Identification of a Highly Expressed Gene Cluster Likely Coding for Benzene Activation Enzymes in a Methanogenic Enrichment Culture

Courtney R. A. Toth[1], Olivia Molenda[1], Camilla Nesbø[1,2], Fei Luo[1‡], Cheryl E. Devine[1], Xu Chen[1], Kan Wu[1], Johnny Z. Xiao[1], Rishika Puri[1], Shen Guo[1], Nancy Bawa[1], Po-Hsiang Wang[1,3], Yifeng Wei[4], Robert Flick[1], and Elizabeth A. Edwards[1*]

[1]Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, Ontario, M5S 3E5, Canada.

[2]Department of Biological Sciences, University of Alberta, Edmonton, Alberta, T6G 2E9, Canada

[3]Graduate Institute of Environmental Engineering, National Central University, Taoyuan City 32001, Taiwan (R.O.C.)

[4]Singapore Institute of Food and Biotechnology Innovation, Agency for Science, Technology and Research (A*STAR), Singapore, 138669, Singapore

[‡]Current Address: Liven Proteins Corporation, Mississauga, Ontario, L5L 1C6, Canada

[*]For correspondence, e-mail: elizabeth.edwards@utoronto.ca; Tel. (+1) 416 946 3506; Fax (+1) 416 978 8605.

Running Title: Anaerobic benzene activation enzymes

Keywords: Benzene, *Desulfobacterota*, ORM2, "*Candidatus* Anaerobenzenivorax", Proteomics, Bioremediation, Biodegradation, Anaerobic Hydrocarbon Degradation, Methanogenesis

ABSTRACT

The Oil Refinery (OR) consortium is a model methanogenic enrichment culture used to study anaerobic benzene degradation. Over half of the culture's bacterial community consists of two closely related *Desulfobacterota* strains, designated ORM2a and ORM2b, whose mechanisms of benzene activation are unknown. Two new metagenomes, including a complete circularized metagenome-assembled genome (MAG) for ORM2a, enabled a thorough investigation of this culture's proteome. Among the proteins identified were Bam-like subunits of an ATP-independent benzoyl-CoA degradation pathway, as well as downstream β-oxidation proteins yielding acetate. The most abundant proteins identified mapped to two ORM2a gene clusters of unknown function. Homologous and syntenic gene clusters were identified in genomes of ORM2b and a sulfate-reducing *Pelotomaculum* that also degrades benzene, as well as in nine contigs assembled from hydrothermal vent metagenomes. Extensive homology and structural predictions suggest that the first cluster – termed the "Magic" gene cluster – encodes for enzymes catalyzing the chemically challenging activation of benzene and subsequent transformation steps yielding benzoyl-CoA. The second ("Nanopod") gene cluster encodes a transmembrane complex that may facilitate benzene transport across the cell membrane. Phylogenomic analyses place ORM2a and ORM2b within a novel genus of strict anaerobes specialized for benzene degradation, which we propose naming "*Candidatus* Anaerobenzenivorax".

41 **IMPORTANCE**

42 Benzene is a widespread, persistent and toxic pollutant that can accumulate in anoxic environments
43 such as groundwater and sediments. Despite decades of study, the biochemical mechanisms by
44 which benzene is activated under anaerobic conditions remain unproven. This study provides
45 strong genetic and proteomic evidence for a new class of enzymes that initiate anaerobic benzene
46 activation and proposes a preliminary model for their underlying biochemistry. These findings lay
47 a foundation for future biochemical studies and expand our understanding of how microbes carry
48 out extreme redox chemistry in the absence of oxygen.

INTRODUCTION

It has been 45 years since anaerobic benzene degradation was first documented by Ward et al. (1), reversing a long-held tenet that aromatic hydrocarbons were recalcitrant under anaerobic conditions. Notorious for its carcinogenicity, stringent environmental regulations, and persistence in anoxic environments, benzene has been the subject of numerous studies seeking to characterize its metabolism in the absence of oxygen (2, 3). Laboratory enrichment culture studies have confirmed unequivocally that anaerobic benzene degradation can be coupled to iron reduction, nitrate reduction, sulfate reduction, and methanogenesis, catalyzed by a handful of specialized microbial clades (2, 4). But despite great efforts, the biochemical mechanism(s) by which benzene is activated in the absence of oxygen remain elusive.

Initial studies using ^{13}C- or ^{14}C-benzene and occasionally $H_2{}^{18}O$ identified three possible activation reactions based on the detection of labelled metabolites: (i) anaerobic hydroxylation of benzene yielding phenol (5-12), (ii) carboxylation of benzene yielding benzoate (6, 8-10, 13, 14), and (iii) methylation of benzene yielding toluene (10). However, neither phenol nor benzoate are diagnostic of a single activation mechanism (15), and supporting evidence of benzene methylation is sparse in literature (16, 17). From 2003-2008, hydrogen and carbon isotope fractionation ratios were evaluated in active benzene-degrading enrichment cultures (18-20). The slope of dual isotope plots ($\Delta\delta\,^{2}H/\Delta\delta\,^{13}C$, a value defined as Λ) from nitrate-reducing cultures (Λ = 8 to 19, n = 5) was substantially lower that than for sulfate-reducing and methanogenic cultures surveyed (Λ = 28 to 31, n = 7), hinting that at least two distinct biochemical benzene activation mechanisms exist in nature (18).

By 2010, advances in DNA sequencing and analysis enabled meta'omic characterization of anaerobic benzene degraders for the first time (21-23). Abu Laban *et al*. (24) sequenced the

metagenome of the iron-reducing enrichment culture BF and identified a protein with homology to a known phenylphosphate carboxylase (PpcA). The corresponding putative anaerobic benzene carboxylation gene (*abcA*) was located on a contig affiliated with *Clostridia*. Expression of the putative *abcA* and neighboring genes (*abcD*, benzoyl-CoA ligase [*bzlA*] and a *ubiX*-like gene) was induced with benzene or benzoate, but not with phenol, suggesting a possible role in benzene activation (24). Syntenic gene clusters have since been identified in the transcriptomes of several nitrate-reducing cultures (25-27), including one previously analyzed by isotopic fractionation (18, 19). Most nitrate- and iron-reducing benzene degraders identified to date have been affiliated with the order *Thermincolales* (25-29); a putative benzene-carboxylating, sulfate-reducing *Thermincolales* was also recently identified in a jet fuel-contaminated aquifer via metagenomics (30). Researchers are now attempting to verify benzene carboxylation using structural and biochemical assays, among other tools (31).

Evidence for benzene hydroxylation has only been shown for one microorganism, *Geobacter metallireducens* GS-15, via gene deletion. In 2013-2014, Zhang *et al.* (5, 23) identified a handful of genes that were upregulated during active phenol or benzene metabolism, and deletion of key genes (Gmet0231, Gmet0232, or *ppsA*) inhibited benzene degradation. The gene *ppsA* encodes for the alpha subunit of phenylphosphate synthase, which catalyzes the first step in anaerobic phenol metabolism (32, 33). Functional annotations suggest that the products of Gmet0231 and Gmet0232 encode oxidoreductases that could potentially catalyze the addition of water to benzene, forming phenol. Further investigation into the exact functions of these enzymes is needed. In 2024, Bullows *et al.* (17) identified a fumarate addition gene cluster consisting of a 3-hydroxybenzylsuccinate synthase and a methyltransferase in the genome of *Geotalea daltonii* strain FRC-32, a close relative of *Geobacter*. Expression of these genes – previously known only

95 to catalyze activation of alkanes and alkyl-substituted aromatic rings (34) – was upregulated
96 during active benzene metabolism but not with toluene. Addition of benzene to *G. daltonii* whole-
97 cell lysates resulted in the near-stoichiometric formation of toluene, but curiously benzene was
98 completely depleted before toluene was formed (17). Whether this gene cluster truly encodes for
99 benzene methylation enzymes remains uncertain.

100 Homologs of known and predicted anaerobic aromatic-activating genes have not been
101 detected in methanogenic benzene-degrading enrichment cultures (4, 21) nor in the well-
102 characterized sulfate-reducing enrichment culture BPL (9, 35). Genes associated with aerobic
103 metabolism (monooxygenases) are not found in the major taxa in these mixed cultures, although
104 recent growth experiments found that aerobic benzene degraders may reside in these cultures at
105 trace abundances (36). Additionally, these enrichments do not possess genes for the first step of
106 anaerobic growth on benzoate (benzoyl-CoA ligase) nor exhibit growth on benzoate (21, 35). To
107 date, metagenomic and proteomic surveys of representative cultures have identified the presence
108 and expression of genes implicated in the ATP-independent metabolism of benzoyl-CoA and
109 downstream β-oxidation intermediates (21, 35). Given this information and previous isotope
110 fractionation data (18, 19) indicating that the mechanism of benzene activation catalyzed by
111 methanogenic cultures is distinct from that in nitrate-reducing cultures – where *abcA* genes are
112 present and highly expressed (25-27), and carboxylation is likely – we surmise that a novel
113 hydrocarbon activation mechanism catalyzing anaerobic transformation of benzene to benzoyl-
114 CoA must exist in our methanogenic culture.

115 In this study, we integrated results from three proteomic experiments on the methanogenic
116 benzene-degrading OR consortium. The culture, maintained on benzene for over 25 years, harbors
117 two closely related strains of benzene-degrading *Desulfobacterota* (ORM2a and ORM2b)

118 consistently detected in high abundance (4, 21, 36-41). Although the relative abundance of each
119 ORM2 strain varies by subculture (39), together they account for 48-92% of bacteria in this
120 consortium (Table S1). They are the only microorganisms whose growth directly coincides with
121 anaerobic benzene degradation activity, and no substrate other than benzene has been shown to
122 support their growth (4, 21, 36, 39). The OR culture also contains methanogenic archaea
123 (*Methanothrix* and *Methanoregula*) and over 100 low abundance microorganisms with various
124 predicted functions (39, 40, 42). Initial proteomic studies were conducted on the OR consortium
125 in the early 2010's (21) but results were difficult to interpret due to the lack of a sound metagenome
126 and low protein yields. In the years since, improvements in culture maintenance (36, 39) and new
127 sequencing efforts (40) have enabled the reconstruction of two high quality metagenomes and 74
128 metagenome-assembled genomes (MAGs; Table S2). The MAG of ORM2a was successfully
129 closed (NCBI accession no. CP113000.1), providing a valuable blueprint for deciphering
130 anaerobic benzene degradation. Here, we discovered that the most abundant proteins expressed by
131 the culture mapped to a unique cluster of syntenic genes found only in other benzene-degrading
132 obligate anaerobes and in a handful of metagenomic contigs from deep hydrocarbon-producing
133 hydrothermal vents.

134

135 **RESULTS**

136 **Features of the ORM2a MAG**

137 The ORM2a MAG consists of a single circular chromosome of 3,300,546 bp with low GC skew
138 and no discernible terminus region (Figures 1 and S1). General genome statistics are summarized
139 in Table S3. To assess metabolic potential, coding sequences predicted by the NCBI Prokaryotic
140 Genome Annotation Pipeline were uploaded to BlastKOALA for automatic KEGG Orthology

141 (KO) assignment (43). Forty-six percent of genes identified (1,431 of 3,135) were assigned KO
142 identifiers and mapped into metabolic pathways using the KEGG Mapper Reconstruct tool (44,
143 45). Key pathways and genes of interest are highlighted below and in Figure 2; additional pathway
144 and genomic features are detailed in Tables S4-S9 and Supplementary Text S1.

145 The ORM2a MAG contains few and mostly incomplete pathways for metabolizing
146 carbohydrates, lipids, amino acids, and nucleic acids. Similarly, pathways for electron transport
147 and ATP generation are sparse, comprising an F-type ATPase, a Nuo-type NADH dehydrogenase
148 with one missing subunit (*nuoG*), and a complete dissimilatory sulfate reduction pathway (Figure
149 2). This finding supports recent experimental evidence that the OR consortium can couple benzene
150 oxidation to sulfate reduction in addition to methanogenesis (4). BLAST searches and manual
151 inspections confirmed the absence of homologs to known hydrocarbon-activating genes and to
152 benzoyl-CoA ligase. The latter is a significant distinction between microorganisms containing
153 putative anaerobic benzene carboxylase genes like *abcA* and those that do not (35). The presence
154 of a complete ATP-independent benzoyl-CoA degradation pathway is consistent with ORM2a's
155 involvement in anaerobic benzene oxidation (Figures 1 and 2).

156 Additional genomic evidence suggests that ORM2a is capable of fermentative metabolism
157 in syntrophic association with methanogenic archaea (Figure 2). Three gene clusters encoding
158 energy-converting hydrogenase (Ech) and formate dehydrogenase (Fdh) complexes were
159 identified that may facilitate interspecies electron transfer (46, 47). The MAG also encodes for a
160 cytoplasmic Mvh-type F_{420}-non-reducing hydrogenase – potentially used for hydrogenotrophic
161 growth on carbon dioxide (48, 49) – and a putative sulfur-reducing hydrogenase operon (Table
162 S6). Sulfhydrogenases (also called sulfur-reducing hydrogenases) are poorly characterized
163 enzymes that can catalyze the reduction of elemental sulfur (S^0) or polysulfides (S_n^{2-}) to hydrogen

164 sulfide (H_2S) using H_2 as an electron donor (50) (Figure 2). This process has been most thoroughly
165 described in *Pyrococcus furiosus* and related thermophilic sulfur-oxidizing archaea (50-53). Under
166 certain conditions, some sulfhydrogenases may also catalyze the reverse reaction, producing H_2
167 (50, 52). In the OR culture, the major sulfur source in the growth medium is ferrous sulfide (FeS)
168 used as a reductant (38). It is unknown whether sulfhydrogenases can oxidize FeS or reduce
169 oxidized forms such as FeS_2. Notably, no visible S^0 or polysulfide formation (e.g., yellow solids)
170 have ever been observed in the 25+ years this culture has been maintained.

171 A draft MAG belonging to ORM2b (designated ORM2b_MAG4) was assembled into 378
172 contigs and is estimated to be 74.6% complete (40). Although the exact taxonomic identity of
173 ORM2b_MAG4 remains uncertain due to the absence of a 16S rRNA gene – precluding direct
174 comparison to a reference clone sequence of ORM2b (NCBI accession no. KT025834) – the high
175 average nucleotide identity (89.94%) between ORM2a and ORM2b_MAG4 supports its correct
176 classification (Figure 1, ring 5; Figure S2). No other MAGs recovered from the OR consortium
177 metagenomes display significant nucleotide identity to ORM2a.

178 **Proteomic analyses**

179 Between 2010 and 2018, three proteomics experiments were performed on subcultures from the
180 same OR culture lineage (OR-b1C, OR-b, and OR-b1A; see Figure S3). Experiments #1 (2010)
181 and #2 (2011) used LC-MS/MS to analyze trypsin-digested crude extracts and recovered 153
182 unique proteins (Table 1) – a very low yield. To improve detection, protein separation by gel
183 electrophoresis was introduced in Experiment #3 (2018). Duplicate crude extracts were prepared:
184 one was directly digested with trypsin (Experiment #3a), as in previous protocols, and yielded
185 similarly few proteins (Table 1). The second crude extract (Experiment #3b) was separated on a

186 15% SDS-PAGE gel, sliced into five equal sections (Figure S4), and subjected to in-gel trypsin
187 digestion and LC-MS/MS. This method yielded over 800 unique proteins (Table 1).

188 In total, 858 unique proteins were identified across all experiments (Table 1). Protein
189 accession numbers, automated annotations, peptide and spectral counts, and amino acid sequences
190 are provided in Table S10, with summary statistics in Tables S11a and S11b. Proteins that mapped
191 to one or more MAG, are putatively involved in anaerobic benzene degradation, and/or were
192 detected in high abundance in Experiment #3b (≥100 spectra) are highlighted in Table 2 and
193 Figures 1-4. Notably, the most high-abundant proteins detected in Experiment #3b were also
194 observed in Experiments #1 and #2 (Table S10), demonstrating strong reproducibility of our
195 results despite years between tests and differences in extraction and analysis methods (Table 1).

196 **Key microorganisms identified in the OR proteomes**

197 Half of all proteins identified mapped to the MAGs of ORM2a and/or ORM2b (Tables 1 and 2).
198 Of these, 348 proteins mapped exclusively to ORM2a (41%), 33 proteins mapped exclusively to
199 ORM2b_MAG4 (4%), and 44 proteins (5%) mapped to both MAGs (Table 1). Most
200 ORM2b_MAG4 proteins identified shared 71-100% amino acid identity to a homologous ORM2a
201 protein or undetected ORM2a gene products. Only six low abundance proteins – associated with
202 protein synthesis or hypothetical functions – were unique to ORM2b_MAG4 (Table S10). Given
203 these similarities, the following sections focus on ORM2a proteins and metabolic pathways of
204 interest.

205 One-quarter of proteins identified mapped to MAGs of methanogenic archaea (Tables 1
206 and S10), specifically acetoclastic *Methanothrix* spp. (170 proteins across 8 MAGs) and
207 hydrogenotrophic *Methanoregula* spp. (19 proteins across 3 MAGs; ~89% fewer than

208 *Methanothrix*). Most high abundance archaeal proteins were affiliated with methanogenesis and
209 mapped to more than one MAG (Table 2). Forty-four low-abundance proteins (5%) mapped to
210 other bacterial MAGs including "*Ca.* Nealsonbacteria" DGGOD1a. The remaining proteins could
211 not be assigned to a specific MAG (180 proteins, 21%) or mapped to contaminant/decoy sequences
212 (20 proteins, 2%).

213 **Benzoyl-CoA degradation to acetyl-CoA**

214 We next searched for enzymes involved in downstream aromatic degradation pathways and
215 detected a complete suite of putative benzoyl-CoA degradation proteins encoded on the ORM2a
216 genome (Figures 1 and 2; Table 2). Ten gene products (designated Proteins #1-10) mapped to a
217 single gene cluster and were annotated as subunits of a Bam-type benzoyl-CoA reductase
218 (BamBCDEFDHI; whose components are described in Wischgoll et al. (54) and Boll et al. (55)),
219 an FdhD-like insertase, and a cyclohexa-1,5-diene-1-carbonyl-CoA hydratase (BamR). Two high
220 abundance proteins annotated as 6-hydroxycyclohex-1-ene-1-carbonyl-CoA dehydrogenase
221 (BamQ, Protein #11) and 6-oxo-cyclohex-1-ene-carbonyl-CoA hydrolase (BamA, Protein #12)
222 were encoded elsewhere in the ORM2a genome (Figure 1). Multiple sequence alignments (Table
223 S12) and maximum likelihood phylogenetic tree analyses (Figure S5) support the functional
224 annotation of all predicted Bam proteins. The only other proteins identified that likely participate
225 in anaerobic benzoyl-CoA degradation were three *bam* genes mapping to ORM2b_MAG4 (Table
226 2). These results corroborate longstanding evidence that ORM2a and ORM2b are key benzene-
227 degrading members of the OR consortium (4, 21, 36, 39, 41).

228 Nine proteins predicted to catalyze the transformation of 3-hydroxypimelyl-CoA to acetyl-
229 CoA were also identified and mapped exclusively to the ORM2a MAG (Figures 1-2; Table 2). In
230 our most successful proteome (Experiment #3b), proteins involved in converting 3-

231 hydroxypimelyl-CoA to glutaconyl-CoA (Proteins #13-17) had higher spectral counts than
232 downstream beta-oxidation enzymes (Proteins #18-21). This observation may indicate a metabolic
233 branch point (Figure 2), as discussed later. Notably, the gene encoding a non-decarboxylating
234 glutaryl-CoA dehydrogenase (Protein #15) is located in the same operon as two genes encoding
235 bifurcating electron transfer flavoproteins (EtfAB-like; Proteins #16-17) (Table 2). EtfAB
236 typically mediate flavin-based electron bifurcation and confurcation reactions, which are used to
237 drive an endergonic redox reaction by coupling it with an exergonic one (56). As illustrated in
238 Figure 2, the EtfAB-like complex may couple the endergonic reduction of NAD^+ with the
239 exergonic oxidation of glutaryl-CoA and ferredoxin, in a manner analogous to EtfAB-butyryl-CoA
240 dehydrogenase (56, 57).

241 **Sulfur metabolism and acetate production**

242 Several unexpected ORM2a proteins were detected in the OR proteomes, including six associated
243 with the dissimilatory reduction of sulfate to hydrogen sulfide – specifically sulfate
244 adenylyltransferase (Sat), adenylylsulfate reductase (AprAB), and dissimilatory sulfite reductase
245 components (DsrAB, DsrC) – despite not adding any sulfate to the culture medium (Figure 2).
246 Additionally, products of the putative sulfhydrogenase operon (HydADBG) were the only
247 hydrogenase proteins detected that mapped to the ORM2a genome (Figure 2 and Table S6). If this
248 enzyme consumes H_2 released during benzene oxidation, it could explain why so few proteins
249 from hydrogenotrophic methanogens (*Methanoregula* spp.) were recovered from the OR
250 proteomes (Table S10). Together, these findings raise new questions about ORM2a's sulfur
251 metabolism that warrant further investigation.

252 The protein folding chaperonin GroEL (Protein #42) and a putative AMP-forming acetyl-
253 CoA synthetase (Acs, #43) were tied as the 5th most abundant proteins in Experiment #3b (Table

254 2). Acs normally catalyzes the ATP-dependent formation of acetyl-CoA molecules from acetate
255 and coenzyme A but given that acetyl-CoA is already produced in abundance during anaerobic
256 benzene metabolism, we surmise that Protein #43 catalyzes the reverse reaction yielding acetate
257 (a substrate for acetoclastic *Methanothrix* spp.), coenzyme A, and ATP (Figure 2). Similar reverse
258 reactions were proposed for the anaerobic benzoate degrader *Syntrophus acidotrophicus* (58, 59)
259 and three syntrophic bacterial MAGs reconstructed from full-scale anaerobic digesters (60). None
260 of these genomes code for acetate kinase or closely related homologs needed to phosphorylate
261 acetate. ORM2a also lacks cytosolic pyrophosphatase but harbors a membrane-bound Na^+-
262 translocating pyrophosphatase (WAC07798.1), suggesting similarly elevated intracellular
263 pyrophosphate levels that may favor ATP production in syntrophic organisms (59). Phylogenetic
264 trees support the correct annotation of Protein #43 (Figure S6). An unclassified porin (Protein #44)
265 whose corresponding gene is localized adjacent to *acs* may facilitate acetate export from ORM2a
266 cells (Figure 2).

267 **Highly expressed ORM2a proteins from gene clusters of unknown function**

268 The three most abundant and consistently detected proteins in all OR proteomes (Proteins #24-26)
269 mapped consecutively to an ORM2a gene cluster of unknown function (Table 2; Figures 1, 3 and
270 4a). Products of several neighboring genes (Proteins #22-23, #27-30) were also detected in
271 multiple OR proteomes. This region, designated herein as the "Magic" gene cluster, is located on
272 a small genomic island flanked by multiple predicted transposases and pseudogenes (Figure 1;
273 Table S7). Five operons were predicted within the "Magic" gene cluster, each encoding for two to
274 three proteins (Figure 4a). Automated annotations from NCBI and IMG provided few insights into
275 their functions (Table 2) with one notable exception: the tenth consecutive gene in this cluster most
276 likely encodes 6-hydroxycyclohex-1-ene-1-carbonyl-CoA dehydrogenase (BamQ, Protein #11), a

277 known component of the benzoyl-CoA degradation pathway (Figure 2). The detection of a
278 neighboring transcription-terminating Rho factor in Experiment #3b (Protein #31) suggests that
279 expression of the "Magic" gene cluster is tightly regulated.

280 Two other highly abundant proteins (#32 and #33, ranked 18th and 4th in Experiment #3b,
281 respectively) mapped to a second ORM2a genomic island of unknown function, designated the
282 "Nanopod" gene cluster (Tables 2 and S7; Figures 1-3), which includes four other genes whose
283 products were detected. Nanopods are tubular structures (<100 nm in diameter) that may facilitate
284 nutrient transport across cell membranes or export toxicants (61, 62). Shetty et al. (61) found that
285 phenanthrene induced nanopod formation in *Delftia acidovorans* Cs1-4 but not structurally similar
286 hydrocarbons (naphthalene and toluene), and that deletion of its DUF1329 gene (DelCs14_1722)
287 impaired growth on phenanthrene. BLASTP alignments suggest that Protein #33 is distantly
288 related to the translated DelCs14_1722 gene sequence (21% amino acid identity), but the high
289 expression of this ORM2a gene cluster and its possible relation to nanopod formation is intriguing.

290

291 **Comparative analyses of the "Magic" and "Nanopod" gene clusters to other genomes and**
292 **proteomes**

293 To determine whether other organisms harbor syntenic regions to the "Magic" and "Nanopod"
294 gene clusters, translated protein sequences were searched against all reference genomes in NCBI
295 and IMG databases (as of August 2024) using BLASTP. Four syntenic contigs were identified
296 (Figure 3; Table 3), along with six shorter contigs (comprising 2-3 genes) containing homologous
297 genes with 51-76% amino acid identity to those in the "Magic" gene cluster (Table S12). These
298 genes also appear to be grouped into operons identical to those in the ORM2a genome (Figures 3-

299 4a). Nine of the ten contigs identified were affiliated to *Deltaproteobacteria* spp. from
300 hydrothermal vent systems (63, 64). Hydrocarbons are important energy sources for microbial
301 communities in these ecosystems and are produced primarily by geological processes (65).
302 Hydrothermal vent MAG *Deltaproteobacteria* AUK045 is notable for containing syntenic contigs
303 to both the "Magic" and "Nanopod" gene clusters (Figure 3; Table 3). Additionally, genes
304 encoding predicted downstream benzoyl-CoA degradation and beta-oxidation pathway
305 components were identified in AUK045 and two other draft *Deltaproteobacteria* bins (Table S12),
306 suggesting metabolic potential for aromatic compound degradation.

307 *Pelotomaculum* candidate BPL (IMG GOLD accession number Ga0073689), the dominant
308 benzene degrader in a sulfate-reducing enrichment culture from Poland (9, 35), provides another
309 compelling comparison. Dong et al. (35) reconstructed this genome in 2017 from metagenome
310 data (2,986,7007 bp across 97 contigs, ~99% completeness) and published concurrent proteomic
311 data from an active benzene-degrading subculture. We identified a highly expressed
312 *Pelotomaculum* contig (Ga0073689_145) producing eight significant alignments to the "Magic"
313 gene cluster, exhibiting 64-76% amino acid identity and e-values ≤ 1.0E-178 (Figure 3; Table 3).
314 The Ga0073689_145 contig also encodes a predicted benzoyl-CoA reductase subunit (BamB) and
315 an enoyl-CoA hydratase, drawing parallels to the "Magic" gene neighborhoods in the ORM2a,
316 ORM2b, and *Deltaproteobacteria* AUK045 MAGs (Figure 3). These genomic and proteomic
317 parallels are striking and suggest a conserved role for the "Magic" cluster in anaerobic benzene
318 degradation.

319 **Predicted functions of the "Magic" proteins**

320 Multiple sequence- and structure-based protein prediction tools were used to investigate the
321 potential biochemical roles of proteins encoded by the "Magic" and "Nanopod" gene clusters.

322 Search results for the "Magic" proteins are summarized below, in Figure 4b and in Tables S13-
323 S14. Structure predictions for the "Magic" proteins and putative enzyme complexes are shown in
324 Figures 5 and S7-S8. Phylogenetic trees were also constructed for each protein; however, most
325 were too distantly related to characterized enzymes to yield meaningful functional insights.

326 Operon #1 of the "Magic" gene cluster encodes Proteins #22 and #23 (Figure 4a). All
327 prediction tools converged on the annotation of Protein #22 as a novel AMP-forming acyl-CoA
328 synthetase (Figure 4b). Phylogenetic analysis placed Protein #22 and its homologs from other
329 anaerobic benzene degraders in a distinct subfamily of acyl-CoA synthetases, separate from well-
330 characterized enzymes such benzoyl-CoA ligase and acetyl-CoA synthetase (Figure S6). DeepFRI
331 Gene Orthology (GO) predictions (66) further suggested that Protein #22 can bind to organic cyclic
332 compounds such as benzene (Table S14b). Since Proteins #22 and #43 are predicted to use similar
333 cofactors (e.g., ATP/AMP, CoA), these reactions may be coupled, with one reaction's energy
334 driving the other (Figure 2). Protein #23 contains a canonical "hot dog" domain typical of (*R*)-
335 specific enoyl-CoA hydratases (67) and is likewise predicted to bind cyclic organic molecules
336 (Table S13 and S14b). Enoyl-CoA hydratases catalyze the second step of β-oxidation (68),
337 converting trans-2-enoyl-CoA molecules into 3-hydroxyacyl-CoA. Structure predictions for
338 Proteins #22 and #23 are shown in Figures S7a and S7b.

339 Operon #3 encodes Proteins #27 and #28, predicted to form a class III CoA-transferase
340 complex (Figures 4b and S7c). These enzymes operate via a ping-pong mechanism to transfer CoA
341 groups between carboxylic acids. They are known to act on diverse substrates including oxalate
342 and benzylsuccinate, which is the fumarate addition product of toluene (69). Proteins #29 and #30,
343 encoded on Operon #4, were poorly annotated by most tools but may form a separate enzyme
344 complex (Figures 4b and S7d). Interestingly, MotifFinder (70) detected weak homology in both

345 proteins to domains in PaaZ, a bifunctional ring-hydrolyzing enzyme involved in phenylacetyl-
346 CoA degradation (71) (Table S13). Operon #5 encodes BamQ (Figure 4b) and two proteins – a
347 GNAT family N-acetyltransferase and a transposase-associated protein – that were not detected in
348 the proteome (Figure 3). Homologs of these two proteins are also not encoded by the MAGs of
349 *Pelotomaculum* candidate BPL or *Deltaproteobacteria* AUK045, suggesting they may not be
350 involved in benzene degradation.

351 Proteins #24-26, encoded on Operon #2, were the three most abundant proteins detected in
352 the OR proteomes (Figure 4a; Table 2). AlphaFold 3 predictions (72) suggest that these proteins
353 form a heterotrimeric complex structurally similar to glycolate dehydrogenase in *Escherichia coli*
354 (Figure 5). This enzyme complex is composed of two putative flavoproteins (GlcDE) and an Fe-S
355 protein (GlcF), which together catalyze the oxidation of glycolate to glyoxylate (73). GlcD likely
356 mediates substrate oxidation via its FAD cofactor and transfers electrons to an unknown acceptor
357 through GlcF, a homolog of heterodisulfide reductase subunit HdrB (73, 74). HdrB reduces the
358 coenzyme M-coenzyme B heterodisulfide to regenerate thiol cofactors, and is critical for
359 methanogenesis (75). The electron acceptor(s) for bacterial HdrB homologs is unknown (76). The
360 function of GlcE is also unknown, though knockout mutations to its corresponding gene abolish
361 glycolate oxidation (73). Although the overall trimeric structures of GlcDEF and Proteins #24-26
362 are similar (Figure 5), their homologous subunits share ≤23% amino acid identity and likely
363 catalyze distinct reactions.

364 Most prediction tools agreed that Proteins #24 and #26 are flavoproteins that may oxidize
365 2-hydroxy acids such as glycolate and D-lactate (Figure 4b). DeepFRI (66) predicted that both
366 proteins can bind cyclic organic compounds, and that Protein #24 may also bind FAD (Table S13b)
367 – a prediction supported by AlphaFold 3 models (Figure 5a). Intriguingly, a putative binding

368 pocket is located adjacent to the FAD-binding site in Protein #24 (Figure S8). Phylogenetic
369 analyses placed Protein #25 in an uncharacterized clade of Hdr-like enzymes (Figures S9 and S10).
370 GO term predictions from PROST (77) suggest that it may have 2-hydroxyacid dehydrogenase
371 and ferredoxin hydrogenase activity (Table S14a). Protein #25 contains four cysteine-rich motifs
372 that could bind to electron-transferring Fe-S clusters (Figures 5a), akin to GlcF and HdrB (73, 74).
373 However, two of the binding motifs lack key conserved cysteine residues, which might impair or
374 prevent electron transfer to ferredoxin or other electron acceptors (Figure S11). Overall, functional
375 inferences suggest that the "Magic" gene cluster encodes enzymes that act on aromatic structures,
376 are rich in Fe-S clusters and FAD binding domains and likely convert benzene stepwise to benzoyl-
377 CoA (see discussion).

378 **Predicted function of the "Nanopod" proteins**

379 Protein prediction tools suggest that the "Nanopod" gene cluster encodes a transmembrane
380 complex (Figure 2, Tables S13-S14). Protein #32 most closely resembles a porin, involved in
381 passive substrate transport, whereas Protein #37 (a predicted RND superfamily efflux protein)
382 appears to be involved in active transport. GO term searches in PROST (77) and DeepFRI (66)
383 suggest that Protein #37 and possibly Protein #34 (an uncharacterized cytoplasmic protein) can
384 bind to organic cyclic structures (Tables S14a and S14b). Most tools agreed that Protein #35 is a
385 universal stress factor. Membrane stress can result from partitioning of hydrophobic compounds
386 such as benzene into lipid bilayers (78). Collectively, it is plausible that the "Nanopod" gene
387 cluster facilitates benzene movement across cell membranes, though the direction(s) of movement
388 is not yet known (Figure 2).

389

390 **Phylogenomic placement of the ORM2a and ORM2b_MAG4 genomes**

391 The genomes of ORM2a and ORM2b have undergone multiple reclassifications in response to
392 ongoing taxonomic revisions (4, 39, 40). Our most recent phylogenomic analyses (Figure 6a)
393 placed ORM2a and ORM2b within an uncharacterized clade of *Desulfobacterota*, distinct from
394 previously assigned taxonomic groups including the candidate order Sva0485 and GTDB classes
395 WTBG01 and UBA8473, as well a clade including *Deltaproteobacteria* MAG AUK045. The
396 ORM2a and ORM2b clade – which appears to represent a novel class within the *Desulfobacterota*
397 – also includes three environmental MAGs, one of which (GCA_021160085.1) was recovered
398 from yet another hydrothermal vent system (63). While these MAGs harbor some downstream
399 functional potential to degrade benzoyl-CoA and its intermediates (Table S12), they lack key genes
400 implicated in benzene activation (Table 3).

401 A complementary 16S rRNA gene-based phylogenetic tree (Figure 6b) revealed a
402 consistent clustering pattern, clearly indicating that ORM2a and ORM2b form a monophyletic
403 group of benzene-degrading specialists (4). These organisms – enriched from geographically
404 diverse origins – share ≥ 99% 16S rRNA gene sequence identity and have only been observed to
405 degrade benzene under sulfate-reducing and/or methanogenic conditions (38, 79, 80).
406 Unfortunately, full genomes are lacking for most of these organisms. Taken together, the
407 phylogenomic evidence presented here strongly supports the delineation of ORM2a and ORM2b
408 as members of a novel genus within the *Desulfobacterota*. We propose the name "*Candidatus*
409 Anaerobenzenivorax" (An.ae.ro.ben.ze.ni.i.vor'ax. Gr pref. *an*, not or without; Gr. n. *aer*, air; N.L.
410 neut. n. *benzenum*, benzene, L. adj. *vorax,* devouring, N.L. neut. n. Anaerobenzenivorax, benzene-
411 devouring without air). The proposed taxonomy is currently being registered in SeqCode (81).

412 **DISCUSSION**

413 Many efforts have been made to elucidate mechanisms underlying anaerobic benzene activation.
414 Recent advances in sequencing accuracy, protein detection, and *in silico* protein structure
415 prediction are now enabling deeper insights into these processes. Because benzene is the only
416 known growth-supporting substrate for organisms like "*Ca.* Anaerobenzenivorax" strains ORM2a
417 and ORM2b, conventional differential growth experiments are not feasible. Instead, we must rely
418 on proteomic snapshots during active benzene metabolism. As shown in this study, however,
419 consistent detection of a highly expressed, conserved gene cluster – across multiple years and
420 independent laboratories – strongly implicates this "Magic" gene cluster in initiating benzene
421 metabolism via a novel mechanism that leads to benzoyl-CoA formation. Although biochemical
422 validation is still required, we offer several early postulates regarding the core biochemistry of this
423 elusive pathway.

424 **Proposed mechanism of anaerobic benzene activation catalyzed by ORM2a**

425 The high activation energy required to functionalize benzene in the absence of oxygen is a long-
426 standing challenge. Based on the enzymes identified herein and the example of toluene activation
427 via fumarate addition (82, 83), we explored options for a radical mechanism. We propose that
428 Proteins #24-26 form a novel oxidoreductase complex that catalyzes the initial activation of
429 benzene via flavin-based radical chemistry. The radical flavin is most likely located in Protein #24
430 (Figure 5a) and transfers electrons to Protein #25, similar to GlcF and heterodisulfide reductase
431 subunits (Figures S9–S10), but with mutations in two of its four cysteine-rich motifs that may
432 preclude the use of an external molecule as an electron sink (Figure S11).

433 In our conceptual model (Figure 7), benzene is attacked by a radical generated by flavin-
434 mediated single-electron oxidation of glutaconyl-CoA, a downstream intermediate of anaerobic
435 benzene degradation. Flavin-mediated generation of an acyl-CoA radical was previously proposed

436 for 4-hydroxybutyryl-CoA dehydratase (84). In ORM2a, proteomic evidence supports the idea that
437 glutaconyl-CoA metabolism diverges into two distinct branches: one that proceeds through beta-
438 oxidation reactions to yield acetyl-CoA (via Proteins #18-21), and another – proposed here – that
439 uses glutaconyl-CoA as an acceptor in the initial benzene activation step (Figures 2 and 7). The
440 addition reaction could occur across *para* positions of the aromatic ring, yielding a
441 cyclohexadienyl-type radical intermediate. Resonance stabilization and internal rearrangement
442 would produce 3-phenylglutaryl-CoA, with the radical localized to the α-carbon adjacent to the
443 CoA-linked carbonyl group. This intermediate may then undergo a β-elimination to remove acetyl-
444 CoA and restore aromaticity, yielding cinnamate as the product. Electron balance calculations
445 (Table S15) show that the proposed activation reaction is redox-neutral: any electrons removed
446 from benzene during activation are internally shuttled within the complex – likely via its flavin
447 and Fe-S cofactors – and returned to the activation product.

448

449 **Transformation of the proposed benzene activation intermediate to benzoyl-CoA**

450 If cinnamate is the product of the initial benzene activation reaction, the most plausible next steps
451 include coenzyme A activation followed by modified β-oxidation (85), yielding benzoyl-CoA and
452 releasing an additional mole of acetyl-CoA (Figure 7). Stoichiometry indicates that three moles of
453 acetyl-CoA and three moles of H_2 or equivalent are produced per mole of benzene oxidized (86).
454 To balance this stoichiometry, we propose that Proteins #22 and #27-28 function in appending
455 CoA moieties to cinnamate and a downstream intermediate, enabling β-oxidation to proceed. The
456 third mole of acetyl-CoA is released by a downstream thiolase (Protein #14) during the conversion
457 of 3-hydroxypimelyl-CoA to glutaryl-CoA (Figures 2 and 7). The proposed metabolic pathway is
458 also electron balanced (Table S15).

β-oxidation consists of four consecutive steps: dehydrogenation, hydration, oxidation, and thiolysis. Multiple protein prediction tools implicate Proteins #23 and #29-30 in hydration reactions (Figure 4b and Tables S13-S14). However, the "Magic" gene cluster lacks clear candidates for dehydrogenation (other than Proteins #24-26), oxidation, and thiolysis steps. It remains unclear whether these functions are carried out by enzymes encoded elsewhere in the genome or by uncharacterized proteins within the "Magic" cluster that catalyze analogous reactions.

Closing remarks

There is little doubt that the "Magic" gene cluster in ORM2, *Pelotomaculum* candidate BPL, and specific hydrothermal vent organisms participate in anaerobic benzene degradation, and likely in its elusive, energetically difficult activation step. By contrast, the "Nanopod" gene cluster – present in ORM2a, ORM2b_MAG4, and *Deltaproteobacteria* AUK045 – likely facilitates benzene transport across the cell membrane. Given that the MAGs of most anaerobic benzene degraders surveyed were incomplete, it is possible that homologous "Nanopod" genes were missed. It is telling that the only genomes with both the "Magic" and "Nanopod" clusters belonged to strict anaerobes with apparent specialization in benzene metabolism. The location of both clusters on genomic islands (Figure 1) suggests acquisition via horizontal gene transfer. Now that these unique gene clusters have been identified, we hope the broader research community will help confirm their functions and continue the search for homologous systems in other datasets and enrichment cultures.

MATERIALS AND METHODS

481 Detailed descriptions of all materials, experimental procedures, and computational analyses are
482 provided as Supplementary Materials.

483

484 **FUNDING AND COMPETITNG INTERESTS STATEMENT**

485 This study was funded by Genomic Application Partnership Program (GAPP) grants awarded to
486 E.A.E. (Project IDs OGI-102 and OGI-173), which were supported by Genome Canada, Ontario
487 Genomics, the Government of Ontario, Mitacs Canada, SiREM, Alberta Innovates, Federated Co-
488 operatives Limited, and Imperial Oil. The authors declare no competing financial interest.

489

490 **ACKNOWLEDGEMENTS**

491 This study was made possible by the contributions of many dedicated students and research staff
492 over the past 15 years. Special recognition is given to Dr. Harry Beller (Lawrence Berkeley
493 National Laboratory), Dr. Johann Heider (Philipps University of Marburg), Dr. Christine Orengo
494 (University College London), and Dr. Joel Roca (University College London) for their invaluable
495 guidance on characterizing proteins of unknown function.

496 **REFERENCES**

497 1. Ward DM, Atlas RM, Boehm PD, Calder A. 1980. Microbial biodegradation and chemical
498 evolution of oil from the Amoco spill. Ambio 9:277-283.
499 2. Vogt C, Kleinsteuber S, Richnow HH. 2011. Anaerobic benzene degradation by bacteria.
500 Microb Biotechnol 4:710-724.
501 3. Meckenstock RU, Boll M, Mouttaki H, Koelschbach JS, Cunha Tarouco P, Weyrauch P,
502 Dong X, Himmelberg AM. 2016. Anaerobic degradation of benzene and polycyclic
503 aromatic hydrocarbons. J Mol Microbiol Biotechnol 26:92-118.
504 4. Toth CRA, Luo F, Bawa N, Webb J, Guo S, Dworatzek S, Edwards EA. 2021. Anaerobic
505 benzene biodegradation linked to the growth of highly specific bacterial clades. Environ
506 Sci Technol 55:7970-7980.
507 5. Zhang T, Tremblay PL, Chaurasia AK, Smith JA, Bain TS, Lovley DR. 2013. Anaerobic
508 benzene oxidation via phenol in *Geobacter metallireducens*. Appl Environ Microbiol
509 79:7800-7806.
510 6. Caldwell ME, Suflita JM. 2000. Detection of phenol and benzoate as intermediates of
511 anaerobic benzene biodegradation under different terminal electron-accepting conditions.
512 Environ Sci Technol 34:1216-1220.
513 7. Vogel TM, Grbić-Galić D. 1986. Incorporation of oxygen from water into toluene and
514 benzene during anaerobic fermentative transformation. Appl Environ Microbiol 52:200-
515 202.
516 8. Kunapuli U, Griebler C, Beller HR, Meckenstock RU. 2008. Identification of intermediates
517 formed during anaerobic benzene degradation by an iron-reducing enrichment culture.
518 Environ Microbiol 10:1703-1712.
519 9. Abu Laban N, Selesi D, Jobelius C, Meckenstock RU. 2009. Anaerobic benzene
520 degradation by Gram-positive sulfate-reducing bacteria. FEMS Microbiol Ecol 68:300-
521 311.
522 10. Ulrich AC, Beller HR, Edwards EA. 2005. Metabolites detected during biodegradation of
523 $^{13}C_6$-benzene in nitrate-reducing and methanogenic enrichment cultures. Environ Sci
524 Technol 39:6681-6691.
525 11. Weiner JM, Lovley DR. 1998. Rapid benzene degradation in methanogenic sediments from
526 a petroleum-contaminated aquifer. Appl Environ Microbiol 64:1937-1939.
527 12. Botton S, Parsons JR. 2007. Degradation of BTX by dissimilatory iron-reducing cultures.
528 Biodegradation 18:371-381.
529 13. Phelps CD, Zhang X, Young LY. 2001. Use of stable isotopes to identify benzoate as a
530 metabolite of benzene degradation in a sulphidogenic consortium. Environ Microbiol
531 3:600-603.
532 14. Holmes DE, Risso C, Smith JA, Lovley DR. 2011. Anaerobic oxidation of benzene by the
533 hyperthermophilic archaeon *Ferroglobus placidus*. Appl Environ Microbiol 77:5926-
534 5933.
535 15. Gieg LM, Toth CRA. 2020. Signature metabolite analysis to determine in situ anaerobic
536 hydrocarbon biodegradation, In: Boll M (ed.) *Anaerobic Utilization of Hydrocarbons, Oils,*
537 *and Lipids*. Handbook of Hydrocarbon and Lipid Microbiology. Springer, Cham, pp 361-
538 390. doi:https://doi.org/10.1007/978-3-319-50391-2_19.
539 16. Boll M, Estelmann S, Heider J. 2020. Catabolic pathways and enzymes involved in the
540 anaerobic degradation of monocyclic aromatic compounds . In: Boll M (ed), *Anaerobic*
541 *Utilization of Hydrocarbons, Oils, and Lipids*. Handbook of Hydrocarbon and Lipid

542 Microbiology. Springer, Cham, pp 85-133. doi:https://doi.org/10.1007/978-3-319-50391-
543 2_6.
544 17. Bullows JE, Kanak A, Shedrick L, Kiessling C, Aklujkar M, Kostka J, Chin K. 2024.
545 Anaerobic benzene oxidation in *Geotalea daltonii* involves activation by methylation and
546 is regulated by the transition state regulator AbrB. Environ Microbiol 90:e00856-24.
547 18. Mancini SA, Devine CE, Elsner M, Nandi ME, Ulrich AC, Edwards EA, Sherwood Lollar
548 B. 2008. Isotopic evidence suggests different initial reaction mechanisms for anaerobic
549 benzene biodegradation. Environ Sci Technol 42:8290-8296.
550 19. Mancini SA, Ulrich AC, Lacrampe-Couloume G, Sleep B, Edwards EA, Lollar BS. 2003.
551 Carbon and hydrogen isotopic fractionation during anaerobic biodegradation of benzene.
552 Appl Environ Microbiol 69:191-198.
553 20. Fischer A, Herklotz I, Herrmann S, Thullner M, Weelink SA, Stams AJ, Schlömann M,
554 Richnow H-H, Vogt C. 2008. Combined carbon and hydrogen isotope fractionation
555 investigations for elucidating benzene biodegradation pathways. Environ Sci Technol
556 42:4356-4363.
557 21. Devine CE. 2013. Identification of key organisms, genes and pathways in benzene-
558 degrading methanogenic cultures. PhD dissertation. University of Toronto.
559 22. Taubert M, Vogt C, Wubet T, Kleinsteuber S, Tarkka MT, Harms H, Buscot F, Richnow
560 HH, von Bergen M, Seifert J. 2012. Protein-SIP enables time-resolved analysis of the
561 carbon flux in a sulfate-reducing, benzene-degrading microbial consortium. ISME J
562 6:2291-2301.
563 23. Zhang T, Tremblay PL, Chaurasia AK, Smith JA, Bain TS, Lovley DR. 2014.
564 Identification of genes specifically required for the anaerobic metabolism of benzene in
565 *Geobacter metallireducens*. Front Microbiol 5:245.
566 24. Abu Laban N, Selesi D, Rattei T, Tischler P, Meckenstock RU. 2010. Identification of
567 enzymes involved in anaerobic benzene degradation by a strictly anaerobic iron-reducing
568 enrichment culture. Environ Microbiol 12:2783-2796.
569 25. Luo F, Gitiafroz R, Devine CE, Gong Y, Hug LA, Raskin L, Edwards EA. 2014.
570 Metatranscriptome of an anaerobic benzene-degrading, nitrate-reducing enrichment
571 culture reveals involvement of carboxylation in benzene ring activation. Appl Environ
572 Microbiol 80:4095-4107.
573 26. Atashgahi S, Hornung B, van der Waals MJ, da Rocha UN, Hugenholtz F, Nijsse B,
574 Molenaar D, van Spanning R, Stams AJM, Gerritse J, Smidt H. 2018. A benzene-degrading
575 nitrate-reducing microbial consortium displays aerobic and anaerobic benzene degradation
576 pathways. Sci Rep 8:4490.
577 27. Melkonian C, Fillinger L, Atashgahi S, da Rocha UN, Kuiper E, Olivier B, Braster M,
578 Gottstein W, Helmus R, Parsons JR, Smidt H, van der Waals M, Gerritse J, Brandt BW,
579 Roling WFM, Molenaar D, van Spanning RJM. 2021. High biodiversity in a benzene-
580 degrading nitrate-reducing culture is sustained by a few primary consumers. Commun Biol
581 4:530.
582 28. Xiao JZ, Nesbø CL, Molenda O, Toth CRA, Edwards EA. 2024. Draft genome of an
583 anaerobic nitrate-reducing, benzene-degrading member of the order *Thermincolales*.
584 Microbiol Resour Announc 13:e00295-24.
585 29. Xiao JZ, Nesbø CL, Molenda O, Toth CRA, Edwards EA. 2024. Metagenomic and
586 genomic sequences from a nitrate-reducing benzene-degrading enrichment culture.
587 Microbiol Resour Announc 13:e00294-24.

588 30. Hidalgo KJ, Centurion VB, Lemos LN, Soriano AU, Baessa MP, Richnow HH, Vogt C,
589 Oliveira VM. 2025. Disentangling the microbial genomic traits associated with aromatic
590 hydrocarbon degradation in a jet fuel-contaminated aquifer. Biodegradation 36:7.
591 31. Roberts GW, Leys D. 2022. Structural insights into UbiD reversible decarboxylation. Curr
592 Opin Struct Biol 75:102432.
593 32. Schleinitz KM, Schmeling S, Jehmlich N, von Bergen M, Harms H, Kleinsteuber S, Vogt
594 C, Fuchs G. 2009. Phenol degradation in the strictly anaerobic iron-reducing bacterium
595 *Geobacter metallireducens* GS-15. Appl Environ Microbiol 75:3912-9.
596 33. Keller AH, Schleinitz KM, Starke R, Bertilsson S, Vogt C, Kleinsteuber S. 2015.
597 Metagenome-based metabolic reconstruction reveals the ecophysiological function of
598 *Epsilonproteobacteria* in a hydrocarbon-contaminated sulfidic aquifer. Front Microbiol
599 6:1396.
600 34. Boll M, Estelmann S, Heider J. 2020. Anaerobic degradation of hydrocarbons: mechanisms
601 of hydrocarbon activation in the absence of oxygen. *In* Boll M (ed), *Anaerobic Utilization*
602 *of Hydrocarbons, Oils, and Lipids*. Handbook of Hydrocarbon and Lipid Microbiology.
603 Springer, Cham, pp 3-29. doi:https://doi.org/10.1007/978-3-319-50391-2_2.
604 35. Dong X, Dröge J, von Toerne C, Marozava S, McHardy AC, Meckenstock RU. 2017.
605 Reconstructing metabolic pathways of a member of the genus *Pelotomaculum* suggesting
606 its potential to oxidize benzene to carbon dioxide with direct reduction of sulfate. FEMS
607 Microbiol Ecol 93:fiw254.
608 36. Guo S, Toth CRA, Luo F, Chen X, Xiao J, Edwards EA. 2022. Transient oxygen exposure
609 causes profound and lasting changes to a benzene-degrading methanogenic community.
610 Environ Sci Technol 20:13036-13045.
611 37. Luo F. 2016. Characterization of the microbial community composition and benzene
612 activation mechanisms in anaerobic benzene-degrading enrichment cultures. PhD
613 dissertation. University of Toronto.
614 38. Ulrich AC, Edwards EA. 2003. Physiological and molecular characterization of anaerobic
615 benzene-degrading mixed cultures. Environ Microbiol 5:92-102.
616 39. Luo F, Devine CE, Edwards EA. 2016. Cultivating microbial dark matter in benzene-
617 degrading methanogenic consortia. Environ Microbiol 18:2923-2936.
618 40. Toth CRA, Molenda O, Nesbø C, Luo F, Devine C, Guo S, Chen X, Edwards EA. 2023.
619 Metagenomic and genomic sequences from a methanogenic benzene-degrading
620 consortium. Microbiol Resour Announc 22:e01342-22.
621 41. Chen X, Toth CRA, Guo S, Luo F, Howe J, Nesbo CL, Edwards EA. 2025. Visualization
622 of syntrophic benzene-fermenting *Desulfobacterota* ORM2 in a methanogenic enrichment
623 culture using fluorescence in situ hybridization. Environ Sci Technol 59:591-602.
624 42. Chen X, Molenda O, Brown CT, Toth CRA, Guo S, Luo F, Howe J, Nesbø C, He C,
625 Montabana EA, Cate JHD, Banfield JF, Edwards EA. 2023. "*Candidatus* Nealsonbacteria"
626 are likely biomass recycling ectosymbionts of methanogenic archaea in a stable benzene-
627 degrading enrichment culture. Appl Environ Microbiol 89:e00025-23.
628 43. Kanehisa M, Sato Y, Morishima K. 2016. BlastKOALA and GhostKOALA: KEGG tools
629 for functional characterization of genome and metagenome sequences. J Mol Biol 428:726-
630 731.
631 44. Kanehisa M, Sato Y. 2020. KEGG Mapper for inferring cellular functions from protein
632 sequences. Protein Sci 29:28-35.

633 45. Kanehisa M, Sato Y, Kawashima M. 2022. KEGG mapping tools for uncovering hidden
634 features in biological data. Protein Sci 31:47-53.
635 46. Welte C, Kallnik V, Grapp M, Bender G, Ragsdale S, Deppenmeier U. 2010. Function of
636 Ech hydrogenase in ferredoxin-dependent, membrane-eound electron transport in
637 *Methanosarcina mazei*. J Bacteriol 192:674-678.
638 47. Schink B, Montag D, Keller A, Müller N. 2017. Hydrogen or formate: Alternative key
639 players in methanogenic degradation. Environ Microbiol Rep 9:189-202.
640 48. Yan Z, Ferry JG. 2018. Electron bifurcation and confurcation in methanogenesis and
641 reverse methanogenesis. Front Microbiol 9:1322.
642 49. Stojanowic A, Mander GJ, Duin EC, Hedderich R. 2003. Physiological role of the F420-
643 non-reducing hydrogenase (Mvh) from *Methanothermobacter marburgensis* Arch
644 Microbiol 180:194-203.
645 50. Ma K, Schicho RN, Kelly RM, Adams MW. 1993. Hydrogenase of the hyperthermophile
646 *Pyrococcus furiosus* is an elemental sulfur reductase or sulfhydrogenase: evidence for a
647 sulfur-reducing hydrogenase ancestor. PNAS 90:5341-5344.
648 51. Schut GJ, Boyd ES, Peters JW, Adams MWW. 2013. The modular respiratory complexes
649 involved in hydrogen and sulfur metabolism by heterotrophic hyperthermophilic archaea
650 and their evolutionary implications. FEMS Microbiol Rev 37:182-203.
651 52. Ma K, Weiss R, Adams MWW. 2000. Characterization of hydrogenase II from the
652 hyperthermophilic archaeon *Pyrococcus furiosus* and assessment of its role in sulfur
653 reduction. J Bacteriol 182:1864-1871.
654 53. Lyratzakis A, Meier-Credo J, Langer JD, Tsiotis G. 2023. Insights into the sulfur
655 metabolism of *Chlorobaculum tepidum* by label-free quantitative proteomics. Proteomics
656 23:e2200138.
657 54. Wischgoll S, Heintz D, Peters F, Erxleben A, Sarnighausen E, Reski R, Van Dorsselaer A,
658 Boll M. 2005. Gene clusters involved in anaerobic benzoate degradation of *Geobacter*
659 *metallireducens*. Mol Microbiol 58:1238-1252.
660 55. Boll M, Einsle O, Ermler U, Kroneck PMH, Ullmann GM. 2016. Structure and function of
661 the unusual tungsten enzymes acetylene hydratase and class II benzoyl-coenzyme A
662 reductase. J Mol Microbiol Biotechnol 26:119-137.
663 56. Buckel W, Thauer RK. 2018. Flavin-based electron bifurcation, ferredoxin, flavodoxin,
664 and anaerobic respiration with protons (Ech) or NAD^+ (Rnf) as electron acceptors: a
665 historical review. Front Microbiol 9:401.
666 57. Demmer JK, Chowdhury NP, Selmer T, Ermler U, Buckel W. 2017. The semiquinone
667 swing in the bifurcating electron transferring flavoprotein/butyryl-CoA dehydrogenase
668 complex from *Clostridium difficile*. Nat Commun 8:1577.
669 58. McInerney MJ, Rohlin L, Moutakki H, Kim U, Krupp RS, Rios-Hernandez L, Sieber J,
670 Struchtemeyer CG, Bhattacharyya A, Campbell JW, Gunsalus RP. 2007. The genome of
671 *Syntrophus aciditrophicus*: Life at the thermodynamic limit of microbial growth. PNAS
672 104:7600-7605.
673 59. James KL, Ríos-Hernández LA, Wofford NQ, Moutakki H, Sieber JR, Sheik CS, Nguyen
674 HH, Yang Y, Xie Y, Erde J, Rohlin L, Karr EA, Loo JA, Ogorzalek Loo RR, Hurst GB,
675 Gunsalus RP, Szweda LI, McInerney MJ. 2016. Pyrophosphate-dependent ATP formation
676 from acetyl coenzyme A in *Syntrophus acidotrophicus*, a new twist on ATP formation.
677 mBio 7:e01208-16.

678 60. Hao L, Michaelsen TY, Singleton CM, Dottorini G, Kirkegaard RH, Albertsen M, Nielsen
679 PH, Dueholm MS. 2020. Novel syntrophic bacteria in full-scale anaerobic digesters
680 revealed by genome-centric metatranscriptomics. ISME J 14:906-918.
681 61. Shetty A, Hickey WJ. 2014. Effects of outer membrane vesicle formation, surface-layer
682 production and nanopod development on the metabolism of phenanthrene by *Delftia*
683 *acidovorans* Cs1-4. PLoS One 9:e92143.
684 62. Baumgarten T, Sperling S, Seifert J, von Bergen M, Steiniger F, Wick LY, Heipieper HJ.
685 201. Membrane vesicle formation as a multiple-stress response mechanism enhances
686 *Pseudomonas putida* DOT-T1E cell surface hydrophobicity and biofilm formation. Appl
687 Environ Microbiol 78:6217-6224.
688 63. Speth DR, Yu FB, Connon SA, Lim S, Magyar JS, Peña-Salinas ME, Quake SR, Orphan
689 VJ. 2022. Microbial communities of Auka hydothermal sediments shed light on vent
690 biogeography and the evolutionary history of thermophily. ISME J 16:1750-1764.
691 64. Dombrowski N, Teske AP, Baker BJ. 2018. Expansive microbial metabolic versatility and
692 biodiversity in dynamic Guaymas Basin hydrothermal sediments. Nat Commun 9:4999.
693 65. Campbell KA. 2006. Hydrocarbon seep and hydrothermal vent paleoenvironments and
694 paleontology: Past developments and future research directions. Palaeogeogr
695 Palaeoclimatol Palaeoecol 232:362-407.
696 66. Gligorijević V, Renfrew PD, Kosciolek T, Leman JK, Berenberg D, Vatanen T, Chandler
697 C, Taylor BC, Fisk IM, Vlamakis H, Xavier RJ, Knight R, Cho K, Bonneau R. 2021.
698 Structure-based protein function prediction using graph convolutional networks. Nat
699 Commun 12:3168.
700 67. Dillon SC, Bateman A. 2004. The Hotdog fold: wrapping up a superfamily of thioesterases
701 and dehydratases. BMC Bioinform 5:109.
702 68. Agnihotri G, Liu H. 2003. Enoyl-CoA hydratase: Reaction, mechanism, and inhibition.
703 Bioorg Med Chem 11:9-20.
704 69. Heider J. 2001. A New Family of CoA-Transferases. FEBS Lett 509:345-349.
705 70. Leung WHP, Chi Tam W, Chang BCH, Halgamuge SK. 2003. Effects of search pattern
706 variations in motif discovery algorithm: MotifFinder. IFAC Proc Vol 36:501-506.
707 71. Ferrández A, Miñambres B, García B, Olivera ER, Luengo JM, García JL, Días E. 1998.
708 Catabolism of phenylacetic acid in *Escherichia coli*. J Biol Chem 273:25974-25986.
709 72. Abramson J, Adler J, Dunger J, al. e. 2024. Accurate structure prediction of biomolecular
710 interactions with AlphaFold 3. Nature 630:493-500.
711 73. Pellicer M-T, Badía J, Aguilar J, Baldomà L. 1996. *glc* locus of *Escherichia coli*:
712 Characterization of genes encoding the subunits of glycolate oxidase and the *glc* regulator
713 protein. J Bacteriol 178:2051-2059.
714 74. Wagner T, Koch J, Ermler U, Shima S. 2017. Methanogenic heterodisulfide reductase
715 (HdrABC-MvhAGD) uses two noncubane [4Fe-4S] clusters for reduction. Science
716 357:699-703.
717 75. Yan Z, Wang M, Ferry JG. 2017. A ferredoxin and $F_{420}H_2$-dependent, electron-bifurcating,
718 heterodisulfide reductase with homologs in the domains *Bacteria* and *Archaea*. mBio
719 8:e02285-16.
720 76. Appel L, Willistein M, Dahl C, Ermler U, Boll M. 2021. Functional diversity of prokaryotic
721 HdrA(BC) modules: Role in flavin-based electron bifurcation processes and beyond. BBA
722 Bioenergetics 1862:148379.

723 77. Kilinc M, Jia K, Jernigan RL. 2023. Improved global protein homolog detection with major
724 gains in function identification. PNAS 120:e2211823120.
725 78. Ramos JL, Duque E, Gallegos MT, Godoy P, Ramos-Gonzales MI, Rojas A, Terán W,
726 Segura A. 2002. Mechanisms of solvent tolerance in Gram-negative bacteria. Ann Rev
727 Microbiol 56:473-768.
728 79. Noguchi M, Kurisu F, Kasuga I, Furumai H. 2014. Time-resolved DNA stable isotope
729 probing links *Desulfobacterales*- and *Coriobacteriaceae*-related bacteria to anaerobic
730 degradation of benzene under methanogenic conditions. Microbes Environ 29:191-199.
731 80. Sakai N, Kurisu F, Yagi O, Nakajima F, Yamamoto K. 2009. Identification of putative
732 benzene-degrading bacteria in methanogenic enrichment cultures. J Biosci Bioeng
733 108:501-507.
734 81. Hedlund BP, Chuvochina M, Hugenholtz P, Konstantinidis KT, Murray AE, Palmer M,
735 Parks DH, Probst AJ, Reysenbach A, Rodriguez-R LM, Rossello-Mora R, Sutcliffe IC,
736 Venter SN, Whitman WB. 2022. SeqCode: a nomenclatural code for prokaryotes described
737 from sequence data. Nat Microbiol 7:1702-1708.
738 82. Leuthner B, Heider J. 2000. Anaerobic toluene catabolism of *Thauera aromatica*: the *bbs*
739 operon codes for enzymes of β oxidation of the intermediate benzylsuccinate. J Bacteriol
740 182:272-277.
741 83. Leuthner B, Leutwein C, Schulz H, Horth P, Haehnel W, Schiltz E, Schagger H, Heider J.
742 1998. Biochemical and genetic characterization of benzylsuccinate synthase from *Thauera*
743 *aromatica*: a new glycyl radical enzyme catalysing the first step in anaerobic toluene
744 metabolism. Mol Microbiol 28:615-628.
745 84. Zhang J, Friedrich P, Pierik AJ, Martins BM, Buckel W. 2015. Substrate-induced radical
746 formation in 4-hydroxybutyryl coenzyme A dehydratase from *Clostridium*
747 *aminobutyricum*. Appl Environ Microbiol 81:1071-1084.
748 85. Heider J, Szaleniec M, Martins BM, Seyhan D, Buckel W, Golding BT. 2016. Structure
749 and function of benzylsuccinate synthase and related fumarate-adding glycyl radical
750 enzymes. J Mol Microbiol Biotechnol 26:29-44.
751 86. Corseuil HX, Monier A, Fernandes M, Schneider M, Nunes CC, do Rosario M, Alvarez
752 PJJ. 2011. BTEX plume dynamics following an ethanol blend release: Geochemical
753 footprint and thermodynamic constrains on natural attenuation. Environ Sci Technol
754 45:3422-3429.

755

Table 1. Summary of proteomic experiments performed on the Oil Refinery (OR) enrichment culture. Experiments differed in sample preparation methods, fractionation approaches, and instrumentation. Experiment #1 analyzed one protein sample, while Experiments #2 and #3 included technical replicates (a and b). All samples were subjected to protein extraction by sonication.

Experiment # (Year)	OR subculture used	Benzene degradation rate (mg/L/day)	Culture volume extracted (mL)	Protein purification methodology	Protein separation methodology	Total # of spectra detected	No. of unique predicted proteins identified	No. of unique predicted proteins assigned to designated MAG (s) (% of unique proteins identified)				
								ORM2a	MAG4_ ORM2b	Both ORM2a and MAG4_ ORM2b	Archaeal MAGs	Other Bacterial MAGs
1 (2010)	OR-b1C	0.3	175	Phenol-methanol/ammonium acetate extraction	None	701	119	29 (24%)	14 (12%)	0 (0%)	37 (31%)	2 (2%)
2a (2011)	OR-1b	0.4	175	Phenol-methanol/ammonium acetate extraction	None	228	49	20 (41%)	3 (6%)	1 (2%)	13 (27%)	1 (2%)
2b (2011)	OR-1b	0.4	175	Phenol-methanol/ammonium acetate extraction	None	457	74	27 (36%)	5 (7%)	0 (0%)	21 (28%)	1 (1%)
3a (2018)	OR-b1A	0.7	50	Filter-aided sample preparation	None	47	24	12 (50%)	0 (0%)	1 (4%)	6 (25%)	2 (8%)
3b (2018)	OR-b1A	0.7	50	Filter-aided sample preparation	15% SDS-PAGE	10,914	825	347 (42%)	25 (3%)	44 (5%)	185 (22%)	50 (6%)
						Average		**39%**	**6%**	**2%**	**27%**	**4%**
						Total # Unique Proteins identified		**348**	**33**	**44**	**189**	**50**

Table 2. Summary of high abundance proteins (≥100 matching spectra in Experiment #3b, bolded) and proteins of interest identified with ≥99% confidence. Proteins shown are mapped to one or more MAG reconstructed from OR metagenomes. Proteins are grouped by predicted biochemical functions and ranked by total spectra counts, with ties indicated by (=) symbols. Additional proteins identified are listed in Table S10.

Protein #	Predicted function (NCBI Prokaryotic Genome Annotation Pipeline version 6.3 and best BLASTP hits)	Protein spectra count (Exp. #3b)	Protein abundance rank	Matching MAG(s)	Protein accession #
Benzoyl-CoA degradation to 3-hydroxypimelyl-CoA					
1	Benzoyl-CoA reductase subunit BamB	80	=26	ORM2a and ORM2b	WAC08761.1, ORM2b_MAG4_01005
2	Benzoyl-CoA reductase subunit BamC	15	=154	ORM2a and ORM2b	WAC08762.1, ORM2b_MAG4_01004
3	Benzoyl-CoA reductase subunit BamD	61	=43	ORM2a	WAC08763.1
4	Formate dehydrogenase accessory sulfurtransferase FdhD	6	=287	ORM2a	WAC09094.1
5	Benzoyl-CoA reductase subunit BamE	68	32	ORM2a	WAC09095.1
6	Benzoyl-CoA reductase subunit BamF	18	=135	ORM2a	WAC08764.1
7	Benzoyl-CoA reductase subunit BamG	6	=287	ORM2a	WAC08765.1
8	Benzoyl-CoA reductase subunit BamH	91	23	ORM2a and ORM2b	WAC08766.1, ORM2b_MAG4_04225
9	Benzoyl-CoA reductase subunit BamI	9	=227	ORM2a	WAC08767.1
10	Cyclohexa-1,5-diene-1-carbonyl-CoA hydratase BamR	49	59	ORM2a	WAC08768.1
11	**6-hydroxycyclohex-1-ene-1-carbonyl-CoA dehydrogenase BamQ**	**103**	**20**	**ORM2a**	**WAC08375.1**
12	**6-oxo-cyclohex-1-ene-carbonyl-CoA hydrolase BamA**	**143**	**13**	**ORM2a**	**WAC07399.1**
3-hydroxypimelyl-CoA degradation to acetyl-CoA					
13	**3-hydroxyacyl-CoA dehydrogenase**	**148**	**9**	**ORM2a**	**WAC08792.1**
14	Thiolase family protein	69	31	ORM2a	WAC08791.1
15	Glutaryl-CoA dehydrogenase (non-decarboxylating)	58	=49	ORM2a	WAC06987.1
16	Electron transfer flavoprotein subunit beta/FixA family protein	39	=69	ORM2a	WAC06988.1
17	Electron transfer flavoprotein subunit alpha/FixB family protein	63	=39	ORM2a	WAC06989.1
18	Glutaconyl-CoA decarboxylase subunit alpha	5	=332	ORM2a	WAC08228.1
19	Enoyl-CoA hydratase-related protein	3	=458	ORM2a	WAC07318.1
20	3-hydroxybutyryl-CoA dehydrogenase	9	=227	ORM2a	WAC09154.1
21	Acetyl-CoA C-acetyltransferase	3	=458	ORM2a	WAC07319.1
Highly expressed "Magic" gene cluster putatively involved in benzene activation					
22	AMP-binding protein	74	29	ORM2a and ORM2b	WAC08366.1, ORM2b_MAG4_03434
23	MaoC family dehydratase	57	52	ORM2a	WAC08367.1
24	**FAD-binding oxidoreductase**	**172**	**3**	**ORM2a**	**WAC08368.1**
25	**(Fe-S)-binding protein**	**219**	**2**	**ORM2a**	**WAC08369.1**
26	**FAD-binding oxidoreductase**	**265**	**1**	**ORM2a**	**WAC08370.1**
27	Type III CoA transferase	45	61	ORM2a	WAC08371.1
28	Type III CoA transferase	58	=49	ORM2a and ORM2b	WAC08372.1, ORM2b_MAG4_03441
29	Hypothetical protein	19	=128	ORM2a and ORM2b	WAC08373.1, ORM2b_MAG4_03442
30	MaoC family dehydratase	50	=57	ORM2a	WAC08374.1
31	Transcription termination factor Rho	25	=100	ORM2a	WAC08378.1
Highly expressed "Nanopod" gene cluster putatively involved in benzene transportation across cell membrane					
32	**Hypothetical protein**	**106**	**=18**	**ORM2a**	**WAC07726.1**
33	**DUF1329 domain-containing protein**	**158**	**4**	**ORM2a**	**WAC07727.1, ORM2b_MAG4_01802**
34	Universal stress protein	6	=287	ORM2a	WAC07728.1
35	Hypothetical protein	7	=269	ORM2a	WAC07729.1
36	YCF48-related protein	65	=36	ORM2a	WAC07730.1
37	MMPL family transporter	60	=45	ORM2a	WAC07731.1
Archaeal proteins associated with methanogenesis					
38	**Acetate-CoA synthetase Acs**	**142**	**14**	***Methanothrix* MAG20**	**MAG20_01236**
39	**Coenzyme-B sulfoethylthiotransferase McrB**	**149**	**8**	***Methanothrix* MAG21**	**MAG21_01157**
40	**Coenzyme-B sulfoethylthiotransferase McrB**	**154**	**7**	***Methanothrix* MAG11, MAG20 and MAG28**	**MAG11_02361, MAG20_02648, MAG28_00841**
41	**Acetate-CoA synthetase Acs**	**125**	**16**	***Methanothrix***	**MAG11_00500**
Other proteins					
42	**Chaperonin GroL**	**156**	**=5**	**ORM2a**	**WAC08671.1**
43	**Acetate-CoA synthetase Acs**	**156**	**=5**	**ORM2a**	**WAC06830.1**
44	Porin	53	53	ORM2a	WAC06831.1

Table 3. BLASTP search results for ORM2a "Magic" and "Nanopod" proteins against MAG4_ORM2b and publicly available sequence databases in GenBank and JGI (as of August, 2024). NCBI accession numbers and percent amino acid identities are shown for best BLASTP hits to select MAGs and bins with significant alignments (e-value < 1.0E-05, coverage ≥ 50%) to ORM2a proteins. Protein accession numbers and best BLASTP hit search results for other ORM2a proteins of interest and other reference genomes are provided in Table S12.

Protein cluster #	ORM2a	ORM2b_MAG4	*Pelotomaculum* candidate BPL MAG	Hydrothermal vent MAG AUK045	Hydrothermal vent bin B94_G9	Hydrothermal vent bin B7_G9	Hydrothermal vent MAG, Pescadero Basin	Wastewater MAG, Germany	Freshwater sediment MAG, Oklahoma
	This study	This study	Ga0073689 (IMG)	GCA_021163815.1	GCA_003646755.1	GCA_003646785.1	GCA_021160085.1	GCA_017996625.1	GCA_016935335.1
Highly expressed "Magic" gene cluster putatively involved in benzene activation									
23	**100%**	**96%**	**76%**	**72%**	**71%**	**71%**	25%	24%	27%
24	**100%**	**95%**	**76%**	**73%**		**73%**			
25	**100%**	**97%**	**74%**	**76%**				22%	
26	**100%**	**97%**	**70%**	**69%**	**66%**	28%	26%	29%	29%
27	**100%**	**94%**	**66%**	**62%**	**61%**			24%	
28	**100%**	**98%**	**74%**	**65%**	**72%**	**69%**			28%
29	**100%**	**98%**	**69%**	**60%**	**62%**	**62%**			29%
30	**100%**	**90%**		**61%**	51%				
31	**100%**	**90%**		**60%**	52%				
11	**100%**	**94%**	**64%**	27%		57%	21%	**62%**	
Highly expressed "Nanopod" gene cluster putatively involved in benzene transportation across cell membrane									
32	**100%**	**84%**		44%		26%	30%	23%	29%
33	**100%**	**94%**		**60%**	32%	36%	31%	23%	33%
34	**100%**	**82%**	32%	43%	31%	29%		28%	31%
35	**100%**	**92%**							
36	**100%**	**88%**		46%			31%	30%	30%
37	**100%**	**92%**		51%	33%		35%	36%	34%

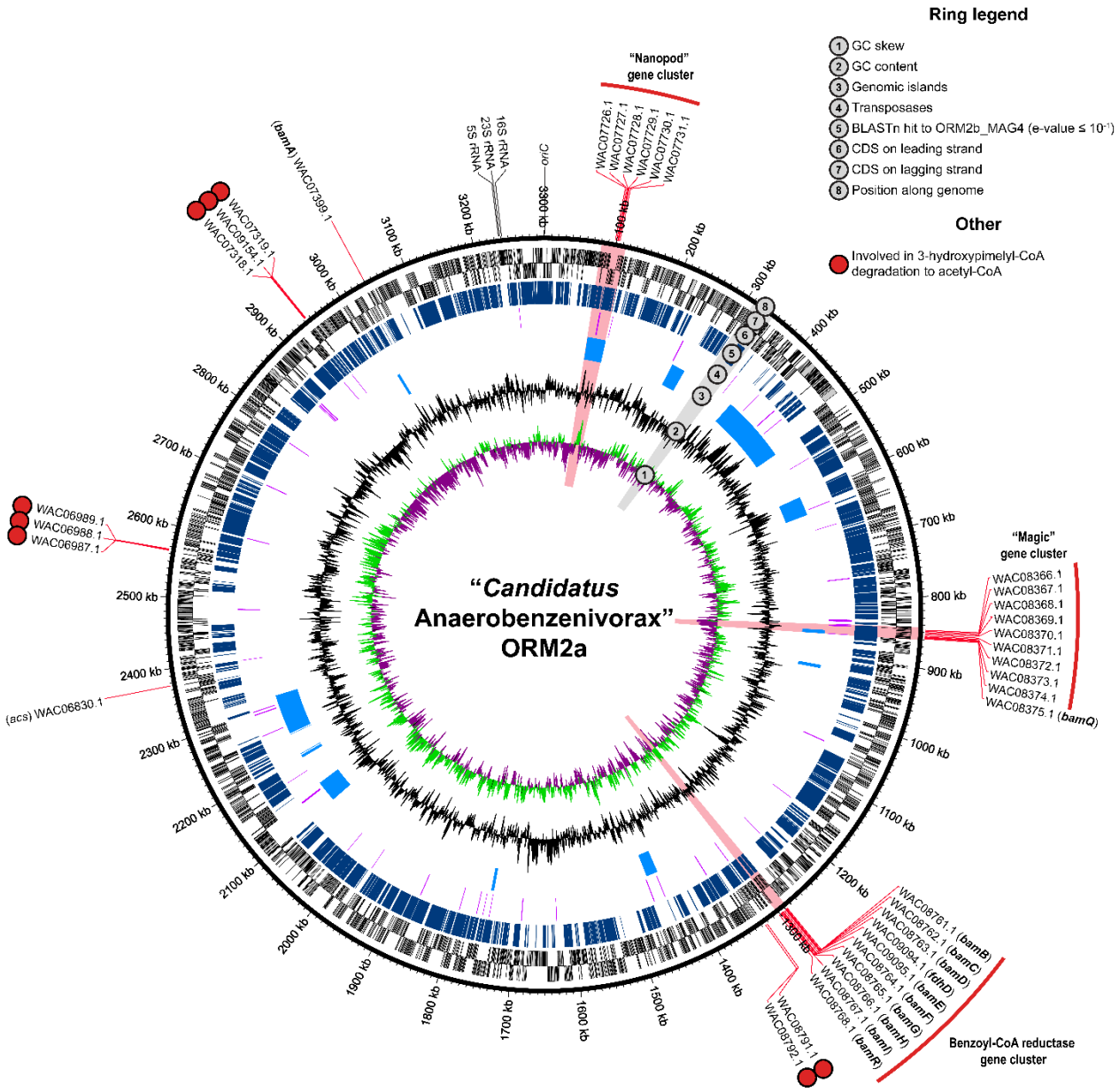
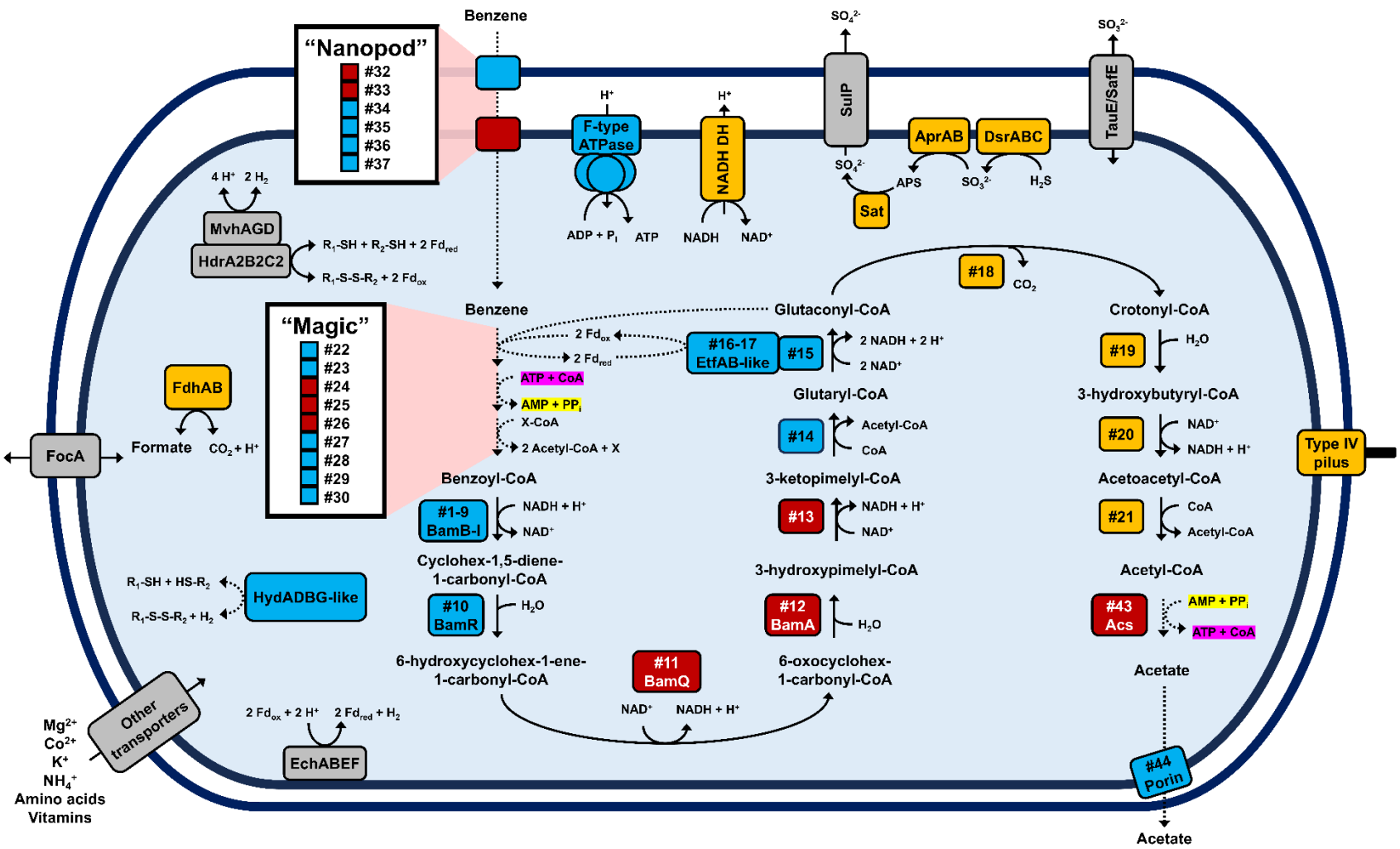


Figure 1. The complete ORM2a genome. Genomic neighborhoods and genes of interest, whose products were recovered with high certainty (≥99% confidence) from the OR proteomes, are highlighted in red and summarized in Table 2. Genes involved in the degradation of 3-hydroxypimelyl-CoA to acetyl-CoA are marked with red circles. Protein accession numbers are provided, and corresponding ORM2a gene locus tags are listed in Table S10. The positions of the origin of replication *(oriC)* and ribosomal RNA genes are also shown in this figure.



Arrow Legend

- Predicted reaction direction
- Proposed reaction and direction
- Unknown reaction direction

Protein Abundance in Exp. #3b

- High (≥100 matching spectra)
- Medium (10-99 spectra)
- Low (1-9 spectra)
- Not detected

Figure 2. Summary of key metabolic pathways identified in the ORM2a proteome. Protein spectral counts from Experiment #3b were used to estimate protein abundances. Additional complete and incomplete metabolic pathways identified in the ORM2a genome are summarized in Table S4. Annotations for numbered proteins are provided in Table 2. Abbreviations (clockwise from top-center): ATPase, ATP synthase; NADH DH; NADH dehydrogenase; SulP, sulfate permease; Sat, sulfate adenylyltransferase; AprAB, adenylyl-sulfate reductase; Dsr, dissimilatory sulfite reductase; TauE/SafE, sulfite exporter; Acs, acetyl-CoA synthetase; EchABEF, energy converting hydrogenase; HydABGD-like, sulfhydrogenase; R-S-S-R, heterodisulfide; FocA, formate-nitrite transporter; FdhAB, formate dehydrogenase; Fd, ferredoxin; HdrA2B2C2, F_{420}-reducing heterodisulfide reductase; Mvh, F_{420}-non-reducing hydrogenase.

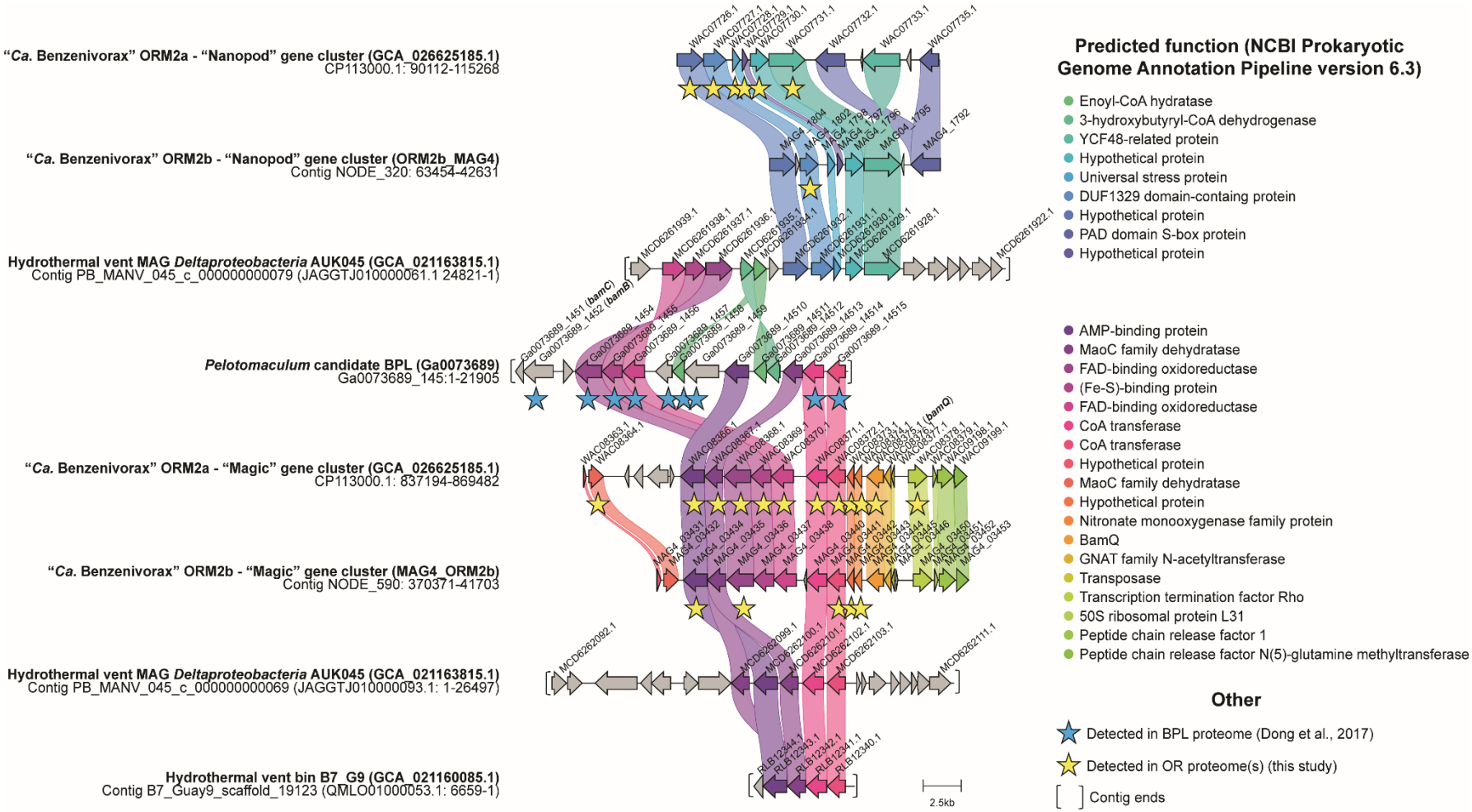
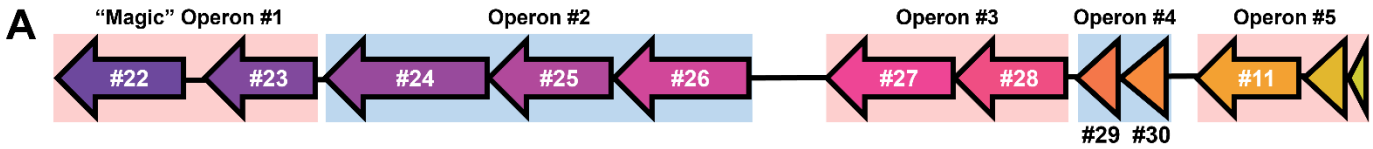


Figure 3. Genomic neighborhoods of all publicly available MAGs in NCBI and IMG producing significant alignments to the "Magic" and "Nanopod" gene clusters in ORM2a and ORM2b_MAG4. Homologous genes (≥40% amino acid identity) are connected and color-coded by function; sequence alignment statistics for relevant proteins are shown in Table 3 and additional homology statistics are provided in Table S12. Corresponding products detected in the proteomes of the OR consortium and the enrichment culture BPL are marked with yellow or blue stars, respectively. The ends of a contig are capped with brackets.



B

Protein #	Closest matching homologous chain description	Z-score
22	AMP-dependent acyl-CoA synthetase and ligase	48.7
23	MaoC dehydratase like domain protein	28.9
24	Probable D-lactate dehydrogenase	29.3
25	Heterodisulfide reductase, subunit A	22.1
26	4-cresol dehydrogenase [hydroxylating] flavoprotein	35.5
27	CaiB/BaiF CoA-transferase family protein	38.5
28	Formyl-CoA:oxalate CoA transferase, subunit YfdW	40.7
29	Outer membrane protein YbgF	11.2
30	(*3R*)-hydroxyacetyl-ACP dehydratase, subunit HadA	17.0
11	6-hydroxycyclohex-1-ene-1-carbonyl-CoA dehydrogenase BamQ	48.8

Figure 4. Structure and functional prediction of proteins encoded by the "Magic" gene cluster. Panel A shows the operon organization of this gene cluster. Panel B summarizes the best Protein Data Bank (PDB) search results for each "Magic" gene product. Z-scores are a measure of protein composite confidence, and Z-scores ≥30 are considered significant. More alignment statistics are provided in Table S13.

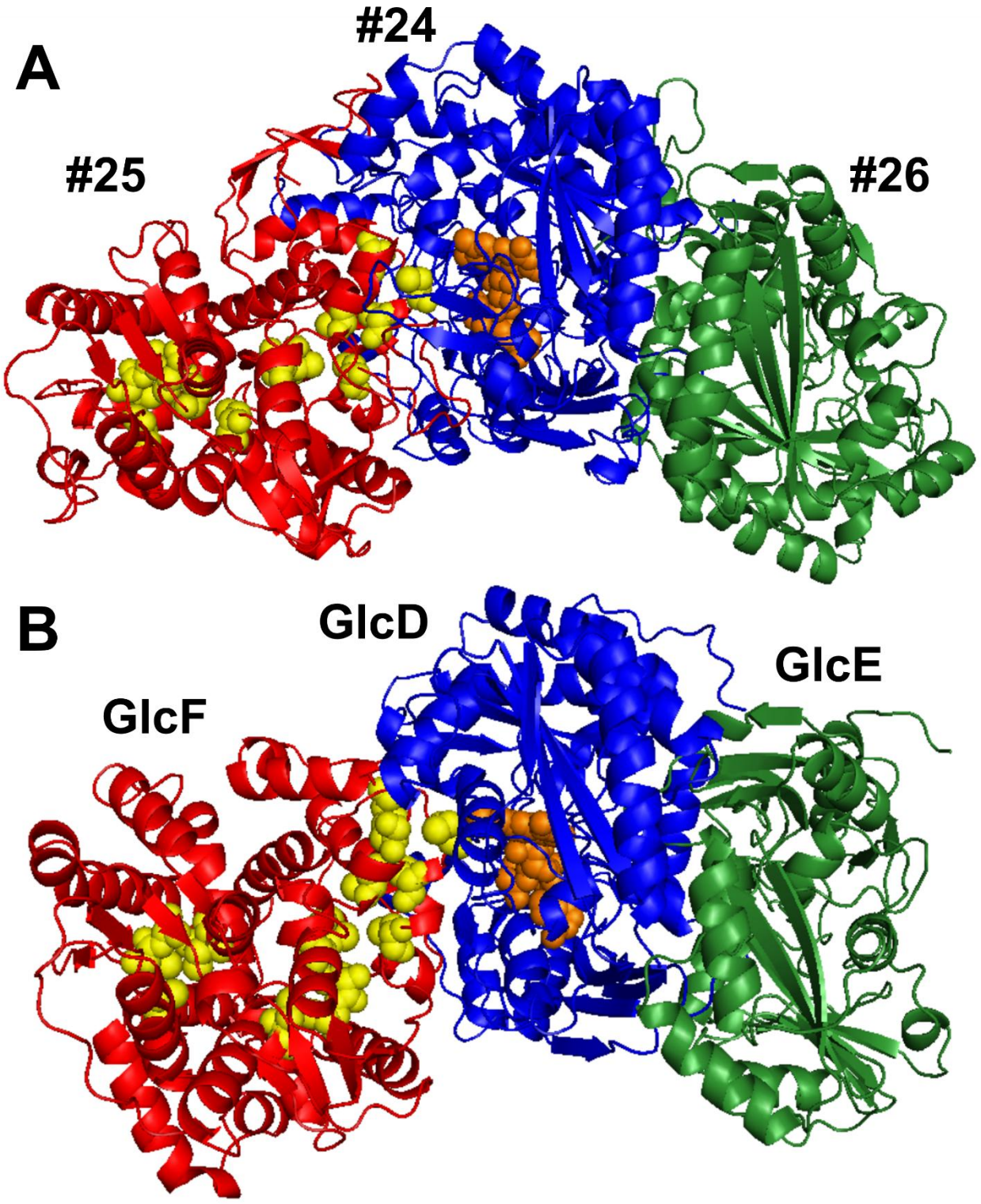


Figure 5. AlphaFold 3 predicted model of A) the Protein #24-26 enzyme complex and B) the glycolate oxidase complex GlcDEF in *Escherichia coli*. Homologous subunits are shown in red, blue, or green. FAD is rendered as orange spheres, and conserved cysteine residues potentially binding [4Fe–4S] clusters are shown as yellow spheres. A putative binding pocket is located near the FAD-binding site in Protein #24, which can be seen in Figure S8a. Structure confidence scores are shown in Figure S8b.

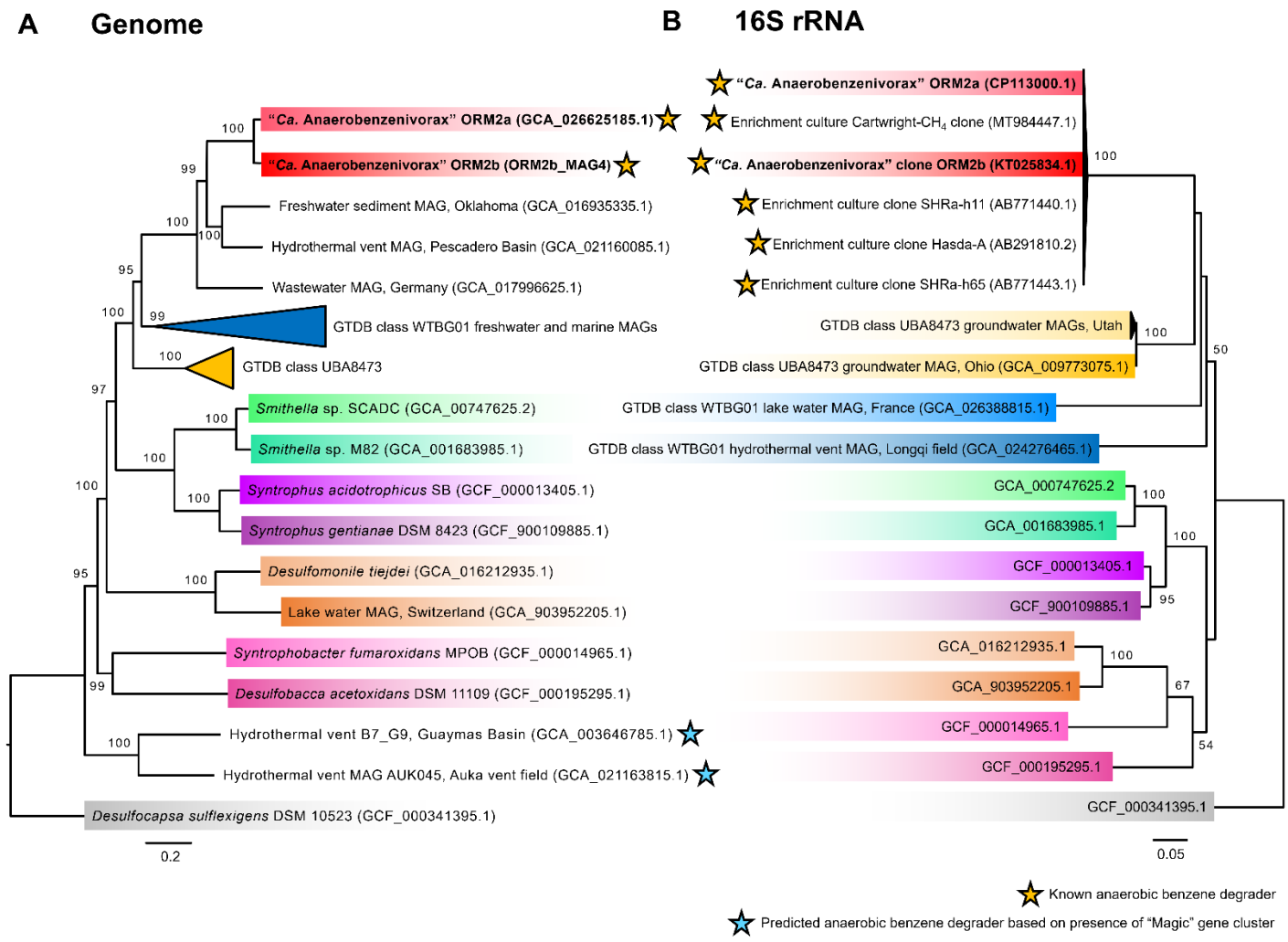


Figure 6. Taxonomic classification of ORM2a and reference *Desulfobacterota* species based on A) genome and B) 16S rRNA gene maximum likelihood phylogenetic trees. Sequences from this study are highlighted in bold. *Desulfocapsa sulflexigens* was used to root both trees. Bootstrap values <60% are not shown.

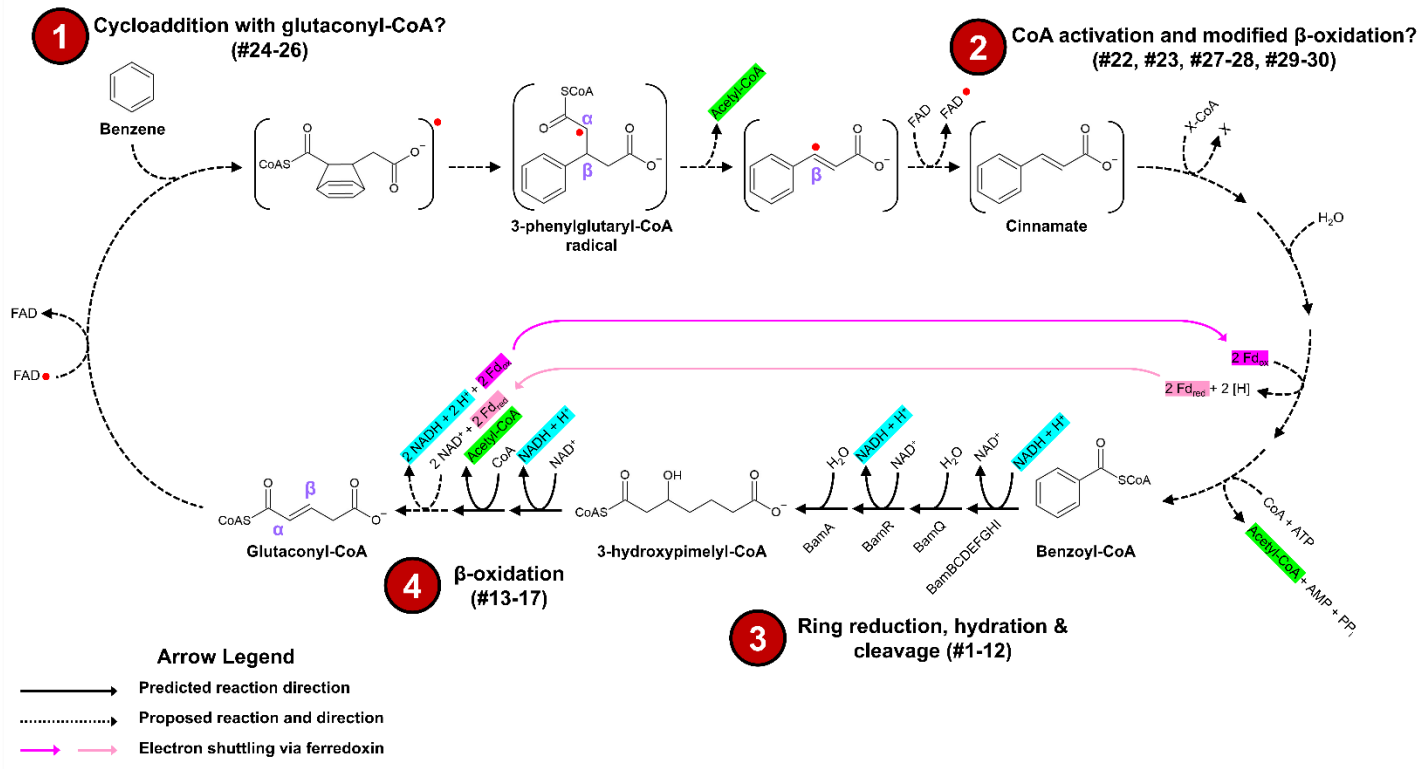


Figure 7. Conceptual model of anaerobic benzene degradation by ORM2a. We propose that a radical flavin-mediated cycloaddition reaction with glutaconyl-CoA is used to activate benzene in ORM2a. Radicals are denoted by red circles. This closed-loop model suggests that glutaconyl-CoA does not undergo further β-oxidation steps yielding acetyl-CoA. Rather, stoichiometric amounts of acetyl-CoA is produced from three separate lyase reactions. Ferredoxin may serve as an electron shuttle between the proposed EtfAb-glutaryl-CoA dehydrogenase complex (Proteins #15-17) and an unknown dehydrogenase during the transformation of cinnamate to benzoyl-CoA. Electron balances for each reaction step are shown in Table S15.