhodospirillum sodomense to Rhodovibrio sodomensis
705 comb. nov., of Rhodospirillum salexigens to Rhodothalassium salexigens comb. nov, and of
706 Rhodospirillum mediosalinum to Roseospira mediosalina comb. nov. International Journal of
707 Systematic Bacteriology. 1998; 48: 793-798.
708 31. Bazylinski DA, Williams TJ, Lefevre CT, Berg RJ, Zhang CL, Bowser SS, et al. Magnetococcus marinus
709 gen. nov., sp. nov., a marine, magnetotactic bacterium that represents a novel lineage
710 (Magnetococcaceae fam. nov., Magnetococcales ord. nov.) at the base of the
711 Alphaproteobacteria. Int J Syst Evol Microbiol. 2013; 63: 801-808.
712 32. Stewart CS, Duncan SH, Cave DR. Oxalobacter formigenes and its role in oxalate metabolism in the
713 human gut. FEMS Microbiol Lett. 2004; 230: 1-7.
714 33. Burgardt NI, Gianotti AR, Ferreyra RG, Ermacora MR. A structural appraisal of sterol carrier protein 2.
715 Biochimica Et Biophysica Acta-Proteins and Proteomics. 2017; 1865: 565-577.

716 34. Yano T, Yagi T, Sled VD, Ohnishi T. Expression and characterization of the 66-kilodalton (NQO3) iron-
717 sulfur subunit of the proton-translocating NADH-quinone oxidoreductase of Paracoccus
718 denitrificans. J Biol Chem. 1995; 270: 18264-18270.
719 35. Ollagnier S, Mulliez E, Gaillard J, Eliasson R, Fontecave M, Reichard P. The anaerobic Escherichia coli
720 ribonucleotide reductase. Subunit structure and iron sulfur center. J Biol Chem. 1996; 271: 9410-
721 9416.
722 36. M NM, Ollagnier-de-Choudens S, Sanakis Y, Abdel-Ghany SE, Rousset C, Ye H, et al. Characterization
723 of Arabidopsis thaliana SufE2 and SufE3: functions in chloroplast iron-sulfur cluster assembly and
724 Nad synthesis. J Biol Chem. 2007; 282: 18254-18264.
725 37. Ollagnier-De Choudens S, Sanakis Y, Hewitson KS, Roach P, Baldwin JE, Munck E, et al. Iron-sulfur
726 center of biotin synthase and lipoate synthase. Biochemistry. 2000; 39: 4165-4173.
727 38. Ollagnier-de Choudens S, Loiseau L, Sanakis Y, Barras F, Fontecave M. Quinolinate synthetase, an
728 iron-sulfur enzyme in NAD biosynthesis. FEBS Lett. 2005; 579: 3737-3743.
729 39. Hewitson KS, Ollagnier-de Choudens S, Sanakis Y, Shaw NM, Baldwin JE, Munck E, et al. The iron-
730 sulfur center of biotin synthase: site-directed mutants. J Biol Inorg Chem. 2002; 7: 83-93.
731 40. Johnson DC, Dean DR, Smith AD, Johnson MK. Structure, function, and formation of biological iron-
732 sulfur clusters. Annu Rev Biochem. 2005; 74: 247-281.
733 41. Fontecave M, Py B, Ollagnier de Choudens S, Barras F. From Iron and Cysteine to Iron-Sulfur Clusters:
734 the Biogenesis Protein Machineries. EcoSal Plus. 2008; 3.
735 42. Kimura S, Sakai Y, Ishiguro K, Suzuki T. Biogenesis and iron-dependency of ribosomal RNA
736 hydroxylation. Nucleic Acids Res. 2017; 45: 12974-12986.
737 43. Yen HW, Chiu CH. The influences of aerobic-dark and anaerobic-light cultivation on CoQ(10)
738 production by Rhodobacter sphaeroides in the submerged fermenter. Enzyme and Microbial
739 Technology. 2007; 41: 600-604.
740 44. Matsumura M, Kobayashi T, Aiba S. ANAEROBIC PRODUCTION OF UBIQUINONE-10 BY PARACOCCUS-
741 DENITRIFICANS. European Journal of Applied Microbiology and Biotechnology. 1983; 17: 85-89.
742 45. Schoepp-Cothenet B, Lieutaud C, Baymann F, Vermeglio A, Friedrich T, Kramer DM, et al.
743 Menaquinone as pool quinone in a purple bacterium. Proc Natl Acad Sci U S A. 2009; 106: 8549-
744 8554.
745 46. Stolper DA, Revsbech NP, Canfield DE. Aerobic growth at nanomolar oxygen concentrations.
746 Proceedings of the National Academy of Sciences of the United States of America. 2010; 107:
747 18755-18760.
748 47. D'Mello R, Hill S, Poole RK. The cytochrome bd quinol oxidase in Escherichia coli has an extremely
749 high oxygen affinity and two oxygen-binding haems: implications for regulation of activity in vivo
750 by oxygen inhibition. Microbiology. 1996; 142 (Pt 4): 755-763.
751 48. Rolfe MD, Ocone A, Stapleton MR, Hall S, Trotter EW, Poole RK, et al. Systems analysis of
752 transcription factor activities in environments with stable and dynamic oxygen concentrations.
753 Open Biol. 2012; 2: 120091.
754 49. Rolfe MD, Ter Beek A, Graham AI, Trotter EW, Asif HM, Sanguinetti G, et al. Transcript profiling and
755 inference of Escherichia coli K-12 ArcA activity across the range of physiologically relevant
756 oxygen concentrations. J Biol Chem. 2011; 286: 10147-10154.
757 50. Lee KM, Park Y, Bari W, Yoon MY, Go J, Kim SC, et al. Activation of Cholera Toxin Production by
758 Anaerobic Respiration of Trimethylamine N-oxide in Vibrio cholerae. Journal of Biological
759 Chemistry. 2012; 287: 11.
760 51. Zhao H, Li X, Johnson DE, Mobley HLT. Identification of protease and rpoN-associated genes of
761 uropathogenic Proteus mirabilis by negative selection in a mouse model of ascending urinary
762 tract infection. Microbiology-Uk. 1999; 145: 185-195.

763 52. Navais R, Mendez J, Perez-Pascual D, Cascales D, Guijarro JA. The yrpAB operon of Yersinia ruckeri
764 encoding two putative U32 peptidases is involved in virulence and induced under microaerobic
765 conditions. Virulence. 2014; 5: 619-624.
766 53. Groenewold MK, Massmig M, Hebecker S, Danne L, Magnowska Z, Nimtz M, et al. A phosphatidic
767 acid binding protein is important for lipid homeostasis and adaptation to anaerobic biofilm
768 conditions in Pseudomonas aeruginosa. Biochem J. 2018.
769 54. Filiatrault MJ, Picardo KF, Ngai H, Passador L, Iglewski BH. Identification of Pseudomonas aeruginosa
770 genes involved in virulence and anaerobic growth. Infect Immun. 2006; 74: 4237-4245.
771 55. Schobert M, Jahn D. Anaerobic physiology of Pseudomonas aeruginosa in the cystic fibrosis lung. Int J
772 Med Microbiol. 2010; 300: 549-556.
773 56. Line L, Alhede M, Kolpen M, Kuhl M, Ciofu O, Bjarnsholt T, et al. Physiological levels of nitrate support
774 anoxic growth by denitrification of Pseudomonas aeruginosa at growth rates reported in cystic
775 fibrosis lungs and sputum. Frontiers in Microbiology. 2014; 5: 11.
776 57. Estellon J, Ollagnier de Choudens S, Smadja M, Fontecave M, Vandenbrouck Y. An integrative
777 computational model for large-scale identification of metalloproteins in microbial genomes: a
778 focus on iron-sulfur cluster proteins. Metallomics. 2014; 6: 1913-1930.
779 58. Rawlings ND, Alan J, Thomas PD, Huang XD, Bateman A, Finn RD. The MEROPS database of
780 proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in
781 the PANTHER database. Nucleic Acids Research. 2018; 46: D624-D632.
782 59. Rouault TA. Mammalian iron-sulphur proteins: novel insights into biogenesis and function. Nature
783 Reviews Molecular Cell Biology. 2015; 16: 45-55.
784 60. Lill R. Function and biogenesis of iron-sulphur proteins. Nature. 2009; 460: 831-838.
785 61. Hille R. Molybdenum-containing hydroxylases. Arch Biochem Biophys. 2005; 433: 107-116.
786 62. Alberge F, Espinosa L, Seduk F, Sylvi L, Toci R, Walburger A, et al. Dynamic subcellular localization of a
787 respiratory complex controls bacterial respiration. Elife. 2015; 4.
788 63. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in Escherichia coli K-12 using
789 PCR products. Proc Natl Acad Sci U S A. 2000; 97: 6640-6645.
790 64. Miller JH (1972) Experiments in Molecular Genetics; Laboratory CSH, editor.
791 65. Cherepanov PP, Wackernagel W. Gene disruption in Escherichia coli: TcR and KmR cassettes with the
792 option of Flp-catalyzed excision of the antibiotic-resistance determinant. Gene. 1995; 158: 9-14.
793 66. Fish WW. Rapid colorimetric micromethod for the quantitation of complexed iron in biological
794 samples. Methods in Enzymology. 1988; 158: 357-364.
795 67. Beinert H. Micro methods for the quantitative determination of iron and copper in biological
796 material. Methods in enzymology. 1978; 54: 435-445.
797 68. Ravcheev DA, Thiele I. Genomic Analysis of the Human Gut Microbiome Suggests Novel Enzymes
798 Involved in Quinone Biosynthesis. Front Microbiol. 2016; 7: 128.
799 69. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in
800 performance and usability. Mol Biol Evol. 2013; 30: 772-780.
801 70. Criscuolo A, Gribaldo S. BMGE (Block Mapping and Gathering with Entropy): a new software for
802 selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol Biol.
803 2010; 10: 210.
804 71. Eddy SR. Accelerated Profile HMM Searches. PLoS Comput Biol. 2011; 7: e1002195.
805 72. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010; 26:
806 2460-2461.
807 73. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm
808 for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015; 32: 268-274.
809 74. Abby SS, Neron B, Menager H, Touchon M, Rocha EP. MacSyFinder: a program to mine genomes for
810 molecular systems with an application to CRISPR-Cas systems. PLoS One. 2014; 9: e110726.

811 75. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2--a multiple sequence
812 alignment editor and analysis workbench. Bioinformatics. 2009; 25: 1189-1191.
813 76. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server.
814 Nucleic Acids Res. 2014; 42: W320-324.
815 77. Garcia PS, Jauffrit F, Grangeasse C, Brochier-Armanet C. GeneSpy, a user-friendly and flexible
816 genomic context visualizer. Bioinformatics (Oxford, England). 2019; 35: 329-331.

817

818

MAIN FIGURES

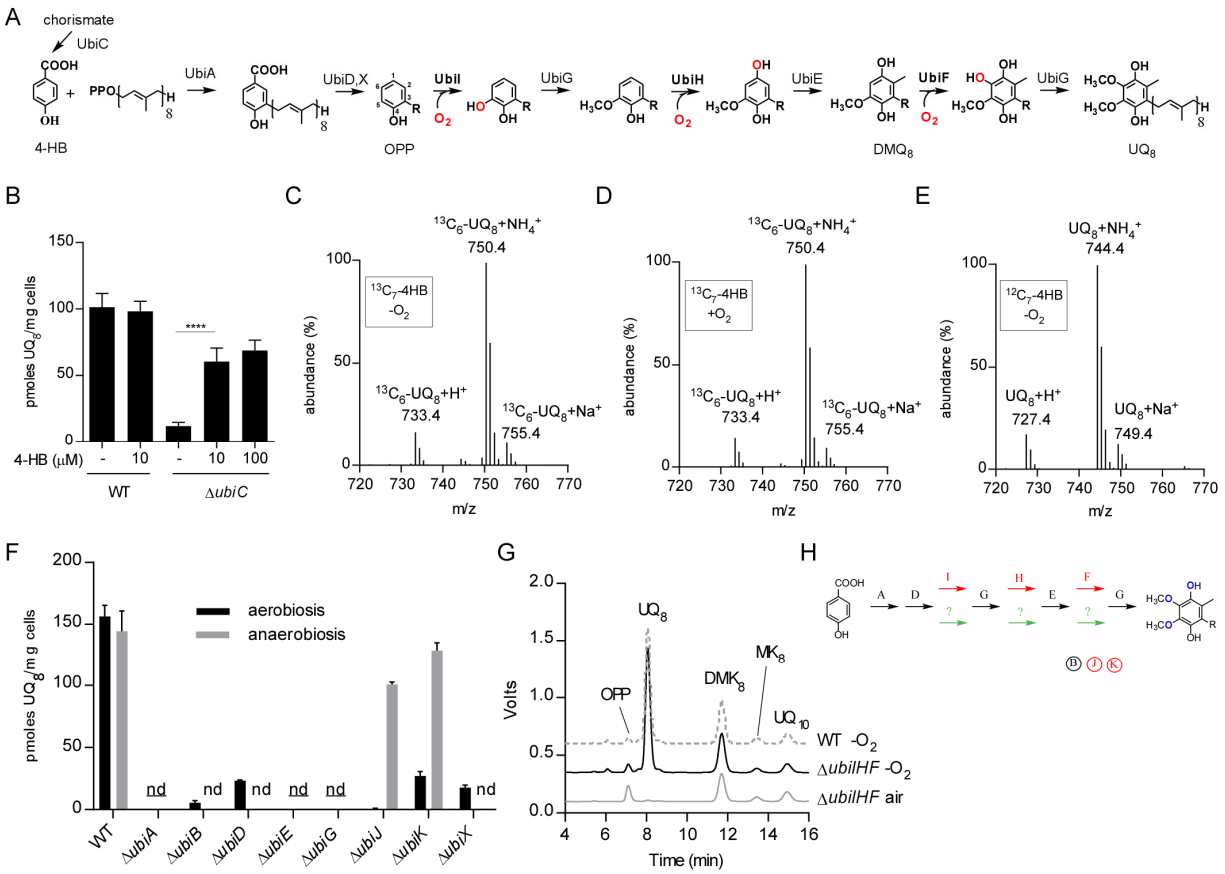


Figure 1: **The aerobic and anaerobic UQ biosynthetic pathways differ only in the hydroxylation steps.** A) O_2-dependent UQ biosynthesis pathway in *E. coli*. The octaprenyl tail is represented by R on the biosynthetic intermediates and the numbering of the aromatic carbon atoms is shown on OPP. Abbreviations used are 4-HB for 4-hydroxybenzoic acid, OPP for octaprenylphenol, DMQ_8 for C6-demethoxy-ubiquinone 8 and UQ_8 for ubiquinone 8. B) UQ_8 quantification of WT and Δ*ubiC* cells grown anaerobically in glycerol-nitrate medium supplemented or not with the indicated concentrations of 4-HB (n=3-6), ****: $p< 0.0001$, unpaired Student's t test. C-E) Mass spectra of UQ_8 obtained by HPLC-MS analysis of lipid extracts from cells grown with $^{13}C_7$-4HB either anaerobically (C) or aerobically (D), or anaerobically with unlabeled 4HB (E). F) UQ_8 quantification from WT and Δ*ubi* cells grown anaerobically in SMGN medium overnight or aerobically in LB medium until OD 0.8. nd = not detected in aerobic and anaerobic conditions, nd= not detected in either aerobic or anaerobic conditions, n=3-4. G) HPLC-ECD analyses (mobile phase 1) of lipid extracts from 1 mg of WT or Δ*ubiHF* cells grown in LB medium in air or anaerobic conditions ($-O_2$). Chromatograms are representative of n=3 independent experiments (UQ_{10} used as standard). H) UQ biosynthesis represented with Ubi enzymes specific to the O_2-dependent pathway (red), to the O_2-independent pathway (green), or common to both pathways (black). The same color code applies to the accessory factors (circled).

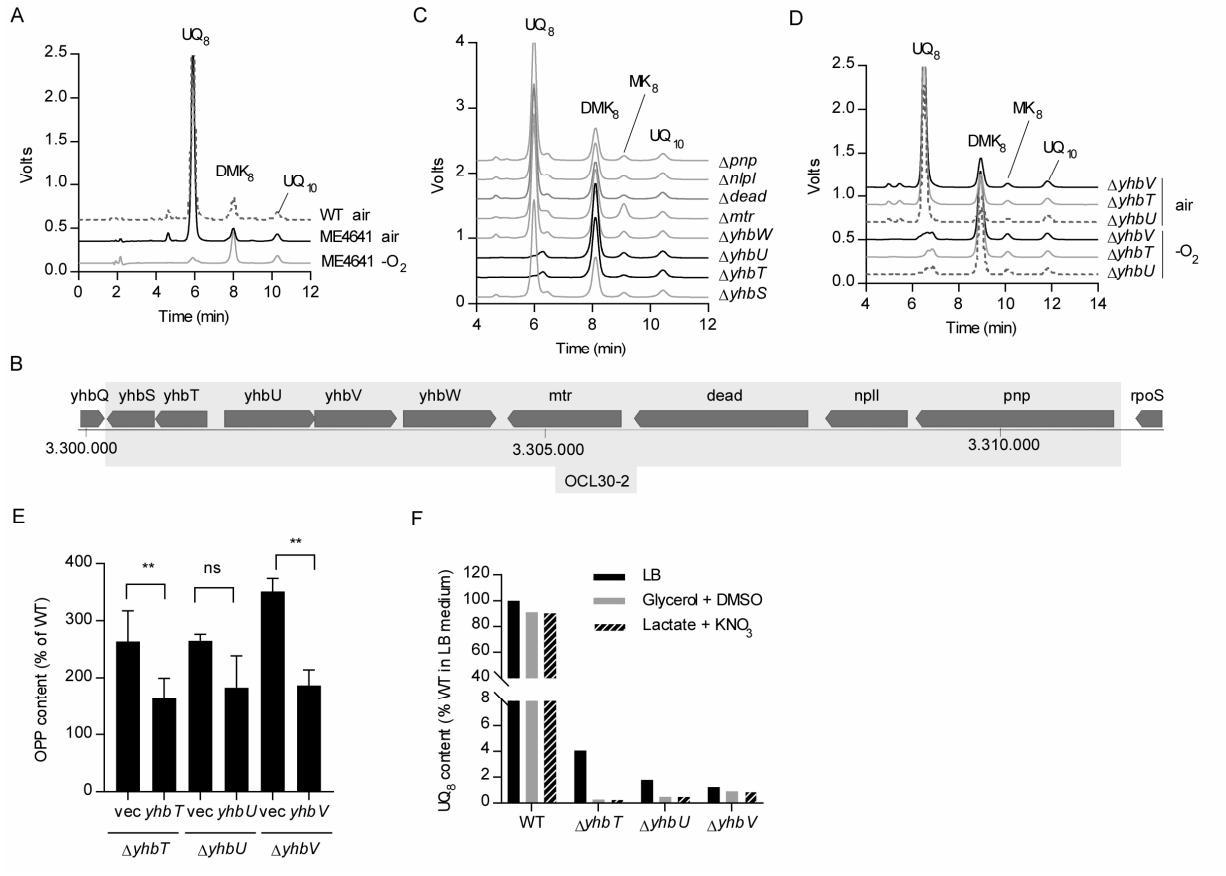


Figure 2: ***yhbT, yhbU* and *yhbV* are essential to the anaerobic biosynthesis of UQ.** A) HPLC-ECD analysis of lipid extracts from ME4641 strain grown in SMGN either aerobically or anaerobically ($-O_2$). B) Genomic region covered by the OCL30-2 deletion in the ME4641 strain. C) HPLC-ECD analysis of lipid extracts from knock-out strains of the individual genes covered by the OCL30-2 deletion grown in SMGN anaerobically. D) HPLC-ECD analysis of lipid extracts from Δ*yhbT*, Δ*yhbU* and Δ*yhbV* strains constructed in the MG1655 background and grown in glycerol nitrate medium either aerobically or anaerobically. HPLC-ECD analyses with mobile phase 2 (A, C, D). E) OPP content (as % of WT, mass detection $M+NH_4^+$) in cells from table 2. The Δ*yhb* strains contain either an empty plasmid or a plasmid carrying the indicated gene and were cultured anaerobically in glycerol nitrate medium containing 0.02% arabinose. mean ±SD (n=3-5), **: $p< 0.01$, unpaired Student's t test. F) UQ_8 content (as % of WT grown in LB medium) of cells cultured anaerobically in SM containing the indicated carbon sources and electron acceptors.

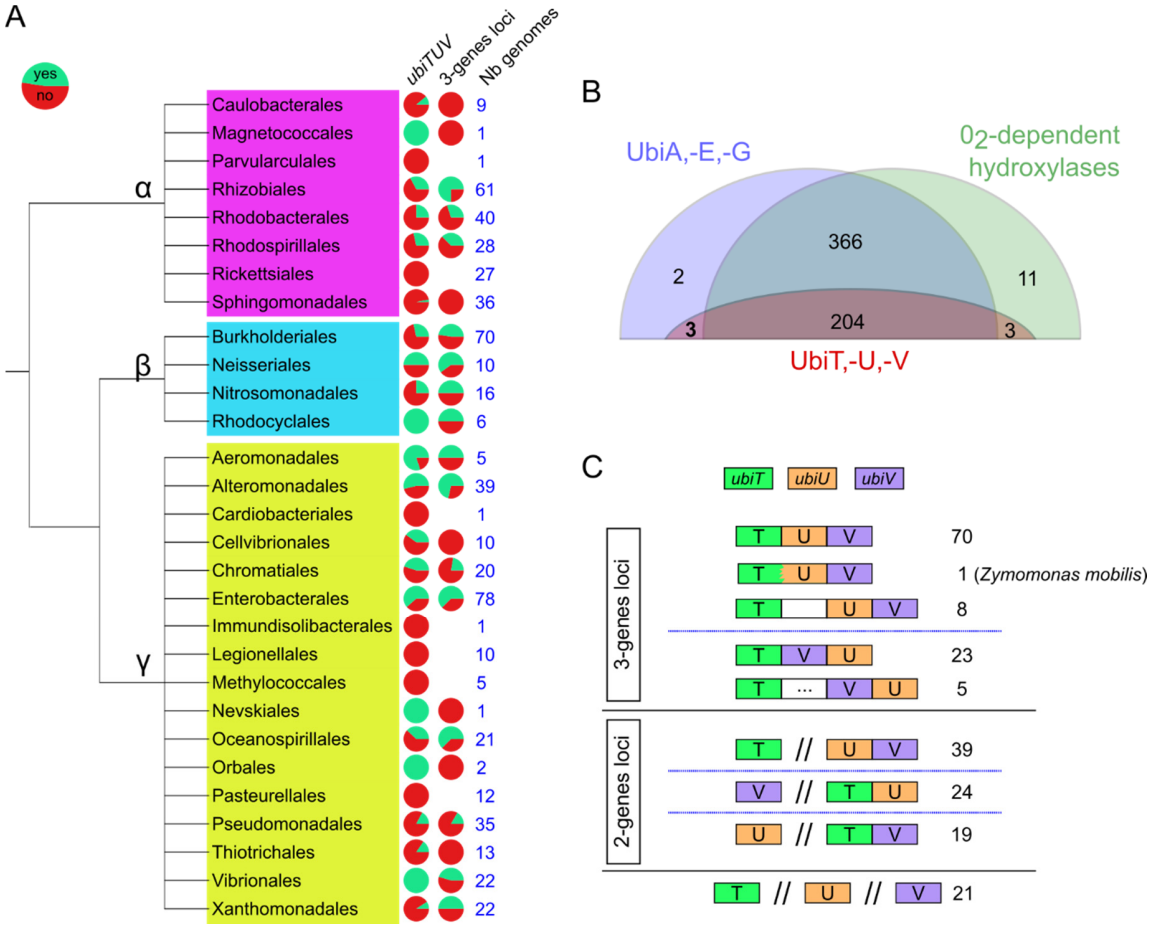
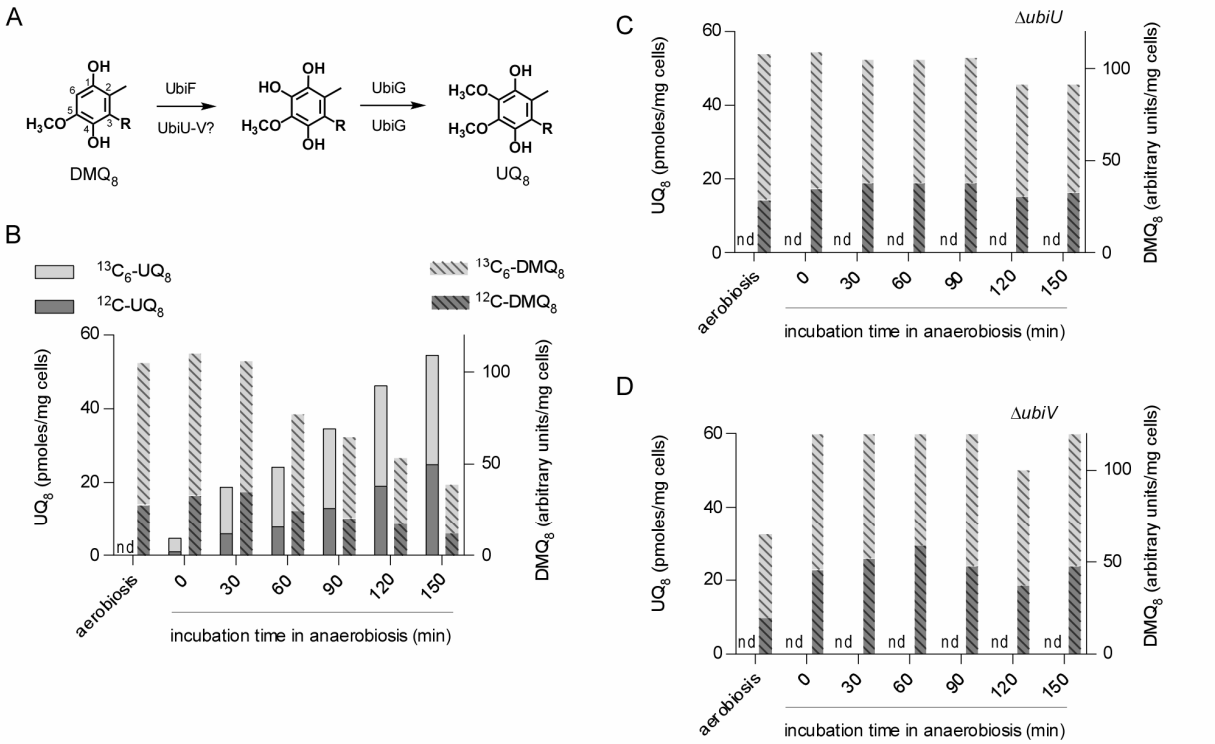


Figure 3: *ubiT*, *-U* and *-V* occurrence and genetic architecture in proteobacterial genomes.

A) The proportion of genomes with (green) and without (red) all three genes "*ubiTUV*" (left column) is indicated for each proteobacterial order known to synthesize UQ. The middle column "3-genes loci" displays the proportion of genomes with the three genes either at a single locus (green) or at different loci (red). The number of genomes analyzed for each order is given in the right column ("Nb Genomes"). B) Occurence in the reference proteobacterial genomes of the marker proteins (UbiA, -E, -G), of the O_2–dependent hydroxylases and of the UbiT,-U, -V proteins. The number in bold represents the 3 genomes (*P. fulvum*, *M. marinus* and *O. formigenes*) containing exclusively the O_2-independent pathway. C) The distinct genetic architectures found for *ubiT*, *-U* and *-V* in genomes where the three genes were present are displayed as boxes with different colors. Number of cases corresponding to each depicted architecture are given on the right. A white box corresponds to a gene found between the genes of interest, and a white box with dots corresponds to two to five genes between the genes of interest.



867

868 **Figure 4**: **UbiU and UbiV are necessary for the anaerobic conversion of DMQ_8 into UQ_8.** A) Conversion of
869 DMQ_8 to UQ_8 with enzymes of the O_2-dependent and the O_2-independent pathways, indicated respectively
870 above and below arrows (numbering of carbon atoms shown on DMQ_8 and polyprenyl tail represented by
871 R). B) Unlabeled (^{12}C) and labeled ($^{13}C_6$) -DMQ_8 and -UQ_8 contents in Δ*ubiC* Δ*ubiF* cells after aerobic growth
872 and transition to anaerobiosis. C-D) Same as in B with Δ*ubiC* Δ*ubiF* Δ*ubiU* cells (C) and Δ*ubiC* Δ*ubiF* Δ*ubiV*
873 cells (D). nd, not detected, results representative of two independent experiments (B-D).

874

875

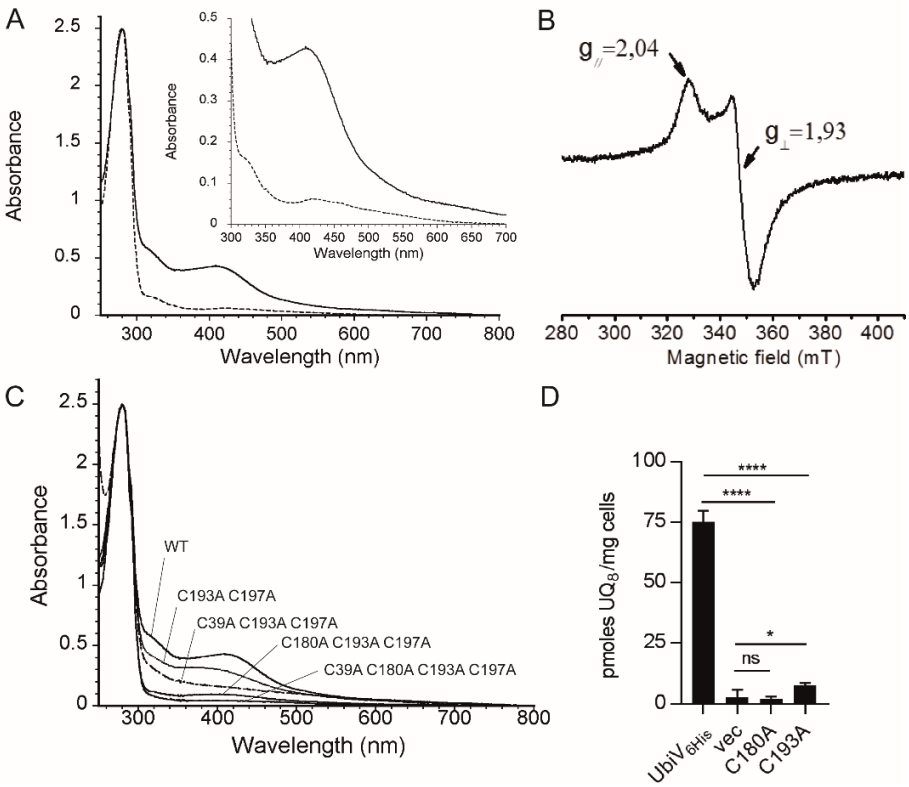
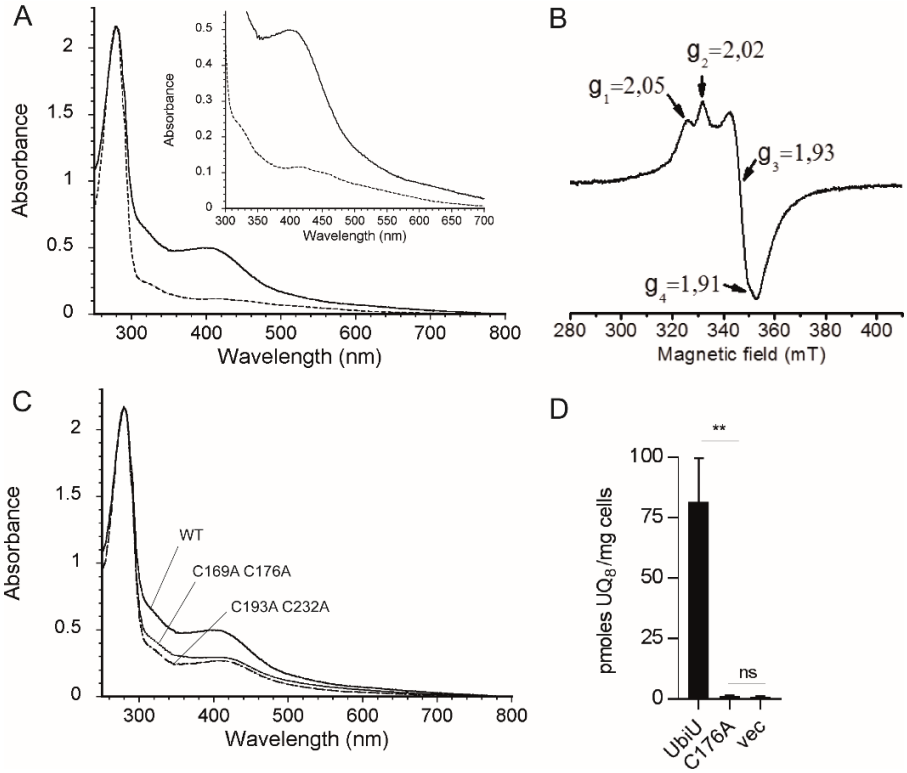


Figure 5: **UbiV binds a [4Fe-4S] cluster**. A) UV-visible absorption spectra of as-purified UbiV (dotted line, 47 µM) and reconstituted holo-UbiV (solid line, 41 µM); Inset: enlargement of the 300-700 nm region. The molar extinction coefficient ε_{410nm} was determined to be 10.8 ± 0.4 mM^{-1} cm^{-1} for holo-UbiV. B) X-band EPR spectrum of 785 µM dithionite-reduced holo-UbiV. Recording conditions: temperature, 10K; microwave power, 10 mW; modulation amplitude, 0.6 mT. C) Comparative UV-visible absorption spectra of WT and different Cys-to-Ala mutants of UbiV after metal cluster reconstitution, with the following concentrations: 41 µM WT, 44 µM C193A C197A, 46 µM C39A C193A C197A, 47 µM C180A C193A C197A, and 54 µM C39A C180A C193A C197A. Proteins were in 50 mM Tris-HCl, pH 8.5, 25 mM NaCl, 15% glycerol, 1m M DTT (A-C). D) UQ_8 quantification of Δ*ubiV* cells transformed with pBAD-$UbiV_{6His}$, pBAD-$UbiV_{6His}$ C180A, pBAD-$UbiV_{6His}$ C193A or empty pBAD and grown overnight in anaerobic MSGN (n=4-5). Mean ±SD, *: p< 0.05, ****: p< 0.0001, unpaired Student's t test.



890

891

892 **Figure 6: The UbiU-V complex binds two [4Fe-4S] clusters**. A) UV-visible absorption spectra of as-purified
893 UbiU-UbiV (dotted line, 17 µM) and reconstituted holo-UbiU-UbiV (solid line, 15.5 µM); Inset: enlargement
894 of the 300-700 nm region. B) X-band EPR spectrum of 339 µM dithionite-reduced holo-UbiU-UbiV;
895 Recording conditions: temperature, 10K; microwave power, 2 mW; modulation amplitude, 0.6 mT. C)
896 Comparative UV-visible absorption spectra of Cys-to-Ala mutants of UbiU in the UbiU-UbiV complex after
897 metal cluster reconstitution with the following concentrations: 15.5 µM WT, 16.0 µM UbiU-C169A C176A
898 and 16.0 µM UbiU-C193A C232A. Proteins were in 50mM Tris-HCl, pH 8.5, 150mM NaCl, 15% glycerol and
899 1mM DTT (A-C). D) UQ_8 quantification of Δ*ubiU* cells transformed with pBAD-UbiU (n=4), pBAD-UbiU
900 C176A (n=2) or pBAD empty vector (n=3) and grown overnight in anaerobic MSGN. mean ±SD, **: $p < 0.01$,
901 ns: not significant, unpaired Student's t test.

902

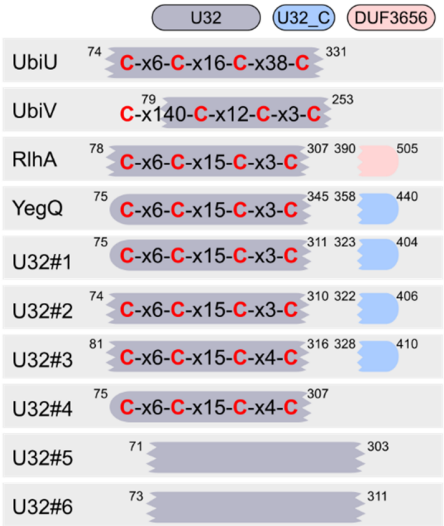


Figure 7: Conserved four cysteines motifs in the U32 protease family.

The conserved 4-cysteines motifs and PFAM domains (colored boxes) found in each U32 protease family are displayed for a set of reference sequences. These motifs were obtained by aligning the sequences listed by Kimura *et al.* [42]. Conserved cysteines are in red, "x6" indicates that 6 residues were found in between two conserved cysteines. Positions of the domains are displayed on the outside of the boxes for the reference sequences. Scrambled extremities show interrupted match for the PFAM domain. No conserved cysteines were found for U32#5 and U32#6 (see main text). Reference sequences were from *E. coli* for UbiU, UbiV, YegQ and RhlA (YHBU_ECOLI, YHBV_ECOLI, YEGQ_ECOLI, YDCP_ECOLI for RlhA). For the rest of the families, the sequence accession numbers were: R7JPV1_9FIRM for U32#1, R6XKQ3_9CLOT for U32#2, S1NZZ5_9ENTE for U32#3, H1YXA1_9EURY for U32#4, H3NJ45_9LACT for U32#5, D5MIQ1_9BACT for U32#6.

SUPPLEMENTARY FIGURES

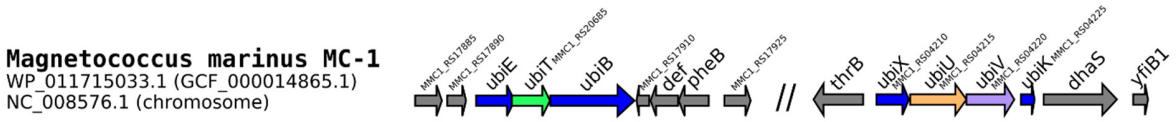


Figure S1: Genetic context of the *ubiT,U,V* genes in *M. marinus.* The scheme was drawn from a figure obtained using the GeneSpy program [77].

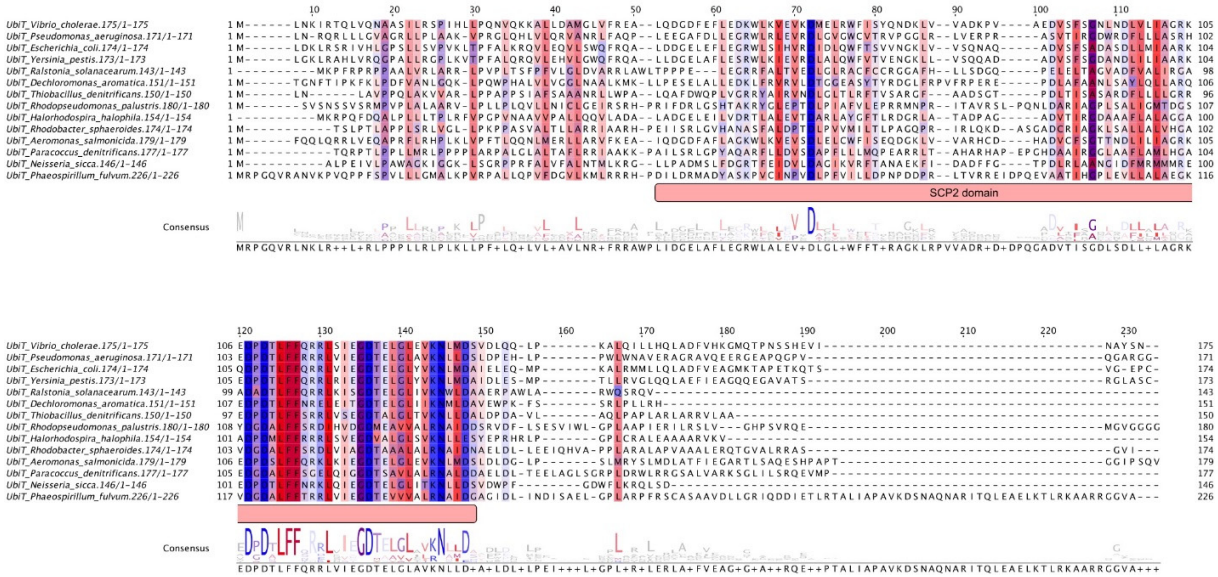


Figure S2: Multiple sequence alignment of UbiT from representative proteobacteria. Sequences were aligned using Mafft (linsi), and the output generated using Jalview and Inkscape. Hydrophobic residues are colored in blue. The position of the SCP2 domain (PF02036) is indicated by a red box. Genbank accession numbers: *Vibrio cholerae*, NP_230303.1, *Pseudomonas aeruginosa*, NP_252600.1, *Escherichia coli*, NP_312065.1, *Yersinia pestis*, YP_002348366.1, *Ralstonia solanacearum*, WP_011004260.1, *Dechloromonas aromatica*, WP_011285876.1, *Thiobacillus denitrificans*, WP_011312695.1, *Rhodopseudomonas palustris*, WP_011666115.1, *Halorhodospira halophila*, WP_011813831.1, *Rhodobacter sphaeroides*, WP_011909347.1, *Aeromonas salmonicida*, WP_005310396.1, *Paracoccus denitrificans*, WP_011750456.1, *Neisseria sicca*, WP_080614297.1, *Phaeospirillum fulvum*, WP_074767212.1.

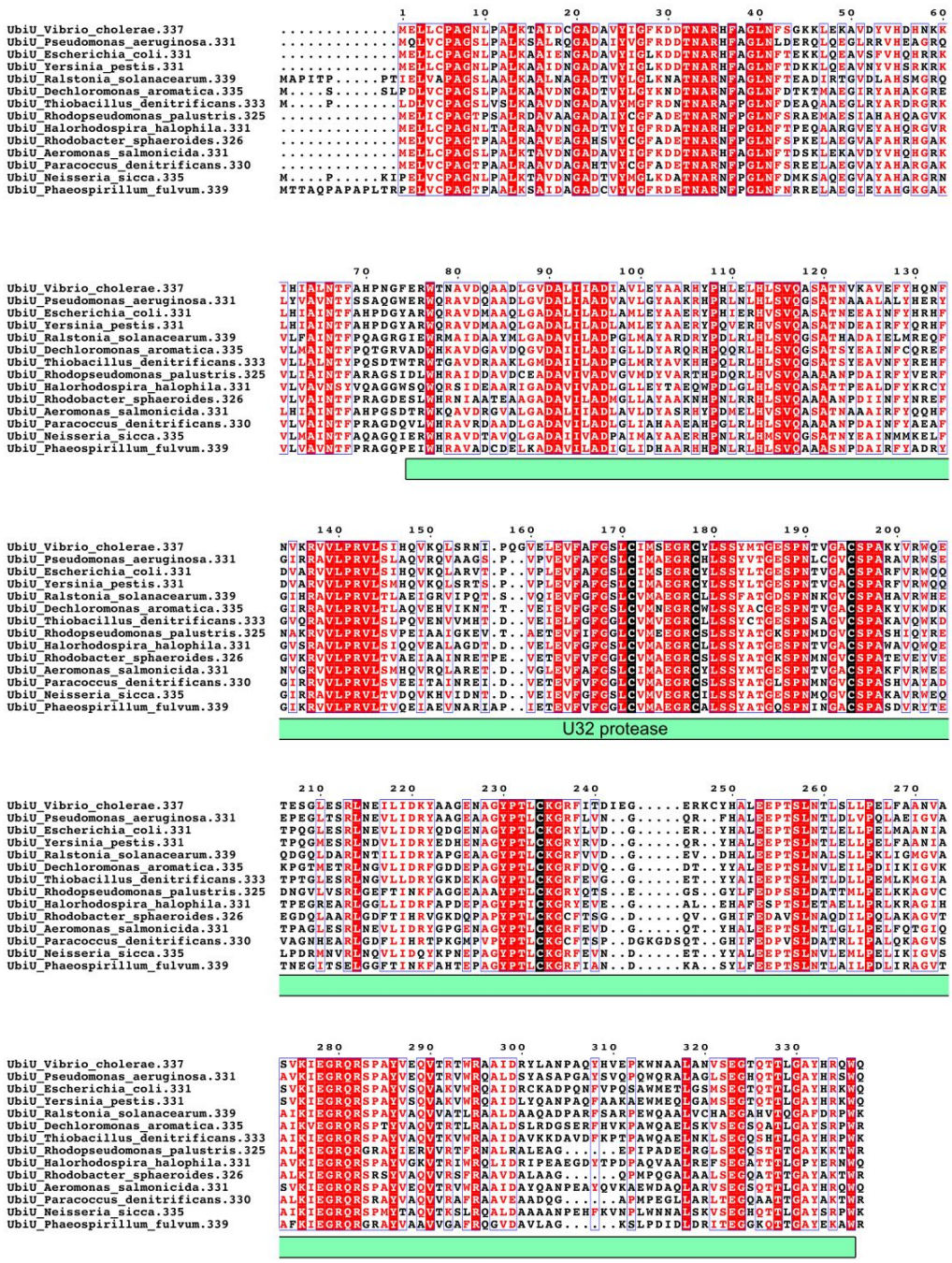


Figure S3: Multiple sequence alignment of UbiU from representative proteobacteria. Sequences were aligned using Mafft (linsi), and the output generated using ESPript and Inkscape. The four conserved cysteines (C169, C176, C193 and C232) involved in iron-sulfur binding are indicated by black columns. The position of the domain U32 protease PF01136 is indicated by a green box. GenBank accession numbers: *Vibrio cholerae*, NP_230301.1, *Pseudomonas aeruginosa*, NP_252602.1, *Escherichia coli*, NP_312066.1, *Yersinia pestis*, YP_002348368.1, *Ralstonia solanacearum*, WP_011004262.1, *Dechloromonas aromatica*, WP_011285878.1, *Thiobacillus denitrificans*, WP_011312698.1, *Rhodopseudomonas palustris*, WP_011666114.1, *Halorhodospira halophila*, WP_011813833.1, *Rhodobacter sphaeroides*, WP_011909346.1, *Aeromonas salmonicida*, WP_005310394.1, *Paracoccus denitrificans*, WP_041530457.1, *Neisseria sicca*, WP_049226489.1, *Phaeospirillum fulvum*, WP_074767209.1.

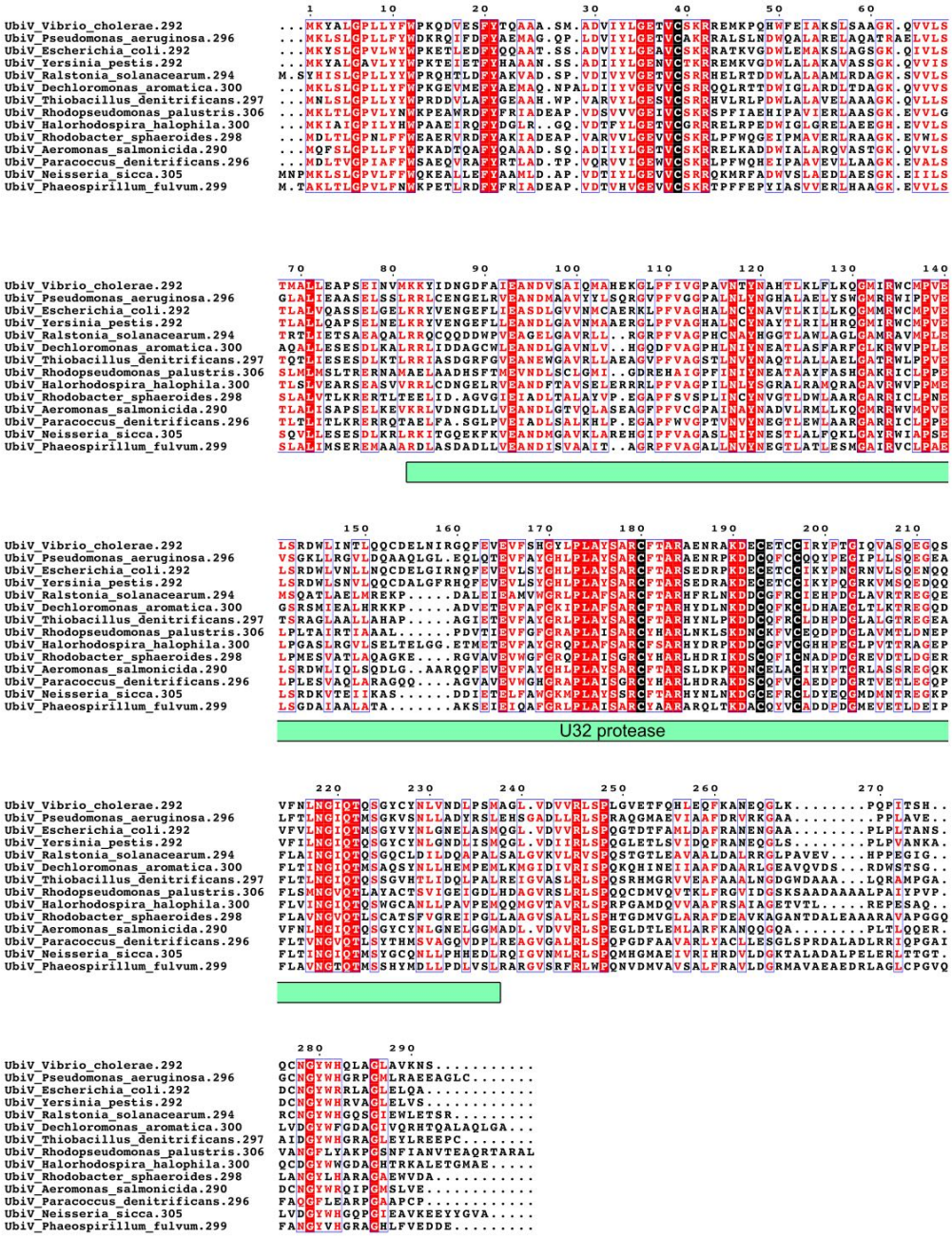


Figure S4: Multiple sequence alignment of UbiV from representative proteobacteria. Sequences were aligned using Mafft (linsi), and the output generated using ESPript and Inkscape. The four conserved cysteines (C39, C180, C193 and C197 - positions in *E. coli*) involved in iron-sulfur binding are indicated by black columns. The position of the domain U32 protease PF01136 is indicated by a green box. GenBank accession numbers: *Vibrio cholerae*, NP_230300.2, *Pseudomonas aeruginosa,* NP_252601.1, *Escherichia coli*, NP_312067.2, *Yersinia pestis*, YP_002348369.1, *Ralstonia solanacearum,* WP_011004261.1, *Dechloromonas aromatica,* WP_011285877.1, *Thiobacillus denitrificans,* WP_011312697.1, *Rhodopseudomonas palustris,* WP_011666113.1, *Halorhodospira halophila,* WP_011813832.1, *Rhodobacter sphaeroides,* WP_011909345.1, *Aeromonas salmonicida,* WP_005310392.1, *Paracoccus denitrificans,* WP_011750458.1, *Neisseria sicca,* WP_080614296.1, *Phaeospirillum fulvum,* WP_074767477.1.

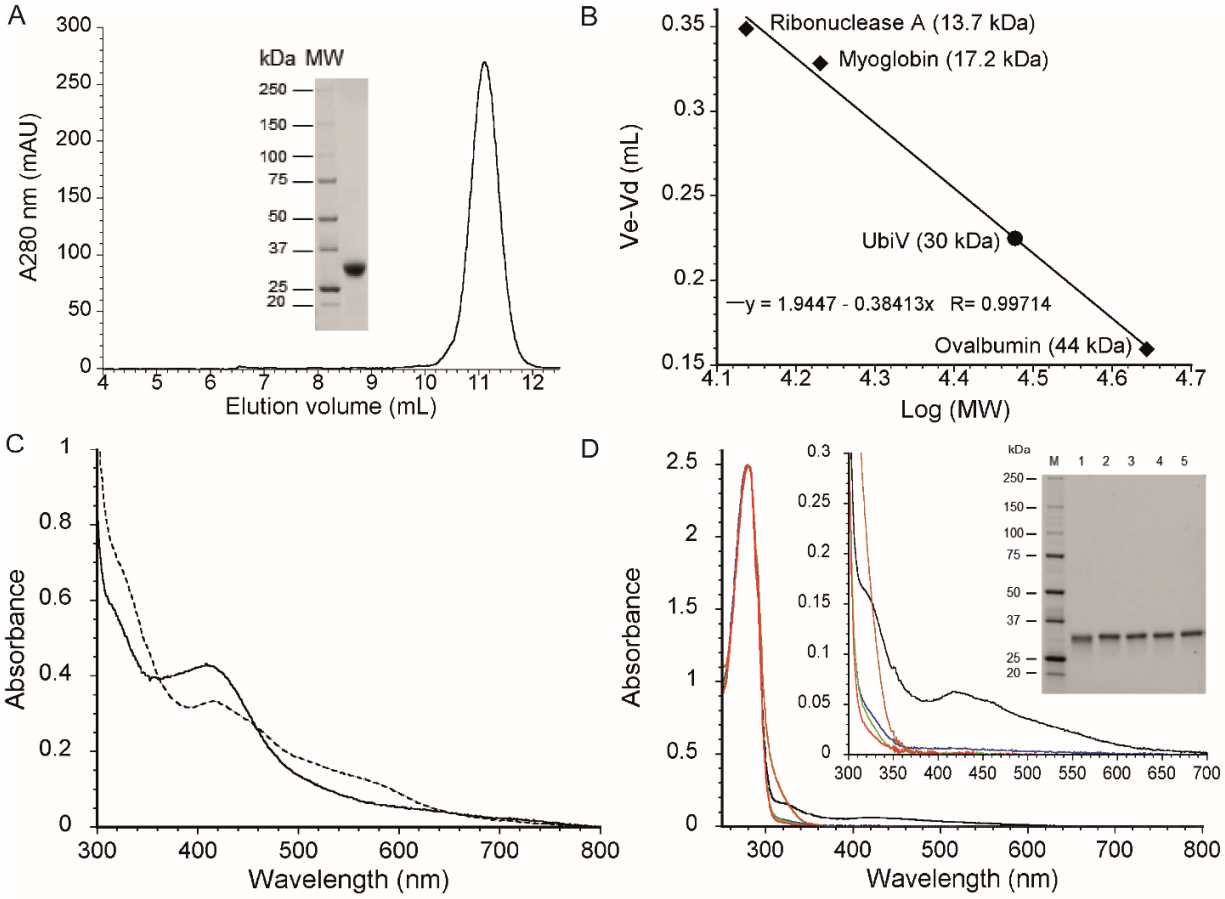


Figure S5: A) Gel filtration chromatogram of aerobically purified UbiV on a Superdex 75 Increase 10/300 GL column; Inset: Coomassie staining SDS-PAGE of aerobically purified UbiV. B) Calibration curve, standard proteins in rhombi, UbiV in circle. C) UV-visible absorption spectra of 41 µM holo-UbiV under anaerobiosis (solid line) and after 20 minutes exposure to air (dotted line). D) Comparative UV-visible absorption spectra of aerobically purified wild-type (black) and different Cys-to-Ala mutants of UbiV (C193A C197A in blue; C39A C193A C197A in green; C180A C193A C197A in pink; C39A C180A C193A C197A in brown). Inset: Coomassie staining SDS-PAGE shows the purifications of UbiV and variants. Lanes are: M, molecular mass marker; 1, wild-type UbiV; 2, UbiV C193A C197A; 3, UbiV C39A C193A C197A; 4, UbiV C180A C193A C197A; 5, UbiV C39A C180A C193A C197. All proteins were in 50 mM Tris-HCl, pH 8.5, 25 mM NaCl, 15% glycerol, 1 mM DTT (A-D).

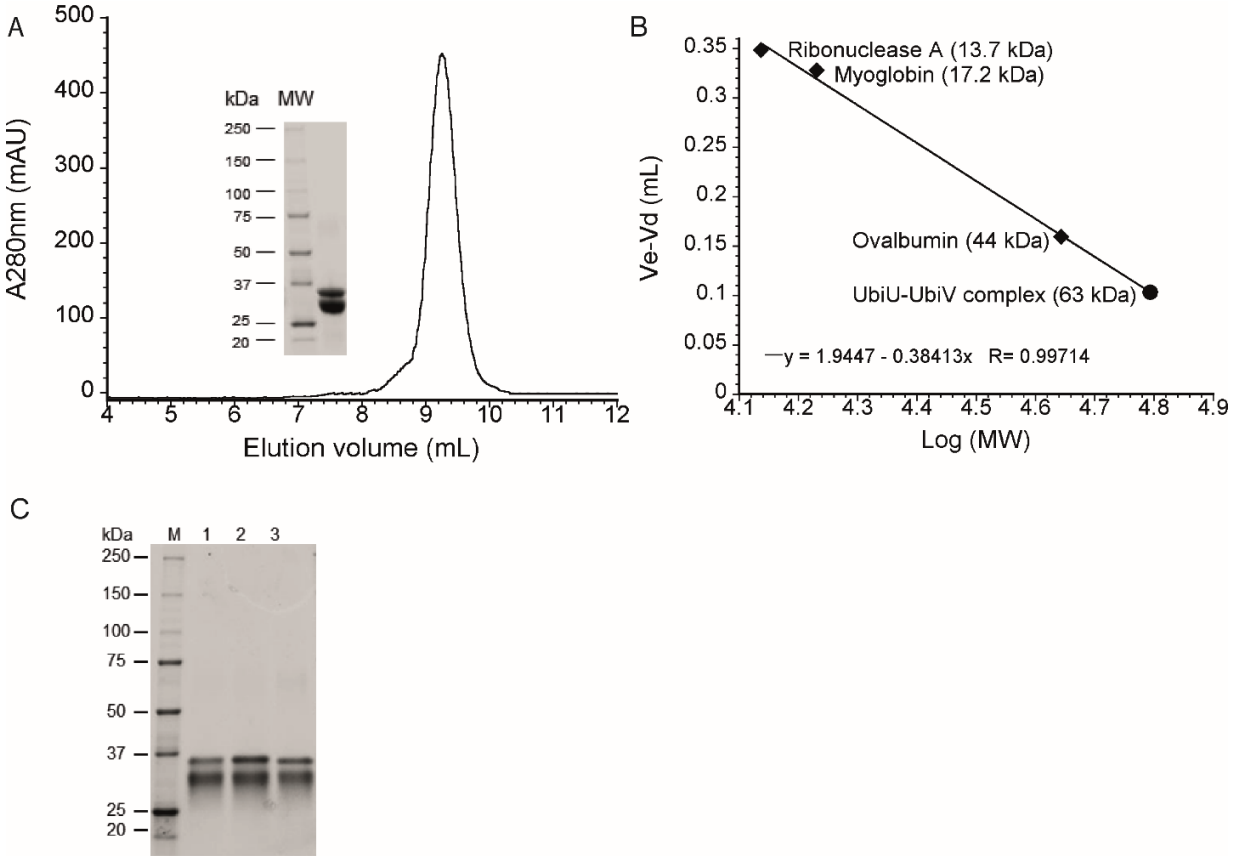


Figure S6: A) Gel filtration chromatogram of aerobically purified UbiU-UbiV on a Superdex 75 Increase 10/300 GL column; Inset: Coomassie staining SDS-PAGE of aerobically purified UbiU-UbiV complex. B) Calibration curve, standard proteins in rhombi, UbiU-UbiV in circle. C) Coomassie staining SDS-PAGE shows the purifications of UbiU-UbiV wild-type and variants proteins: Lanes are: M, molecular mass marker; 1, wild-type UbiU-UbiV complex; 2, UbiU C169A C176A - UbiV wild-type; 3, UbiU C193A C232A - UbiV wild-type. All proteins were in buffer 50 mM Tris-HCl, pH 8.5, 150 mM NaCl, 15% glycerol, 1 mM DTT (A-D).

1033 **Table S1:** UQ levels in strains from the medium- and large- deletion collections

1034 **Table S2:** Occurrence of genes from the UQ pathways in bacterial genomes.

1035 **Table S3:** List of oligonucleotides and strains used in this study

1036 **Table S4:** References of the protein sequences used to build the HMM profiles.

1037