A new gene responsible for *para*-aminobenzoate biosynthesis

Yasuharu Satoh,[1] Masahiro Kuratsu,[2] Daiki Kobayashi,[1] and Tohru Dairi[1]*

[1]Faculty of Engineering, Hokkaido University, N13-W8, Kita-ku, Sapporo 060-8628, JAPAN.

[2]Kyowa Hakko Bio Co. Ltd., 1-6-1, Ohtemachi, Chiyoda-ku, Tokyo 100-8185, JAPAN

*Corresponding author: Tel&FAX: +81 11 706 7815, E-mail: dairi@eng.hokudai.ac.jp

Running title: A new gene responsible for pABA biosynthesis

Keywords: *para*-Aminobenzoate, folate, *Nitrosomonas europaea, Lactobacillus fermentum*, *Chlamydia trachomatis*

ABSTRACT

Folate is an essential cofactor in all living cells for one-carbon transfer reactions. *para*-Aminobenzoate (pABA), a building block of folate, is usually derived from chorismate in the shikimate pathway by reactions of aminodeoxychorismate synthase (PabA & B) and 4-amino-4-deoxychorismate lyase (PabC). We previously suggested that an alternative pathway for pABA biosynthesis would operate in some microorganisms such as *Lactobacillus fermentum* and *Nitrosomonas europaea* since these bacteria showed a prototrophic phenotype to pABA despite the fact that there are no orthologs of *pabA*, *B*, and *C* in their genome databases. In this study, a gene of unknown function, *NE1434*, was obtained from *N. europaea* by shotgun cloning using a pABA-auxotrophic *Escherichia coli* mutant (*ΔpabABC*) as a host. A tracer experiment using [U-$^{13}C_6$]glucose suggested that pABA was *de novo* synthesized in the transformant. An *E. coli ΔpabABCΔaroB* mutant carrying the *NE1434* gene exhibited a prototrophic phenotype to pABA, suggesting that compounds in the shikimate pathway including chorismate were not utilized as substrates by NE1434. Moreover, the *CT610* gene, an ortholog of *NE1434* located in the folate biosynthetic gene cluster in *Chlamydia trachomatis*, also complemented pABA-auxotrophic *E. coli* mutants. Taken together, these results suggest that NE1434 and CT610 participate in pABA biosynthesis.

Folate, composed of a pterin moiety, *para*-aminobenzoate (pABA) and glutamic acid(s), is an essential cofactor for all living cells and plays critical roles in a diverse range of metabolic pathways, mainly in one-carbon transfer reactions such as amino acid interconversions, and purine and pyrimidine biosynthesis. Most bacteria and plants have the pathway for *de novo* synthesis of folate (Figure 1) (1). The first step in forming the pterin moiety is conversion of GTP into 7,8-dihydroneopterin triphosphate by GTP cyclohydrolase I (FolE). After dephosphorylation by phosphatases, the resulting 7,8-dihydroneopterin is converted to 6-hydroxymethyl-7,8-dihydropterin with concomitant release of glycolaldehyde by 7,8-dihydroneopterin aldolase (FolB) and then pyrophosphorylated by 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase (FolK). The resulting compound, 6-hydroxymethyl-7,8-dihydropterin pyrophosphate, is attached to pABA—which is supplied from chorismate by two enzymes, aminodeoxychorismate (ADC) synthase (PabA/B) and ADC lyase (PabC)—by dihydropteroate synthase (FolP) to yield dihydropteroate, which is glutamylated by dihydrofolate synthase (FolC) and then reduced by dihydrofolate reductase (FolA) to give tetrahydrofolate.

We previously showed that *Lactobacillus fermentum* IFO 3956 exhibited a prototrophic phenotype to pABA in a single-omission growth test (2), despite the fact that the strain lacks orthologs of *pabA*, *-B*, and *-C* (3) (Figure 1), suggesting that the strain utilizes an alternative pathway to synthesize pABA. Similarly, *Nitrosomonas europaea* NBRC14298 (ATCC19718) also has a complete set of folate biosynthetic gene orthologs, except for *pabA*, *-B*, and *-C*, in

its genome (KEGG, http://www.genome.jp/kegg/; (4)), despite its chemoautotrophic phenotype. In this paper, we performed shotgun cloning experiments with a pABA-auxotrophic *E. coli* mutant (*ΔpabABC*) as a host and *L. fermentum* IFO 3956 and *N. europaea* NBRC14298 as DNA donors. We successfully obtained a complementary gene, *NE1434*, from *N. europaea.* A tracer experiment using [U-$^{13}C_6$]glucose suggested that NE1434 was involved in *de novo* pABA biosynthesis. We found that *Chlamydia trachomatis* possessed a *NE1434* ortholog (*CT610*) in its folate biosynthetic gene cluster and confirmed that the *CT610* gene also complemented the pABA-auxotrophic *E. coli* mutant (*ΔpabABC*).

MATERIALS AND METHODS

Bacterial strains and cultures

Strains used in this study are summarized in Table 1. *Lactobacillus fermentum* IFO 3956 and *Nitrosomonas europaea* NBRC 14298 (ATCC 19718) were obtained from the Biological Resource Center, National Institute of Technology and Evaluation (NITE), Tokyo, Japan. *Escherichia coli* JM109 (Nippon Gene Co., Ltd, Tokyo, Japan) was routinely used for plasmid construction. For pABA auxotrophy complementation assays, a *pabA*, *-B*, and *-C* gene deleted mutant derived from *E. coli* BW25113 was used (2).

The media used were LB broth medium (Lennox; Life Technologies Corp., Carlsbad, CA, USA) and M9 minimal medium [M9 minimal salts (Becton, Dickinson and Company, Franklin Lakes, NJ, USA), 0.4 % (w/v) glucose, 5 mM $MgSO_4$, 0.1 mM $CaCl_2$]. For growth

on plates, 1.5% (w/v) agar was added into the media. Ampicillin (Ap), chloramphenicol (Cm), and kanamycin (Km) were added to the media at concentrations of 100, 30, and 25 μg mL^{-1} if necessary.

Plasmids construction

Plasmids used in this study are listed in Table 2. All primers were designed using the genomic DNA sequence. A *pabB* homolog, *NE2150* (4), was amplified by PCR using PrimeSTAR GXL DNA polymerase (Takara Bio Inc., Shiga, Japan), genomic DNA of *N. europaea* as a template, and the primers PB1 and PB2, whose sequences are shown in Table S1. The PCR reaction was carried out as recommended by the manufacturer. The amplified DNA fragment was treated with *Nde*I and *Bam*HI and cloned into the same sites of the pET28 vector (Merck Ltd., Tokyo, Japan) to express NE2150 as an *N*-terminal His-tag fused protein. The plasmid was designated pET-NE2150.

For expression of *C*-terminal His-tag fused NE1434 (4), the DNA fragment was amplified by PCR using the method described above with genomic DNA of *N. europaea* and the primers PB3 and PB4 (Table S1). The amplified DNA fragment was cloned into the *Bam*HI and *Hin*dIII sites of the pUC18 vector (Takara Bio Inc.) to construct pUC-NE1434H.

For *in vivo* expression of the *CT610* (5) and *pqqC* genes (6), DNA fragments were amplified by PCR with genomic DNA of *C. trachomatis* strain UW-3/Cx (ATCC VR-885D) and *P. putida* KT2440 using appropriate sets of primers (PB5 to PB8, Table S1). Each of the

amplified DNA fragments was cloned into the *Bam*HI and *Hin*dIII sites of the pUC18 vector (Takara Bio Inc.) to construct pUC-CT610H and pUC-pqqC, respectively. The DNA sequences of all amplified fragments were analyzed with the BigDye Terminator v3.1 Cycle Sequencing Kit using the ABI PRISM 3130 genetic analyzer (Applied Biosystems Japan Ltd., Tokyo, Japan).

Shotgun cloning of a gene related to pABA biosynthesis from *L. fermentum* and *N. europaea*

Genomic DNA from *L. fermentum* and *N. europaea* was partially digested with *Sau*3AI, and approximately 7-kbp DNA fragments were separated by agarose gel electrophoresis and then purified. The DNA fragments were ligated into the pSTV29 cloning vector (pACYC184 ori, P_{lac}, Cm^r; Takara Bio Inc.) digested with *Bam*HI. pABA-auxotrophic *E. coli ΔpabABC* (2) was transformed with the ligation mixture and plated onto M9 minimal media plates.

pABA auxotrophy complementation assay

For the pABA auxotrophy complementation assay, the *E. coli* mutants *ΔpabABC* (2), *ΔpabABCΔaroB*, *ΔpabABCΔaroC*, and *ΔpabABCΔaroD* were transformed with appropriate plasmids. Each of the transformants (triplicate) grown on LB plates was dispersed in 200 μL of sterilized distilled water and then a 50 μL aliquot was inoculated into 5 mL of M9 medium supplemented with or without pABA. The growth was assessed at 30°C since *N. europea*

prefers a cultivation temperature at 20 to 30°C. Cell growth of the transformants was monitored at one-hour intervals using a TAITEC OD-monitorC&T system (Saitama, Japan) and compared to that of parental strain (Figure 2B).

Construction of *aroB, -C,* and *-D* disrupted mutants of *E. coli ΔpabABC*

To construct the disruptants, the Quick & Easy *E. coli* Gene Deletion Kit (Gene Bridges GmbH, Heidelberg, Germany) was used. Briefly, the genes encoding *aroB* (b3389), *aroC* (b2329), or *aroD* (b1693) (7) were replaced with a kanamycin-resistance gene cassette by homologous recombination and then the cassette was removed from the chromosome by a FLP-recombinase according to the manufacturer's protocol (Figure S3). The sequences of the primers are shown in Table S1.

Analysis of pABA productivity

pABA production was examined under the cultivation condition as previously described (8). An aliquot (60 μL) of overnight culture was inoculated into 3 mL of fresh LB medium and cultured at 30°C for 3 h. Then, IPTG was added at a final concentration of 0.5 mM and cultivation was continued for an additional 3 h. After the cells were harvested and washed with the same amount of M9 medium, the collected cells were suspended in 3 mL of M9 medium including Ap and 0.5 mM IPTG, followed by cultivation at 30°C. Culture samples were collected at appropriate time intervals (24, 48, and 72 h) to analyze pABA productivity

by LC/MS.

Supernatants of the culture (5 μL) were analyzed using an Agilent Technologies 1100 series HPLC system (Agilent Technologies Inc., Santa Clara, CA, USA) equipped with a Bruker Daltonics micrOTOF-HS (Bruker Daltonics Inc., Billerica, MA, USA). The analytical conditions were as follows: InertSustain C18 column (150 mm × 2.1 mm ID, 3 μm; GL Sciences Inc., Tokyo, Japan); column temperature, 35°C; 5% (v/v) acetonitrile solution with 0.1% (v/v) formic acid by isocratic flow; flow rate, 0.2 mL/min. pABA was eluted at approximately 10 min and detected as m/z = 138.056 ± 0.010 and 145.079 ± 0.010, which correspond to the $[M+H]^+$ ions for non-labeled and ^{13}C-labeled pABA.

RESULTS

Shotgun cloning experiment with the *E. coli ΔpabABC* mutant.

We previously showed that *L. fermentum* IFO 3956 was prototrophic to pABA and that LAF1336, an ortholog of FolP in *L. fermentum* IFO 3956, utilized 6-hydroxymethyl-7,8-dihydropterin pyrophosphate and pABA to form dihydropteroate (2), despite the fact that the strain has no orthologs of PabA, -B, and -C (3). These facts strongly suggested that the strain biosynthesizes pABA via a new enzyme/pathway. We therefore used genomic DNA of *L. fermentum* IFO 3956 as donor DNA to obtain a DNA fragment that could complement the pABA-auxotrophic *E. coli* mutant (*ΔpabABC*). Approximately 7-kbp DNA fragments partially digested with *Sau*3AI were used for shotgun cloning and screening

153 was carried out on the basis of complementation of pABA auxotrophy of the *E. coli*
154 *ΔpabABC* mutant on M9 minimal plates without supplementation of pABA. However, no
155 transformants were obtained even though we performed several trials. A reason for the failure
156 in complementation using *L. fermentum* genome is discussed later.

157 We then used *N. europaea* NBRC 14298 as the DNA donor. The strain also has a
158 complete set of orthologs of the folate biosynthetic genes, except for pABA-biosynthesis
159 related genes, in its genome (GTP cyclohydrolase I, NE1163; dihydroneopterin aldolase,
160 NE0223; 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase, NE0070; dihydropteroate
161 synthase, NE0529; dihydrofolate synthase, NE0696; dihydrofolate reductase, NE0567) (4).
162 As for pABA biosynthetic genes, only one ortholog, NE2150, which was annotated as TrpE,
163 was hit by a BLAST search (KEGG) when *E. coli* PabB was used as a query sequence.
164 Microorganisms generally possess two orthologs of PabB; one is PabB, which is responsible
165 for pABA biosynthesis, and the other is an anthranilate synthase subunit (TrpE), which is
166 responsible for tryptophan biosynthesis (9). To examine whether *NE2150* participates in
167 pABA biosynthesis, we prepared recombinant NE2150 and used it for *in vitro* assays (Figure
168 S1A). When chorismate and NH_4Cl were used as substrates, the formation of anthranilate was
169 confirmed (Figure S1B); in contrast, no product was observed when NE2150 was omitted,
170 demonstrating that *NE2150* actually encodes anthranilate synthase (TrpE) and that *N.*
171 *europaea* synthesizes pABA via a new enzyme/pathway.

172 To obtain a gene related to pABA biosynthesis from *N. europaea*, we performed

173 shotgun cloning using the method described above. We obtained dozens of colonies on M9
174 minimal plates without pABA. The DNA sequences of plasmids randomly prepared from the
175 transformants revealed that they had a common region containing four genes, *NE1431* to
176 *NE1434* (Figure 2A). To determine the essential gene(s) for complementation, the obtained
177 plasmids were digested with *Hin*dIII and then self-ligated (Figure 2A). The constructed
178 plasmid including *NE1434* gene again complemented the *E. coli ΔpabABC* mutant, showing
179 that NE1434 was essential.

180 Cell growth of *E. coli* BW25113 (parental strain) and *E. coli ΔpabABC* mutant
181 expressing *NE1434* gene in M9 minimum medium was compared. As shown in Figure 2B,
182 the transformant showed the almost the same growth curve as that of the parental strain,
183 showing that NE1434 gene would be involved in pABA biosynthesis.

184 NE1434 was previously annotated as pyrroloquinoline-quinone (PQQ) synthase (*pqqC*).
185 However, the low similarity between NE1434 and PqqC (24% identity and 40% similarity for
186 *Klebsiella pneumoniae* KPN01811, 25% identity and 39% similarity for *Pseudomonas putida*
187 PP0378 (10)) and the absence of other genes related to PQQ biosynthesis in the flanking
188 region suggested that NE1434 was a new enzyme responsible for pABA biosynthesis.

189

190 **Tracer experiment using [U-$^{13}C_6$]glucose.**

191 We next performed a tracer experiment using [U-$^{13}C_6$]glucose to estimate the precursor
192 compound of pABA. Since milligram quantities of labeled pABA are required for ^{13}C-NMR

193 analysis and the concentration of coenzymes and vitamins including folate in cells is
194 generally quite low, we first examined pABA productivity in the *E. coli ΔpabABC* mutant
195 harboring pUC-NE1434H. After the transformant was cultured in M9 medium, the culture
196 broth and cell extracts were analyzed by HPLC and LC/MS as described in the Material and
197 Methods. pABA was detected only with LC/MS ($[M+H]^+$ = 138.056, Figure 3, left column)
198 and the productivity was too low to purify labeled pABA. Then, we examined whether pABA
199 was *de novo* synthesized in the transformant using [U-$^{13}C_6$]glucose as the sole carbon source.
200 As shown in Figure 3 (right column), one peak corresponding to ^{13}C-labeled pABA ($[M+H]^+$
201 = 145.079) was specifically detected, confirming that the *NE1434* gene product participates in
202 *de novo* pABA production.
203
204 **Elucidation of the substrate of NE1434.**
205 We tried to elucidate the substrate of NE1434 by several methods. In the known
206 pathway, pABA is derived from chorismate in two steps catalyzed by ADC synthase (PabAB)
207 and ADC lyase (PabC). We therefore prepared recombinant NE1434 fused with a His-tag at
208 the C-terminus (Figure S2) and incubated it with chorismate. However, no product was
209 formed under the various conditions tested such as addition of co-factors and metals and
210 different pH conditions.
211 Besides chorismate, 3-dehydroquinate, an intermediate in the shikimate pathway (11), is
212 also used as a precursor for pABA biosynthesis. We therefore examined whether the *NE1434*

gene product utilized a compound in the shikimate pathway by a genetic method. We constructed *aroB* (3-dehydroquinate synthase), *aroC* (chorismate synthase), and *aroD* (3-dehydroquinate dehydratase I)-knockout mutants from the *E. coli ΔpabABC* mutant. These strains required aromatic amino acids, 4-hydroxy benzoate, and pABA for their growth. However, all of the mutants harboring the *NE1434* gene were able to grow in M9 medium without pABA (Figure S3), suggesting that none of the intermediates in the shikimate pathway, except for 3-deoxy-D-arabinoheptulosonate-7-phosphate, could be the precursor for NE1434-mediated pABA biosynthesis. To exclude the possibility that NE1434 utilized aromatic amino acids and 4-hydroxy benzoate, which were added into the medium, as precursors for pABA biosynthesis, we used these compounds as substrates for NE1434. However, no products were formed. Taking these results together, we concluded that a new enzyme/pathway for pABA biosynthesis is operating in *N. europaea.*

We also tried to isolate a pABA auxotrophic mutant from *E. coli ΔpabABC* expressing the *NE1434* gene because a gene complimenting the mutant is thought to participate in biosynthesis of the substrate of NE1434. Therefore, the *NE1434* gene was integrated into the *tnaA* locus (tryptophanase gene) in *E. coli ΔpabABC* (Figure S4). The strain was treated with *N*-methyl-*N*'-nitro-*N*-nitrosoguanidine and pABA auxotroph were screened. We screened approximately 30,000 colonies including penicillin-concentrated auxotrophs. However, no mutants exhibiting a pABA auxotrophic phenotype were obtained, suggesting that an essential gene might be responsible for biosynthesis of the substrate of NE1434.

233

234 **Distribution of NE1434 orthologs.**

235 The *NE1434* gene was previously annotated as *pqqC*, which is a member of the
236 TenA/Thi-4/PqqC family of Pfam (http://pfam.sanger.ac.uk/). We investigated the
237 distribution of NE1434 orthologs using BLASTP and MEGA5 (12). Orthologs were collected
238 from the genome databases of prokaryotes that have no gene clusters for PQQ biosynthesis
239 (10), and therefore could be responsible for pABA biosynthesis like NE1434, using the
240 BLASTP program with NE1434 as a query sequence and a cutoff e-value of 10^{-24}. The
241 orthologs were aligned together with PqqC from *K. pneumoniae* and *P. putida*, both of which
242 were confirmed to participate in PQQ biosynthesis, to construct a phylogenetic tree (Figure 4).
243 The orthologs were distributed among two groups of prokaryotes. One was
244 ammonia-oxidizing microorganisms such as *Nitrosomonas*, *Nitrosopumilus*, *Thaumarchaeota,*
245 and the nitrogen-fixing *Trichodesmium erythraeum*. The other included obligate intracellular
246 bacteria such as *Chlamydia*, *Chlamydophila*, *Wolbachia*, *Rickettsia*, and *Neorickettsia.*

247 In the case of *Chlamydia*, *Chlamydia trachomatis* has a folate biosynthetic gene cluster
248 including *folA* (*CT614*), *-P* (*CT613*), and *-X* (*CT612, folB* ortholog) genes (5). However,
249 *pabA*, *-B*, and *-C* genes are missing from the genome as in *N. europaea.* However, the *CT610*
250 gene, an ortholog of *NE1434* (40% identity and 60% similarity), is located downstream of the
251 *folAPX* cluster in the same direction. To confirm whether *CT610* participates in pABA
252 biosynthesis, the gene was cloned into a cloning vector (pUC-CT610H) and a pABA

auxotrophy complementation assay with the *E. coli ΔpabABC* mutant was carried out. As shown in Figure 5, the *CT610* gene complemented pABA auxotrophy in the same manner as *NE1434*, suggesting that *CT610* is also involved in pABA biosynthesis in *C. trachomatis*. In contrast, the *PP0378* gene (pUC-pqqC), which is an ortholog of the *pqqC* gene located in the PQQ biosynthetic gene cluster in *P. putida* KT2440 (6) and has low similarity to NE1434 (25% identity and 39% similarity), was unable to complement pABA-auxotrophic *E. coli ΔpabABC* (Figure 5) suggesting that *PP0378* is involved in PQQ biosynthesis and that only orthologs with a high e-value to NE1434 are responsible for pABA biosynthesis.

DISCUSSION

We previously showed that some bacteria lacked orthologs of *pabA*, *-B*, and *-C,* which are essential for pABA biosynthesis, though they exhibited a prototrophic phenotype to pABA (2). In this study, we showed that *NE1434*, which is annotated as *pqqC* but has low similarity, was responsible for *de novo* pABA biosynthesis in *N. europaea.* The *CT610* gene in *C. trachomatis*, which is an ortholog of *NE1434* located in the folate biosynthetic gene cluster, also complemented the pABA-auxotrophy of the *E. coli ΔpabABC* mutant (Figure S5A). In contrast, we could not obtain any genes that complemented the pABA auxotrophic phenotype when the genome of *L. fermentum* IFO 3956 was used as a DNA source. Considering that it has no orthologs of the *NE1434* gene, two or more genes located at separated loci might be responsible for biosynthesis of pABA in *L. fermentum* IFO 3956.

Otherwise, the strain may biosynthesizes pABA in a totally different manner using a different compound as the substrate.

We tried to identify the substrate of NE1434 by several approaches. Since intermediate compounds in the shikimate pathway such as chorismate and 3-dehydroquinate are used as substrates for pABA biosynthesis (11), we first examined this possibility. However, recombinant NE1434 did not react with chorismate and none of the *aroB*, -*C*, and -*D* knockout mutants of *E. coli ΔpabABC* harboring the *NE1434* gene required pABA for their growth, suggesting that intermediate compounds in shikimate pathway are not used as the substrate of NE1434.

We were not able to estimate the substrate of NE1434 by the tracer experiment, *in vitro* analysis with recombinants, or genetic experiments. Since NE1434 has significant similarity to PqqC, whose substrate and product are already known, studies of the detailed reaction mechanism of PqqC would be a shortcut to identify the substrate of NE1434. Otherwise, another approach such as analysis of the crystal structure of NE1434 might give us a clue. The following aspects might also important in identifying it, though we cannot conceive an experimental plan utilizing them. One is the fact that orthologs of NE1434 are specifically distributed in the genomes of ammonia-oxidizing bacteria and archaea (Figure 4), which possess no *pabA*, -*B*, and -*C* homologs in their genomes. The other is that CT610 was reported to function as a modulator of host cell apoptosis via binding to the death domains of tumor necrosis factor family receptors (13).

The folate biosynthetic pathway has been a target for the development of antibacterial and antiparasitic drugs (1). Sulfonamide and diaminopyrimidine antibiotics are successful examples. Phylogenetic analysis (Figure 4) showed that obligate intracellular pathogenic bacteria such as *C. trachomatis* and *Neorickettsia sennetsu* possess NE1434 orthologs. Since an NE1434 ortholog is absent in mammalian cells, NE1434 orthologs would be an attractive target for development of antibiotics with high specificity to human pathogens.

ACKNOWLEDGEMENTS

We thank Prof. Tatsufumi Okino in Hokkaido University for his helpful suggestions. This work was supported by the Grants-in-Aid for Scientific Research (23108101 and 24651235 to T. Dairi and 2471023902 to Y. Satoh) from the Japan Society for the Promotion of Science (JSPS) and "Strategic Molecular and Materials Chemistry through Innovative Coupling Reactions" of Hokkaido University.

Author contributions

Y.S., M.K., and T.D. designed the experiments. Y.S. M.K., and D.K. performed the experiments. Y.S. and T.D. wrote the manuscript.

REFERENCES

1. **Bermingham, A. and Derrick, J. P.:** The folic acid biosynthesis pathway in

bacteria: evaluation of potential for antibacterial drug discovery, BioEssays, **24**, 637−648 (2002).

2. **Kuratsu, M., Hamano, Y., and Dairi, T.:** Analysis of the *Lactobacillus* metabolic pathway, Appl. Environ. Microbiol., **76**, 7299−7301 (2010).

3. **Morita, H., Toh, H., Fukuda, S., Horikawa, H., Oshima, K., Suzuki, T., Murakami, M., Hisamatsu, S., Kato, Y., Takizawa, T., Fukuoka, H., Yoshimura, T., Itoh, K., O'Sullivan, D. J., McKay, L. L., Ohno, H., Kikuchi, J., Masaoka, T., and Hattori, M.:** Comparative genome analysis of *Lactobacillus reuteri* and *Lactobacillus fermentum* reveal a genomic island for reuterin and cobalamin production, DNA Res., **15**, 151−161 (2008).

4. **Chain, P., Lamerdin, J., Larimer, F., Regala, W., Lao, V., Land, M., Hauser, L., Hooper, A., Klotz, M., Norton, J., Sayavedra-Soto, L., Arciero, D., Hommes, N., Whittaker, M., and Arp, D.:** Complete genome sequence of the ammonia-oxidizing bacterium and obligate chemolithoautotroph *Nitrosomonas europaea*, J. Bacteriol., **185**, 2759−2773 (2003).

5. **Stephens, R. S., Kalman, S., Lammel, C., Fan, J., Marathe, R., Aravind, L., Mitchell, W., Olinger, L., Tatusov, R. L., Zhao, Q., Koonin, E. V., and Davis, R. W.:** Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*, Science, **282**, 754−759 (1998).

6. **Nelson, K. E., Weinel, C., Paulsen, I. T., Dodson, R. J., Hilbert, H.,**

Martins dos Santos, V. A., Fouts, D. E., Gill, S. R., Pop, M., Holmes, M., Brinkac, L., Beanan, M., DeBoy, R. T., Daugherty, S., Kolonay, J., Madupu, R., Nelson, W., White, O., Peterson, J., Khouri, H., Hance, I., Chris Lee, P., Holtzapple, E., Scanlan, D., Tran, K., Moazzez, A., Utterback, T., Rizzo, M., Lee, K., Kosack, D., Moestl, D., Wedler, H., Lauber, J., Stjepandic, D., Hoheisel, J., Straetz, M., Heim, S., Kiewitz, C., Eisen, J. A., Timmis, K. N., Düsterhöft, A., Tümmler, B., and Fraser, C. M.: Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440, Environ. Microbiol., **4**, 799−808 (2002).

7. **Blattner, F. R., Plunkett III, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., Shao, Y.:** The complete genome sequence of *Escherichia coli* K-12, *Science*, **277**, 1453−1462 (1997).

8. **Satoh, Y., Tajima, K., Munekata, M., Keasling, J. D., Lee, T. S.:** Engineering of a tyrosol-producing pathway, utilizing simple sugar and the central metabolic tyrosine, in *Escherichia coli*, *J. Agric. Food Chem.*, **60**, 979−984 (2012).

9. **Goncharoff, P. and Nichols, B. P.:** Nucleotide sequence of *Escherichia coli* *pabB* indicates a common evolutionary origin of *p*-aminobenzoate synthetase and anthranilate synthetase, J. Bacteriol., **159**, 57−62 (1984).

10. **Shen, Y. Q., Bonnot, F., Imsand, E. M., RoseFigura, J. M., Sjölander, K., and Klinman, J. P.:** Distribution and properties of the genes encoding the biosynthesis of the bacterial cofactor, pyrroloquinoline quinone, Biochemistry, **51**, 2265−2275 (2012).

11. **Porat, I., Sieprawska-Lupa, M., Teng, Q., Bohanon, F. J., White, R. H., and Whitman, W. B.:** Biochemical and genetic characterization of an early step in a novel pathway for the biosynthesis of aromatic amino acids and *p*-aminobenzoic acid in the archaeon *Methanococcus maripaludis*, Mol. Microbiol., **62**, 1117−1131 (2006).

12. **Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S.:** MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, Mol. Biol. Evol., **28**, 2731−2739 (2011).

13. **Schwarzenbacher, R., Stenner-Liewen, F., Liewen, H., Robinson, H., Yuan, H., Bossy-Wetzel, E., Reed, J. C., and Liddington, R. C.:** Structure of the *Chlamydia* protein CADD reveals a redox enzyme that modulates host cell apoptosis, *J. Biol. Chem.*, **279**, 29320−29324 (2004).

FIGURE LEGENDS

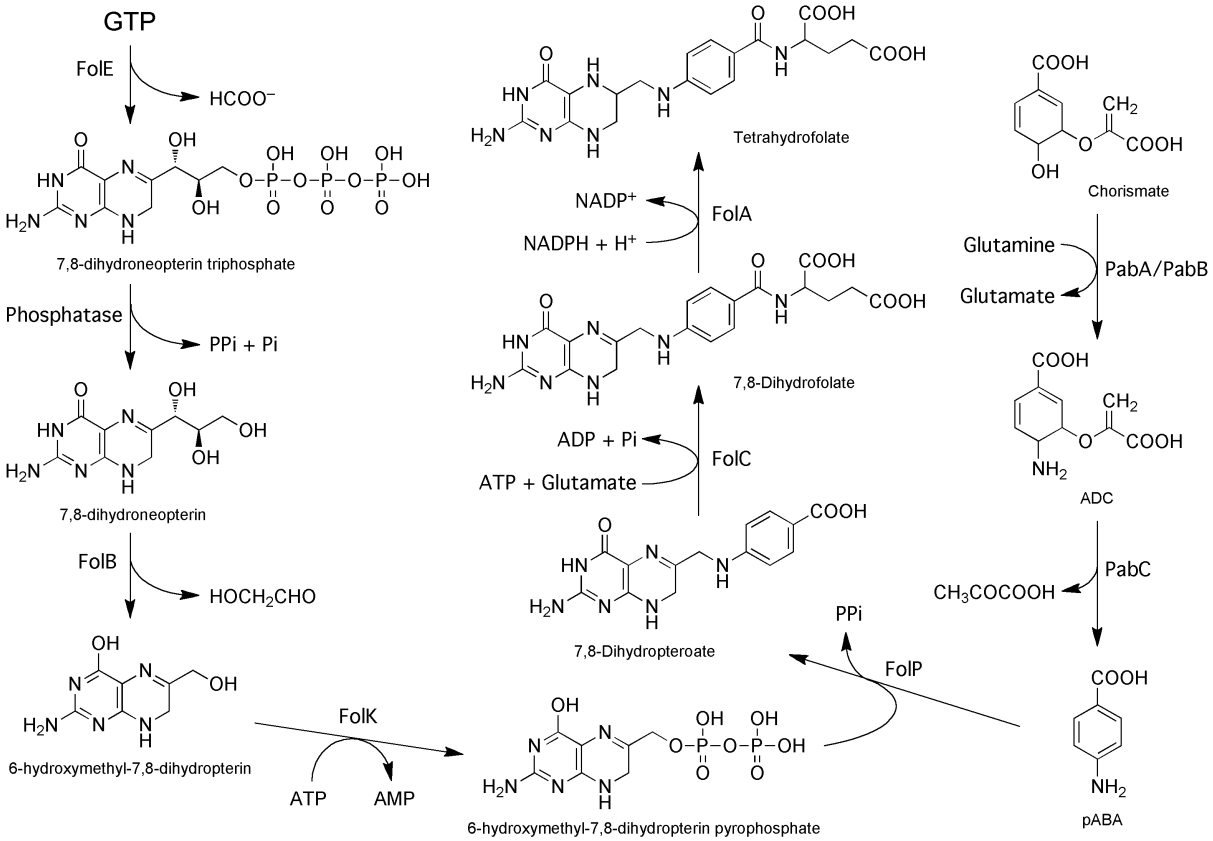
Figure 1 Folate biosynthetic pathway.

373 FolE, GTP cyclohydrolase I (LAF1339, NE1163); FolB, 7,8-dihydroneopterin aldolase,
374 (LAF1341, NE0223); FolK, 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase
375 (LAF1340, NE0070); FolP, dihydropteroate synthase (LAF1336, NE0529); FolC,
376 dihydrofolate synthase (LAF1338, NE0696); FolA, dihydrofolate reductase (LAF0888,
377 NE0567).
378
379 **Figure 2** Shotgun cloning experiment with the *E. coli ΔpabABC* mutant and genomic DNA
380 of *N. europaea* NBRC 14298.
381 **(A)** DNA fragments obtained by shotgun cloning are shown schematically. The predicted
382 function of each gene is as follows: NE1429, anhydro-N-acetylmuramic acid kinase; NE1430,
383 M23/M37 family peptidase; NE1431, tyrosyl-tRNA synthetase; NE1432, transcriptional
384 regulator; NE1433, serine hydroxymethyltransferase; NE1434, pyrroloquinoline-quinone
385 synthase; NE1435, transcription regulator. Parentheses indicate restriction sites in pSTV29
386 cloning vector.
387 **(B)** Growth curves of *E. coli* BW25113 (parent strain) harboring pUC19 (blue dashed line), *E.*
388 *coli* BW25113 harboring pUC-NE1434H (blue line), *E. coli ΔpabABC* mutant harboring
389 pUC-NE1434H (red line), *E. coli ΔpabABC* mutant harboring pUC19 (red dashed line), and
390 *E. coli ΔpabABC* mutant supplemented pABA in M9 minimum medium are shown.
391
392 **Figure 3** Tracer experiment using [U-$^{13}C_6$]glucose as the sole carbon source.
393 The *E. coli ΔpabABC* mutant expressing the *NE1434* gene was cultivated with unlabeled (left

394 column) and [U-$^{13}C_6$]glucose (right column). After cultivation for 72 h, the culture broth was
395 analyzed by LC/MS. The upper two chromatograms were selected for m/z = 138.056 ± 0.010,
396 which corresponds to the $[M+H]^+$ ion for pABA with all unlabeled carbon. The lower two
397 chromatograms were selected for m/z = 145.079 ± 0.010 corresponding to the $[M+H]^+$ ion for
398 ^{13}C-labeled pABA. The upper line (blue) and lower line (red) in all four chromatograms
399 indicate *E. coli ΔpabABC* harboring pUC-NE1434H and *E. coli ΔpabABC* harboring pUC18
400 (control), respectively.
401
402 **Figure 4** Phylogenetic tree of NE1434 and its orthologs.
403 The sequences of the NE1434 gene product and its orthologs were aligned using MEGA5 (8).
404 A phylogenetic tree was constructed using the neighbor-joining method. The numbers on the
405 tree indicate percentages from bootstrap sampling with 1000 replications.
406
407 **Figure 5** Complementation assay of the *E. coli ΔpabABC* mutant using the *CT610* gene
408 from *C. trachomatis* and the *pqqC* gene from *P. putida.*
409 For the pABA auxotrophy complementation assay, pUC-CT610H and pUC-pqqC were
410 introduced into the *E. coli ΔpabABC* mutant and the growth of the transformants was
411 examined. Experimental procedures were the same as those used for the same experiment
412 with *E. coli ΔpabABC* harboring pUC-NE1434H.
413

414 **Figure 1**

415

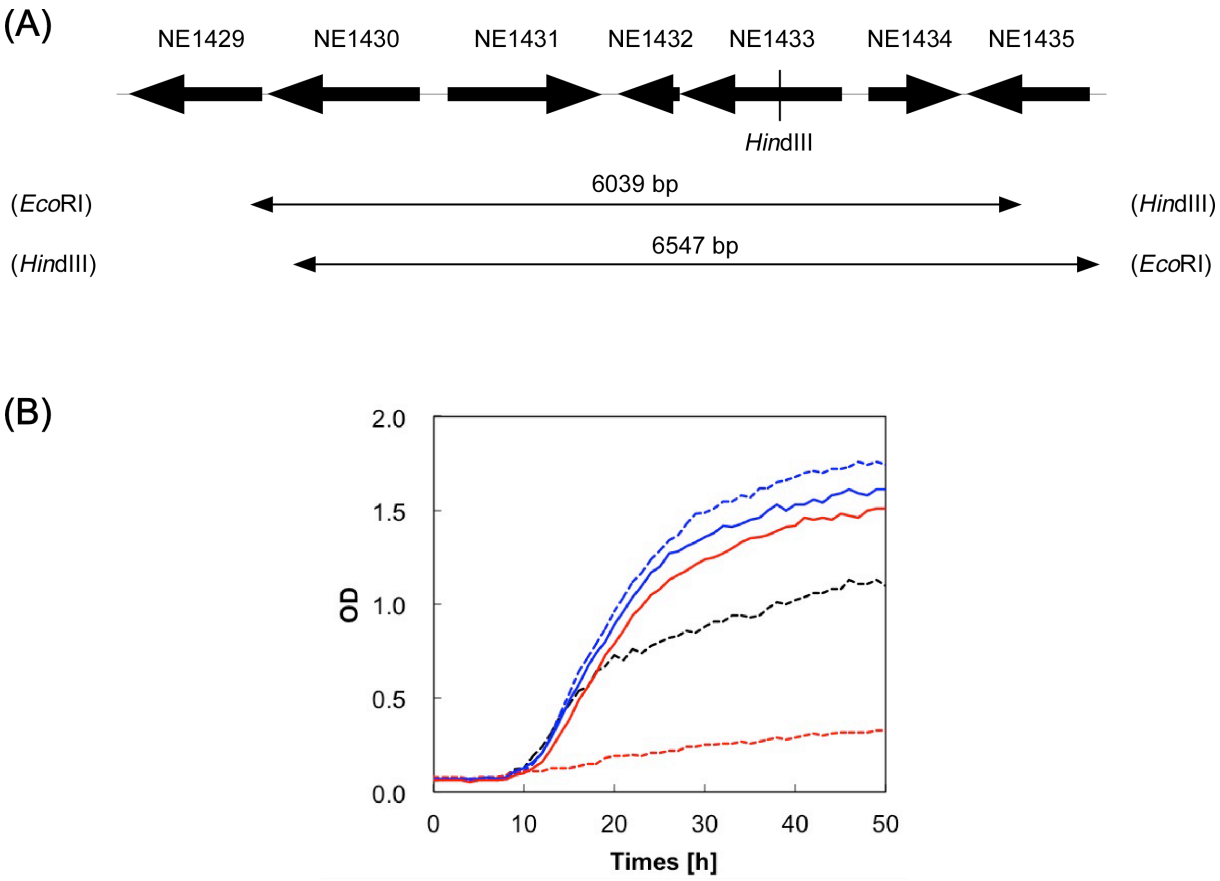


416

417

418 **Figure 2**

419



420

421

Figure 3

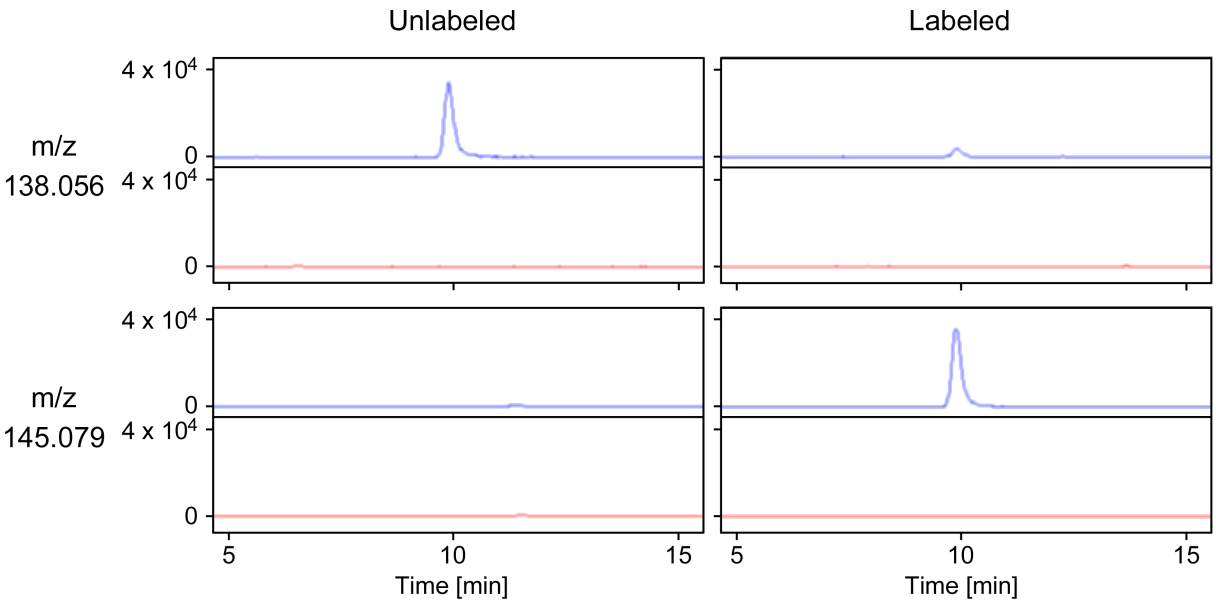

Unlabeled
Labeled
m/z
138.056
m/z
145.079
4 x 10^4
0
5
10
15
Time [min]

Figure 4

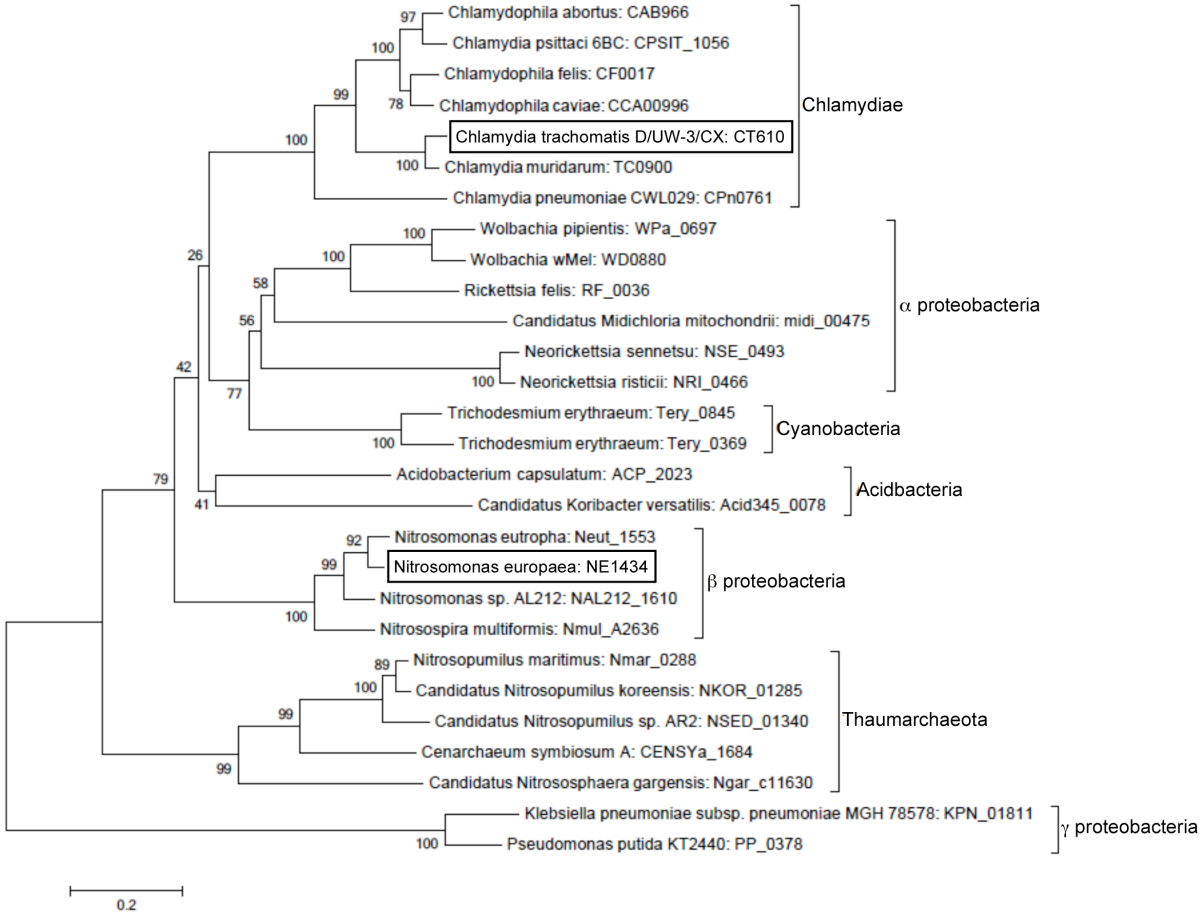

Chlamydophila abortus: CAB966
Chlamydia psittaci 6BC: CPSIT_1056
Chlamydophila felis: CF0017
Chlamydophila caviae: CCA00996
Chlamydia trachomatis D/UW-3/CX: CT610
Chlamydia muridarum: TC0900
Chlamydia pneumoniae CWL029: CPn0761
Chlamydiae
Wolbachia pipientis: WPa_0697
Wolbachia wMel: WD0880
Rickettsia felis: RF_0036
Candidatus Midichloria mitochondrii: midi_00475
Neorickettsia sennetsu: NSE_0493
Neorickettsia risticii: NRI_0466
α proteobacteria
Trichodesmium erythraeum: Tery_0845
Trichodesmium erythraeum: Tery_0369
Cyanobacteria
Acidobacterium capsulatum: ACP_2023
Candidatus Koribacter versatilis: Acid345_0078
Acidbacteria
Nitrosomonas eutropha: Neut_1553
Nitrosomonas europaea: NE1434
Nitrosomonas sp. AL212: NAL212_1610
Nitrosospira multiformis: Nmul_A2636
β proteobacteria
Nitrosopumilus maritimus: Nmar_0288
Candidatus Nitrosopumilus koreensis: NKOR_01285
Candidatus Nitrosopumilus sp. AR2: NSED_01340
Cenarchaeum symbiosum A: CENSYa_1684
Candidatus Nitrososphaera gargensis: Ngar_c11630
Thaumarchaeota
Klebsiella pneumoniae subsp. pneumoniae MGH 78578: KPN_01811
Pseudomonas putida KT2440: PP_0378
γ proteobacteria
0.2

Figure 5

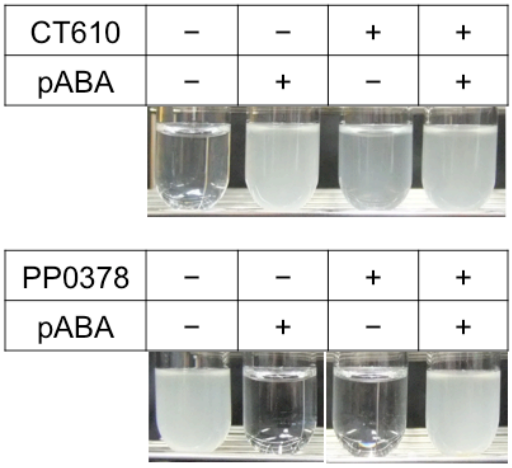


Table 1. Strains used in this study

Strains	Description	Source
E. coli JM109	*F'[traD36, proAB, lacI^q, lacZ△M15], △(lac-proAB), hsdR17 ($r_k^-m_k^+$), recA1, endA1, relA1, supE44, thi-1, gyrA96, e14$^-$(mcrA$^-$)*	Nippon Gene
E. coli BW25113	*rrnB △lacZ4787 hsdR514 △(araBAD)567△(rhaBAD)568 rph-1*	NIG
E. coli △pabABC	BW25113 derivative, *△pabA, -B,* and *-C*	(2)
L. fermentum IFO 3956	wild type	NBRC
N. europaea NBRC 14298	wild type	NBRC
C. trachomatis UW-3/Cx ATCC VR-885	wild type	ATCC
P. putida NBRC 100650	KT2440 derivative; *rpoN*::Km^r,	NBRC

NIG; National Institute of Genetics, Shizuoka, Japan

Table 2. Plasmids used in this study

Plasmids	Description	Source
pET28	N-terminus His-tag fused protein expression vector, Km^r	Novagen
pET-NE2150	pET28 derivative; N-terminus His-tag fused NE2150	This study
pSTV29	cloning vector, Cm^r	Takara Bio
pUC18	cloning vector, Ap^r	Takara Bio
pUC-NE1434H	pUC18 derivative, C-terminus His-tag fused NE1434	This study
pUC-CT610H	pUC18 derivative, CT610 from *C. trachomatis*	This study
pUC-pqqC	pUC18 derivative, *pqqC* from *P. putida* KT2440	This study

436