Adaptive Evolution of *Escherichia coli* K-12 MG1655 during Growth on a Nonnative Carbon Source, 1,1,2-Propanediol

Dae-Hee Lee and Bernhard Ø. Palsson

Department of Bioengineering, University of California—San Diego, La Jolla, California 92093-0412

Received 10 February 2010/Accepted 23 April 2010

Laboratory adaptive evolution studies can provide key information to address a wide range of issues in evolutionary biology. Such studies have been limited thus far by the inability of workers to readily detect mutations in evolved microbial strains on a genome scale. This limitation has now been overcome by recently developed genome sequencing technology that allows workers to identify all accumulated mutations that appear during laboratory adaptive evolution. In this study, we evolved *Escherichia coli* K-12 MG1655 with a nonnative carbon source, 1,1,2-propanediol (1,1,2-PDO), for ~700 generations. We found that (i) experimental evolution of *E. coli* for ~700 generations in 1,2-PDO-supplemented minimal medium resulted in acquisition of the ability to use 1,1,2-PDO as a sole carbon and energy source so that the organism changed from an organism that did not grow at all initially to an organism that had a growth rate of 0.35 h\(^{-1}\); (ii) six mutations detected by whole-genome resequencing accumulated in the evolved *E. coli* mutant over the course of adaptive evolution on 1,1,2-PDO; (iii) five of the six mutations were within coding regions, and IS5 was inserted between two *fuc* regulons; (iv) two major mutations (mutations in *fucO* and its promoter) involved in 1,1,2-PDO catabolism appeared early during adaptive evolution; and (v) multiple defined knock-in mutant strains with all of the mutations had growth rates essentially matching that of the evolved strain. These results provide insight into the genetic basis underlying microbial evolution for growth on a nonnative substrate.

Evolution of microorganisms in the laboratory offers the possibility of relating acquired mutations to increased fitness of the organism under the conditions used. Complete identification of mutations over defined evolutionary periods is necessary to fully understand the evolutionary change because spontaneous mutation is the foundational biological source of phenotypic variation (52). Since microbes grow rapidly and have large population sizes and since ancestors can be preserved by freezing them for later direct comparison of evolved types, laboratory evolution using microorganisms provides a powerful context for studying the genetics of evolutionary adaptation (5, 12, 14, 19, 43) due to the advent of new technologies for genome-wide detection of mutations (30, 33). A large number of studies of experimental evolution with various microorganisms have been carried out using natural carbon sources, especially glucose (12, 19, 47, 55), since glucose is the preferred growth requirements and to achieve optimal overall performance in the new conditions (20, 21, 34). However, a comprehensive analysis of the genetic basis of adaptation to nonnative carbon sources has not been performed.

The K-12 MG1655 strain of *E. coli* is not able to utilize 1,1,2-propanediol (1,1,2-PDO) as a sole carbon and energy source. However, *E. coli* has an enzyme, 1,1,2-PDO oxidoreductase (POR), which is involved in fermentative *t*-fucose metabolism and catalyzes the oxidation of 1,1,2-PDO to *t*-lactaldehyde (Fig. 1A). The *E. coli* POR is encoded by the *fucO* gene of the fucose regulon (11, 23), which consists of two divergent operons (*fucAO* and *fucPIKUR*) under positive control of FucR (Fig. 1B) (9). FucR is activated by fuculose-1-phosphate, which is the inducer of the *fuc* regulon (3). In *E. coli*, fucose metabolism is initiated by the sequential actions of a permease (encoded by *fucP*), an isomerase (encoded by *fucI*), a kinase (encoded by *fucK*), and an aldolase (encoded by *fucA*). The aldolase catalyzes the cleavage of fuculose-1-phosphate to dihydroxyacetone phosphate and *t*-lactaldehyde. Under aerobic respiratory conditions, *t*-lactaldehyde is oxidized to *t*-lactate by an NAD-linked aldehyde dehydrogenase with broad functions (encoded by *aldA*). 1-Lactate is then oxidized to pyruvate by a flavin adenine dinucleotide (FAD)-dependent lactate dehydrogenase (encoded by the *lldD* gene of the *lldPRD* operon [formerly the *letPRD* operon]). Under anaerobic fermentative conditions, however, redox balance requires oxidation of the *t*-lactaldehyde as a hydrogen acceptor at the expense of NADH (Fig. 1A). This reaction is catalyzed by the POR. The terminal fermentation product, 1,1,2-PDO, is then released by a permease (57). Although the POR catalyzes the oxidation of 1,1,2-PDO to 1,1,2-lactaldehyde, 1,1,2-PDO cannot be utilized by wild-type (WT) *E. coli* as a sole carbon source under aerobic conditions because this compound cannot induce expression of the *fuc* regulon (11). Indeed, the *fuc* regulon...
lon was not expressed under any conditions when a database of 213 expression profiles produced in our laboratory was examined (38). Furthermore, even if the POR is expressed, it is oxidatively inactivated by a metal-catalyzed oxidation (MCO) mechanism (7).

Sridhara et al. (48) previously described E. coli mutants that were isolated from an E. coli K-12 derivative treated with the mutagen ethyl methanesulfonate and were able to grow aerobically on L-1,2-PDO as a sole carbon source. Previous studies showed that an IS5 insertion between the fucAO and fucPIKUR operons caused constitutive expression of the fucAO operon (9, 41) at a level that enabled the E. coli mutant to grow on L-1,2-PDO. In addition, mutations resulting in increased resistance to MCO under aerobic conditions were found in the N-terminal domain of POR (39). However, at present, little is known about the accumulated genome-wide mutations and their effects on the fitness in E. coli that has acquired the ability to use L-1,2-PDO because previous studies have focused on mutations in POR and its regulatory region.

In an attempt to investigate the genetic basis of adaptive
evolution of E. coli during growth on L-1,2-PDO, we first isolated an E. coli mutant able to use L-1,2-PDO using experimental evolution without a mutagen, and we then characterized this evolved E. coli mutant. Using whole-genome sequencing, we identified all accumulated mutations of the evolved E. coli mutant related to the known ancestor and also determined the fitness benefits and phenotypic behaviors of the mutations discovered. Our results offer a systematic view of the genetic basis underlying microbial adaptation to a non-native substrate.

Materials and Methods

Strains and media. All strains and plasmids used in this study are listed in Table S1 in the supplemental material. A glyceral-evolved E. coli K-12 MG1655 strain [GC strain; F’ Δ(gfucose 1,5-lactone operon)] (30) was used as a parent strain for adaptive evolution on 1,2-PDO. Evolution of the GC strain was carried out at 37°C using 200 ml of M9 minimal medium supplemented with 2 g/liter of 1,2-PDO (catalogue number 398039; Sigma Aldrich) in 500-ml Erlenmeyer flasks containing magnetic stir bars for aeration. M9 minimal medium contained (per liter of deionized water) 0.8 g of NaH2PO4, 0.5 g of NaCl, 7.5 g of Na2HPO4·2H2O, and 3.0 g of KH2PO4. The following components were sterilized separately and then added (per liter [final volume] of medium): 2 ml of 1 M MgSO4·7H2O, 0.1 ml of 1 M CaCl2, and 0.5 ml of a trace element solution containing (per liter) 1 g of FeCl3·6H2O, 0.18 g of ZnSO4·7H2O, 0.12 g of CuCl2·2H2O, 0.12 g of MnSO4·H2O, and 0.18 g of CoCl2·6H2O. During the early stage of adaptive evolution, the minimal medium was also supplemented with 2 g/liter of glyceral, and the concentration of this compound was gradually decreased during the course of evolution while the 1,2-PDO concentration was increased to keep the total carbon source concentration in the minimal medium 2 g/liter. E. coli K-12 MG1655 (∼ ATCC 700620) obtained from the American Type Culture Collection (Manassas, VA) and the GC strain were used as parent strains for construction of all knock-in mutants in this study. During the knock-in process, the strains were cultured on Luria-Bertani (LB) medium supplemented with 50 µg/ml of kanamycin or 100 µg/ml of ampicillin when necessary.

Generation of mutant strains. The knock-in E. coli mutants were generated by homologous recombination using the gene gorging method or the lambda Red recombinase system (15, 29). The gene gorging method introduces a desired mutation without direct selection and therefore leaves no antibiotic resistance gene or other trace of having been used. Plasmids pKD46, pKD13, and pACBSR were used for introduction of the recombinase gene, for homologous recombination with the target gene, and for removal of antibiotic resistance markers, respectively. In order to confirm the genotypes of all mutants, colonies were isolated from solid medium and tested using PCR. Wild-type E. coli colonies were used as control strains for negative control experiments.

Adaptive evolution. At the start of adaptive evolution, the GC strain was cultured on solid M9 minimal medium containing 2 g/liter of glyceral and incubated overnight at 37°C. A single colony was selected from the plate that was incubated overnight at 37°C. A single colony was selected from the plate that was incubated overnight at 37°C. A single colony was selected from the plate that was incubated overnight at 37°C. A single colony was selected from the plate that was incubated overnight at 37°C. A single colony was selected from the plate that was incubated overnight at 37°C. A single colony was selected from the plate. The OD600 nm (OD600) was determined, and cells were transferred into fresh medium. The total number of generations was estimated daily by calculating the starting optical density of the OD600 of 1. Triplicate cultures were evolved concurrently under identical conditions. The OD600 was typically between 1 and 99 bp long. Mutations were not permitted to overlap. The number of alignment mismatches allowed. This iterative limit increased from 0 to 5, and unaligned reads were used as input for the next iteration, which had a new mismatch score limit. The final mismatch score limit was then used to compile the read alignments into a nucleotide resolution alignment profile. The consistency and coverage were then assessed to identify likely polymorphic locations. Locations at which coverage was greater than 10× and for which indels were observed or the count for a single-nucleotide polymorphism (SNP) was greater than twice the count for the nucleotide matching the reference sequence were considered to be likely polymorphic locations. False-negative rates were determined by this sequencing method by carrying out polymorphism identification analysis using an E. coli reference sequence which had 100 SNPs, deletions, and insertions added at random and known locations. Insertion sizes were randomly and uniformly distributed between 1 and 4 bp, and deletions were between 1 and 99 bp long. Mutations were not permitted to overlap. The rates of detection of SNPs, deletions, and insertions were determined separately by determining the fraction of each type of mutation that was marked as polymorphic by the script described above when sequence data from an end point were mapped on the mutated reference genome.

qPCR. RNA samples were taken from exponentially growing cells and added to 2 volumes of RNA Protect (Qiagen). Total RNA was isolated using an RNeasy kit (Qiagen). Reverse transcription was performed with 10 µg of total RNA. The reverse transcription mixture (60 µl) contained 10 µg total RNA, 75 µg random primers, 1× 1st Strand buffer, 10 mM dithiothreitol, 0.5 mM dNTPs, 20 U of SuperScript III reverse transcriptase, and 1,500 U of Superscript II. The mixture was incubated in an iCycler (Bio-Rad) at 25°C for 10 min, at 37°C for 1 h, and then at 42°C for 1 h. The reaction mixture was incubated at 70°C for 10 min to inactivate the Superscript II. The RNA was then degraded by adding 20 µl of 1 N NaOH and incubation at 65°C for 30 min. After the incubation, 20 µl were stored at −80°C every day over the course of evolution. Primers used in this study are described in the supplemental material.

Phenotype assessment. The OD500 was measured to monitor batch cell growth. To measure the growth rate, the substrate uptake rate (SUR), and the by-product secretion rate, each population was grown in batch culture at 37°C under aerobic conditions. Aerobic cultivation was conducted in 500-ml Erlenmeyer flasks containing 200 ml of M9 minimal medium with trace elements and 2 g/liter of L-1,2-PDO as the sole carbon source. The temperature was controlled at 37°C by using a circulating water bath, and aeration was controlled with a stir bar (~1,000 rpm). Samples were taken from the batch cultures periodically through experimental growth, filtered through a 0.22-µm filter, and stored at −20°C for substrate uptake and by-product secretion analyses. The depletion of L-1,2-PDO and the secretion of acetate and other metabolites in the medium were determined using filtered supernatants by performing high-performance liquid chromatography (HPLC) (Waters). Samples were injected into an Aminex HPX-87H column (Bio-Rad) connected to a cation-H guard column (Bio-Rad) at 65°C. Metabolites were eluted with 5 mM sulfuric acid at a flow rate of 0.5 ml/min for 30 min. Detection was performed using a differential refractive index detector at 35°C, and the compounds were compared to standards. The identities of metabolites and organic acids in the fermentation broth were further verified with enzymatic kits (R-Biopharm).

Solexa resequencing. Five micrograms of genomic DNA isolated from a single colony (eBOP12-6) of the end point eBOP12 population was used to generate a genomic DNA library using an Illumina genomic DNA library generation kit by following the manufacturer’s protocol (Illumina Inc., San Diego, CA). Briefly, bacterial genomic DNA was fragmented by nebulization. The ends of fragmented DNA were repaired by T4 DNA polymerase, Klenow DNA polymerase, and the inclusion of phosphorylated termini. The repaired DNA was then used to add an A base to the 3’ end of the DNA fragments. After ligation of the adapters to the ends of the DNA fragments, the ligated DNA fragments were subjected to electrophoresis on a 6% × 1× Tris-borate-EDTA (TBE) gel. DNA fragments ranging from 190 bp to 220 bp long were recovered from the gel and purified using a Qiagen minigel purification kit. Finally, the adapter-modified DNA fragments were enriched by PCR. The final concentration of the genomic DNA library was determined by using a NanoDrop instrument (Thermo Scientific), and the results were validated by using a 6% × 1× TBE gel. The genomic DNA library was used to generate a cluster on a Flowcell by following the manufacturer’s protocol. The V2 genomic sequencing primer was used for all DNA sequencing. A 36-cycle sequencing program was used with an Illumina genome analyzer II by following the manufacturer’s protocol.

Genome sequence assembly and identification of polymorphism. The Solexa output for the resequencing run was first curated to remove any sequences containing a period. We then used MosaicAligner, developed by M. P. Stromberg and G. T. Marth (unpublished data), to iteratively align reads with the E. coli reference sequence (gi 48999473); in each iteration a limit was placed on the number of alignment mismatches allowed. This iterative limit increased from 0 to 5, and unaligned reads were used as input for the next iteration, which had a new mismatch score limit. A new mismatch score limit was then used to compile the read alignments into a nucleotide resolution alignment profile. The consistency and coverage were then assessed to identify likely polymorphic locations. Locations at which coverage was greater than 10× and for which indels were observed or the count for a single-nucleotide polymorphism (SNP) was greater than twice the count for the nucleotide matching the reference sequence were considered to be likely polymorphic locations. False-negative rates were determined by this sequencing method by carrying out polymorphism identification analysis using an E. coli reference sequence which had 100 SNPs, deletions, and insertions added at random and known locations. Insertion sizes were randomly and uniformly distributed between 1 and 4 bp, and deletions were between 1 and 99 bp long. Mutations were not permitted to overlap. The rates of detection of SNPs, deletions, and insertions were determined separately by determining the fraction of each type of mutation that was marked as polymorphic by the script described above when sequence data from an end point were mapped on the mutated reference genome.
of 1 N HCl was added to neutralize the solution. QIAquick PCR purification kits were used to clean up the cDNA synthesis product. Following purification, the cDNA was quantified and then directly used in real-time quantitative PCR (qPCR). The 50-μl qPCR mixtures contained 25 μl of SYBR green PCR master mixture (Qiagen), 0.2 μM forward primer, 0.2 μM reverse primer, and cDNA as a template. Each qPCR was performed in triplicate in the iCycler using the following conditions: 95°C for 15 min and then 40 cycles of denaturation 94°C for 15 s, annealing at 52°C for 30 s, and extension at 72°C for 30 s. The gene expression of evolved strains was analyzed using minimal medium containing glycerol and was compared to that of the wild-type strain under same growth conditions. In order to determine the binding affinity of each primer set, a standard curve was constructed for each primer, and the reaction efficiency was obtained by using it. Using the standard curve, the relative quantity of cDNA was determined for each gene by normalizing the quantity to the quantity of acpP (acyl carrier protein) cDNA in the same sample. acpP was chosen as the internal control gene since it is constitutively expressed in wild-type *E. coli* K-12 (13).

**POR expression, purification, and assay.** The WT and mutant alleles of the *fucO* gene encoding POR were amplified by PCR using genomic DNA of the GC strain and the L-1,2-PDO-evolved mutant as templates, respectively, and were cloned into pGEX-6P-1 (GE Healthcare Life Sciences). Potential clones were checked for the presence of the correct insert in the correct orientation by restriction enzyme analysis, followed by Sanger sequencing. *E. coli* BL21 (New England Biolabs) was used for expression of WT and mutant alleles of the *fucO* gene and was cultured in LB medium supplemented with 100 μg/ml ampicillin at 30°C. Protein overexpression was induced with isopropyl-β-D-thiogalactoside (IPTG), and cell extracts were prepared by sonication and centrifugation. WT and mutant POR encoded by *fucO* were purified using a GSTrap Fast Flow column (GE Healthcare Life Sciences), and the glutathione S-transferase (GST) moiety was removed by on-column cleavage with PreScission protease (GE Healthcare Life Sciences), resulting in intact POR. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was conducted to confirm the purity of the purified samples. POR was routinely assayed using a method described previously (39). Glycoaldehyde was used as a substrate because it has been used as an alternative substrate for the forward reaction (l-lactaldehyde → t-L-1,2-PDO) of the enzyme (6, 39) and is commercially available. The activity of POR was measured in the forward direction and in the reverse direction (t-L-1,2-PDO → l-lactaldehyde) by monitoring NAD(H) release at 340 nm. The l isomer of L-1,2-PDO (catalogue number 540250; Sigma) was used for the POR assay. A calibration curve was constructed using enzyme standards and chemicals purchased from Sigma Chemicals. The protein content was determined with the Quant-it assay system (Invitrogen) by following the manufacturer's instructions.

**Growth rate measurement.** Growth rate experiments with knock-in *E. coli* mutants were performed by measuring the OD{}_{600} of triplicate cultures at several time points at which the OD{}_{600} was 0.03-0.05. A positive maximal growth rate then steadily increased during laboratory evolution. Although glucose is a poor carbon source for WT strain MG1655 cell growth in M9 minimal medium (Fig. 2A). After the glycerol was completely removed from the medium, the average growth rate was 0.0035 ± 0.0005 h⁻¹ in 1,2-PDO-supplemented minimal medium. However, the growth rate then steadily increased during laboratory evolution over the following ~450 generations. Laboratory evolution of the GC strain on 1,2-PDO-supplemented minimal medium stopped once a stable growth rate was achieved and resulted in the evolved eBOP12, eBOP13, and eBOP14 populations. The number of cell divisions that occurred during evolution was estimated based on the amount of cells transferred each day and the doubling time. The total number of cell divisions was 3.85 × 10^{11} ± 0.3 × 10^{11} (Fig. 2A).

**Allele frequency estimation.** In order to estimate the temporal appearance of acquired mutations, the approximate time point that each mutation was fixed in the relevant population was estimated by screening the frozen stocks of cultures saved at intermediate time points during each evolution with t-L-1,2-PDO. The predominant presence or absence of each mutation at a time point was determined by performing PCR for the 200- to 300-bp regions surrounding each mutation, followed by Sanger sequencing. If the mutation was not fixed, we considered the larger peak the dominant signal for two peaks at one position in a Sanger sequencing chromatogram.

**RESULTS**

**Adaptive evolution.** Initially, WT strain MG1655 and the GC strain were each examined to determine their abilities to utilize L-1,2-PDO as a sole source of carbon and energy. Both strains showed no growth in M9 minimal medium supplemented with 4 g/liter of a racemic L-1,2-PDO mixture. Therefore, another carbon source was needed to support the growth of these strains in minimal medium containing 1,2-PDO. We assumed that *E. coli* could grow on a natural carbon source and then utilize the 1,2-PDO during adaptation. Although glucose is often the preferred carbon for *E. coli* cell growth, it inhibits utilization of other carbon sources by a mechanism known as catabolite repression (16). Unlike the non-phosphotransferase system (PTS) sugars (e.g., lactose, melibiose, maltose, gluconate, and glucose-6-phosphate) that cause catabolite repression in *E. coli* (31, 32), glycerol is a non-PTS carbon source for which there is not catabolite repression (18). However, it is a poor carbon source for WT strain MG1655 cell growth in M9 minimal medium. Our lab has generated several evolved *E. coli* mutants that grow well on minimal medium containing glycerol (20). One of these glycerol-evolved strains, the GC strain, has only one mutation in the *glpK* gene (encoding glycerol kinase), which was revealed by genome sequencing (30), and it showed the smallest change in the global gene expression profile (20) when its profile was compared to the WT strain MG1655 profile. Thus, the initial evolution experiment was conducted to generate a 1,2-PDO-evolved mutant by further experimental evolution of the GC strain with 2 g/liter of glycerol. Then the amount of glycerol in the medium was decreased gradually while the concentration of 1,2-PDO was increased to keep the total concentration of the two carbon sources 2 g/liter. A rapid decrease in the growth rate was observed when the amount of glycerol was reduced and the concentration of 1,2-PDO was increased during serial passage in M9 minimal medium (Fig. 2A). When the glycerol was completely removed from the medium, the average growth rate was 0.0035 ± 0.0005 h⁻¹ in 1,2-PDO-supplemented minimal medium. However, the growth rate then steadily increased during laboratory evolution over the following ~450 generations. Laboratory evolution of the GC strain on 1,2-PDO-supplemented minimal medium stopped once a stable growth rate was achieved and resulted in the evolved eBOP12, eBOP13, and eBOP14 populations. The number of cell divisions that occurred during evolution was estimated based on the amount of cells transferred each day and the doubling time. The total number of cell divisions was 3.85 × 10^{11} ± 0.3 × 10^{11} (Fig. 2A).

**Phenotypic characterization.** Phenotypic characterization of the evolved populations revealed that the populations at the three evolutionary end points had similar metabolic phenotypes. Three evolved populations, eBOP12, eBOP13, and eBOP14, had acquired the ability to grow on a racemic substrate for the forward reaction (l-lactaldehyde → t-L-1,2-PDO) of the enzyme (6, 39) and is commercially available. The activity of POR was measured in the forward direction and in the reverse direction (t-L-1,2-PDO → l-lactaldehyde) by monitoring NAD(H) release at 340 nm. The l isomer of L-1,2-PDO (catalogue number 540250; Sigma) was used for the POR assay. A calibration curve was constructed using enzyme standards and chemicals purchased from Sigma Chemicals. The protein content was determined with the Quant-it assay system (Invitrogen) by following the manufacturer's instructions.

**Allele frequency estimation.** In order to estimate the temporal appearance of acquired mutations, the approximate time point that each mutation was fixed in the relevant population was estimated by screening the frozen stocks of cultures saved at intermediate time points during each evolution with t-L-1,2-PDO. The predominant presence or absence of each mutation at a time point was determined by performing PCR for the 200- to 300-bp regions surrounding each mutation, followed by Sanger sequencing. If the mutation was not fixed, we considered the larger peak the dominant signal for two peaks at one position in a Sanger sequencing chromatogram.
evolution, the biomass yield, SUR, and product secretion rate were monitored for the three evolved populations (Fig. 2B). The results showed that there was a 9.6-fold increase in the average dry weight yield compared to the yield of the ancestral population after utilization of L-1,2-PDO was started. On average, a 5.8-fold increase in the SUR, from 1.24 ± 0.6 mmol g\(^{-1}\) h\(^{-1}\) to 7.17 ± 0.51 mmol g\(^{-1}\) h\(^{-1}\), was observed after the evolved populations started to utilize L-1,2-PDO. The major carbon flow in the evolved populations grown aerobically on L-1,2-PDO was acetate carbon flow, which decreased from 4.3 ± 0.1 mmol g (dry weight)\(^{-1}\) h\(^{-1}\) to 3.4 ± 0.2 mmol g (dry weight)\(^{-1}\) h\(^{-1}\) at the end point of evolution (Fig. 2B). Ethanol, another metabolite derived from intracellular acetyl coenzyme A (acetyl-CoA), was not secreted. Lactate and succinate were also produced, but at very low levels.

**Whole-genome sequencing.** To determine the genetic basis of adaptation, the evolved eBOP12 population was selected because it had the highest SUR and growth rate of the three evolutionary end point populations (see Table S2 in the sup-
The eBOP12 population was grown on solid minimal medium supplemented with L-1,2-PDO, and 10 random colonies were isolated for analysis. Of the 10 clones isolated, a single clone, eBOP12-6, was selected because it had the highest growth rate on minimal medium containing L-1,2-PDO (see Table S3 in the supplemental material). Genomic DNA of the eBOP12-6 clone was extracted and sequenced by using the Solexa sequencing system. When SNPs and indels were accounted for, a total of eight putative mutations were found to have accumulated in the evolutionary lineage from the GC strain to eBOP12-6. Six of these mutations were confirmed by PCR amplification and Sanger sequencing, which revealed three SNPs, one deletion, and two insertions, including a transposable element. Five of the six mutations were in coding regions, and there was an IS5 insertion in the fur operon. All confirmed mutations are listed in Table 1.

An SNP was found in the furO gene encoding the POR (Table 1), which catalyzes the first step of L-1,2-PDO catabolism in E. coli (see Fig. S1 in the supplemental material). It seems likely that this mutation allowed the evolved population to utilize L-1,2-PDO by overcoming metal-catalyzed inactivation (39). An insertion sequence (IS5) was found in the region between the furCO and furPIKUR operons (see Fig. S1 in the supplemental material), which might have caused constitutive activation of the furCO operon (9, 41). The L-1,2-PDO-evolved eBOP12-6 clone contained two genes (ilvG and ylbE1) that are not expressed in WT E. coli K-12 MG1655 due to internal frameshift mutations (22). The ilvG and ylbE1 genes, which encode acetohydroxy acid synthase (AHAS) II and a predicted protein, respectively, were expressed in the eBOP12-6 clone by deletion and insertion of single nucleotides in their coding regions, respectively. In addition, a synonymous mutation occurred in the coding region of the ylbE1 gene (Table 1). The ilvG gene is known to be involved in biosynthesis of valine and isoleucine by conversion of pyruvate to 2-acetohydroxy acid. Because L-1,2-PDO enters central metabolism as pyruvate, it seems likely that the mutation in ilvG improves the flow of carbon for amino acid biosynthesis. The evolved eBOP12-6 clone also had a mutation in the mID gene encoding 23S rRNA, which is the RNA component of the large subunit (50S subunit) of the E. coli ribosome. The mID mutation found in the eBOP12-6 clone is located in domain V (U2016 to G2625) of 23S rRNA of E. coli, which has been identified as the peptidyl transferase center that acts as a general protein folding modulator by binding to unfolded proteins (45). Furthermore, the mutation is close to one of the five recently identified sites in domain V (G2553 to C2556) where proteins bind during refolding (44).

Analysis of the mutations found. To determine where the six mutations that accumulated in the eBOP12-6 clone were fixed in the eBOP12 population at the end point of adaptive evolution, the genotypes of 15 randomly selected clones were determined for each mutation locus. Five mutations appear to have been fixed during the 700-generation evolution, whereas one locus (mID) was fixed at a frequency of nearly 0.87 (Table 1). In addition, we screened the six mutations that were found in the eBOP12 population in the eBOP13 and eBOP14 populations by Sanger sequencing. The synonymous mutation in ylbE1 was not detected in the eBOP13 and eBOP14 populations at the end point of adaptive evolution. However, four other mutations (furO, ilvG, ylbE1, and IS5 insertion) were all fixed in both populations (see Table S4 in the supplemental material).

In order to determine the approximate time of appearance of each mutation in the eBOP12 population, the frozen stock samples obtained at intermediate times during laboratory evolution were screened for the appearance of each mutation found in the end point population by Sanger sequencing of PCR-amplified mutated regions (Fig. 3). An SNP was considered present if the dominant signal peak resulting from Sanger sequencing indicated that the mutation was present, although at times lower levels of SNPs were observed in the population as nondominant peaks in the sequencing trace. Sequencing of intermediate points showed that mutations occurred over the course of adaptive evolution. The increase in the growth rate was greatest during the first approximately 100 generations (250 to 350 generations in Fig. 3) after the evolved populations started to utilize L-1,2-PDO as the sole carbon source, which corresponded to the appearance of two major mutations (furO and its promoter region) that are directly involved in L-1,2-PDO catabolism in the eBOP12-6 clone. The eBOP13 and eBOP14 populations also had two mutations during the first approximately 100 generations after the evolved populations began to take up L-1,2-PDO (see Fig. S2 in the supplemental material). It was reasonable to expect that the furO mutation and IS5 insertion would be detected simultaneously when the eBOP12 population started to utilize only L-1,2-PDO.
PDO. However, the fucO mutation was not dominant in the sequence trace in this screen, suggesting that it was not fixed yet in the eBOP12 population when L-1,2-PDO utilization started (~250 generations). For the unfixed fucO mutation, we screened several individual colonies of the eBOP12 population after ~250 generations. Of 15 eBOP12 colonies, 10 had only the IS5 insertion, 1 had only the fucO mutation, 1 had both the fucO and IS5 mutations, and the remaining 3 had neither mutation, suggesting that the fucO mutation was present at a very low frequency in the adaptive culture after ~250 generations.

To evaluate the contributions of individual mutations that enabled organisms to grow on minimal medium containing L-1,2-PDO and that improved fitness, we introduced each mutation into the WT and GC strains using a site-directed mutagenesis strategy called gene gorging (29). The growth rate recovery of the mutants constructed using M9 minimal medium containing L-1,2-PDO is shown in Fig. 4. A growth rate recovery of 0% indicates that the mutant did not grow in minimal medium containing L-1,2-PDO, while a mutant with a growth rate recovery of 100% grew at the same rate as the eBOP12-6 clone. We found that mutations in the fucO gene and its promoter region resulted in the largest growth rate recoveries in the GC (61%) and WT (53%) strains (Fig. 4). Furthermore, these mutations were essential for utilization of L-1,2-PDO in E. coli because a single fucO or IS5 insertion mutation did not allow growth of the WT or GC strain in minimal medium containing L-1,2-PDO (Fig. 4). In addition to the GC strain mutant, the WT strain mutant with two mutations in fucO and its promoter region could grow in minimal medium containing L-1,2-PDO. This result suggests that adaptation to growth with L-1,2-PDO as the only carbon source did not occur with a background of previous glycerol adaptation. Besides single mutants, we also made multiple mutants to reconstruct the eBOP12-6 clone isolated from the eBOP12 population. Mutations in ilvG and/or ylbE with a background of fucO and its promoter mutation also resulted in growth rate recoveries from 11% to 22%. The effect of rldD mutation on the growth rate was not significant. The growth rates of reconstructed strains with all mutations almost matched the growth rate of the evolved eBOP12 clone in L-1,2-PDO-supplemented minimal medium. This indicates that we were able to identify all of the important mutations in this evolved clone through whole-genome sequencing and that there was not an influential epigenetic component of the adaptation to L-1,2-PDO.

**Functional characterization of mutations.** To verify the exact nature of the mutations and their phenotypic consequence, the fucO alleles were amplified by PCR using genomic DNA of the GC and eBOP12-6 strains and cloned into pGEX-6P-1 containing GST as a fusion partner. WT or mutant POR expressed in E. coli BL21 was purified by affinity chromatography. When kinetic constants were investigated by using reciprocal plots for saturating concentrations of the corresponding enzymes, the $K_m$ values of the mutant POR for L-1,2-PDO and glycoaldehyde were 1.2-fold and 2.1-fold higher than those of the WT POR, indicating that the mutant POR had low affinity for substrates of the reduction and oxidation reaction. However, the $V_{\text{max}}$ of the mutant POR for L-1,2-PDO was 10-fold.
greater than that of the WT POR, while the \(V_{\text{max}}\) values of the two enzymes for glycoaldehyde were similar. No difference in the \(K_m\) and \(V_{\text{max}}\) values for NAD and NADH was found between the two enzymes (Table 2). Thus, the ability to grow and to grow rapidly on \(1,1,2\)-PDO can be partially attributed to altered kinetic properties of POR.

To determine the effect of IS5 inserted between the \(fucAO\) and \(fucPIKUR\) operons on \(fucAO\) expression, the gene expression of the \(fucAO\) operon in the eBOP12-6 clone was analyzed using minimal medium containing glycerol or \(1,1,2\)-PDO and compared to that of the parental GC strain grown on glycerol-containing M9 minimal medium (Fig. 5). In glycerol-containing minimal medium, the expression of the \(fucAO\) operon in evolved clone eBOP12-6 was increased 11-fold. The \(fucAO\) expression of the eBOP12-6 clone in minimal medium containing \(1,1,2\)-PDO was also upregulated 12-fold compared with that of the GC strain in glycerol-containing minimal medium, indicating that the \(fucAO\) operon was constitutively expressed in the evolved eBOP12-6 strain regardless of the carbon source. In order to determine the effect of \(iibG\) and \(ylbE1\) mutations in this study, the gene expression of the \(iibGMEDA\) and \(fdrA-k_{\text{ylbE1}}-ybcF-ybcF\) operons in the eBOP12-6 clone in glycerol-containing minimal medium was analyzed and compared to that in the parental GC strain (Fig. 5). It is known that mutations that restore the reading frame of the \(iibG\) gene result in a 5- to 10-fold increase in expression of the \(iibEDA\) genes (37, 54). Consistent with previous studies, the \(iibG\) mutation (single nucleotide deletion at bp 977) resulted in 3- to 4-fold upregulation of the expression of the distal genes (\(iibE\) and \(iibA\)) in the same operon of the eBOP12-6 clone (Fig. 5). Because \(1,1,2\)-PDO enters central metabolism via pyruvate, it seems likely that the mutation in the \(iibG\) gene improves the flow of carbon for amino acid biosynthesis. However, there was no difference in expression of the genes downstream of \(yibE\) (\(yibF\) and \(ybcF\)) (data not shown).

**DISCUSSION**

Laboratory adaptive evolution can provide key information to address a wide range of issues in evolutionary biology. In this study, we used whole-genome sequencing to obtain an integrative understanding of the genetic and phenotypic changes during evolution of \(E.\ coli\) in response to a nonnative carbon source, \(1,1,2\)-PDO. We found that (i) experimental evolution of \(E.\ coli\) in 1,2-PDO-supplemented minimal medium revealed that preevolved \(E.\ coli\) acquired the ability to use \(1,1,2\)-PDO as a sole carbon and energy source; (ii) six mutations detected by whole-genome sequencing accumulated in the evolved \(E.\ coli\) clones over the course of adaptive evolution with \(1,1,2\)-PDO; (iii) five of the six mutations were in the coding region, and the IS5 insertion was between two \(fuc\) regulons; (iv) two major mutations (\(fucO\) and IS5 insertion) involved in \(1,1,2\)-PDO catabolism appeared early during adaptive evolution; and (v) multiple defined knock-in mutant strains with all of the mutations had growth rates essentially matching that of the evolved strain.

We found that \(E.\ coli\) could acquire the ability to grow on \(1,1,2\)-PDO through laboratory evolution without a mutagen. Although an \(E.\ coli\) mutant able to utilize \(1,1,2\)-PDO was reported previously, this mutant was isolated after mutagen treatment (48). Later, a spontaneous \(1,1,2\)-PDO-utilizing mutant was isolated from an \(E.\ coli\) K-12 derivative pregrown on fucrose (24), suggesting that POR-induced \(E.\ coli\) K-12 was
used to generate the 1,1,2-PDO-evolved *E. coli* strain. In contrast to our assumption concerning catabolite repression and gradual adaptation, Hacking and Lin (24) cultured WT cells pregrown on fucose with glucose to support cell growth in minimal medium containing 1,2-PDO. In addition, evolutionary trajectories of fitness during the overall evolution were not shown in previous studies. We used a natural non-PTS substrate without catabolite repression, glycerol, to support the initial growth of glycerol-evolved *E. coli* in minimal medium containing 1,2-PDO. The amount of glycerol was decreased gradually for ~250 initial generations, while the concentration of 1,2-PDO was increased. During adaptation using this two-substrate strategy, *E. coli* could utilize 1,1,2-PDO as a sole carbon source, and there was a significant increase in the growth rate on 1,1,2-PDO over the course of experimental evolution (Fig. 2). This result indicates that experimental approaches to studying evolution with microorganisms can be expanded to nonnative carbon sources. Based on fitness, the best representative (clone eBOP12-6) of the eBOP12 population at the end point of evolution was selected to determine genetic mechanisms of evolution with a nonnative carbon source.

Mutations can be found using new whole-genome sequencing technology. Recent next-generation DNA sequencing technology allows us to determine how many mutations are acquired during experimental evolution, where the mutations occur, the order of acquisition of mutations, and the contribution of mutations to fitness. Whole-genome sequencing with targeted Sanger sequencing revealed that only a small number of mutations accumulated during adaptation to 1,1,2-PDO. The first study of *E. coli* that adaptively evolved in glycerol-containing minimal medium found that only a small number of mutations (two or three mutations) are required for adaptation to the new environment (30). This result suggests that marked changes in phenotype can be mediated by as few as two mutations, in agreement with results for *Saccharomyces cerevisiae*, maize, and the influenza virus (49, 53, 56). None of the mutations were found to occur in the intergenic or promoter region.

In contrast, the number of mutations found in the present study of 1,1,2-PDO-evolved *E. coli* was higher (6 mutations); one of the mutations occurred in the promoter region, although the rest of the mutations were found in coding regions. The results suggested that adaptive evolution of *E. coli* with a nonnative carbon source, 1,1,2-PDO, is more complex than adaptive evolution with a native carbon source. No large-scale genomic duplication or deletion distinguished the evolved strain from its ancestor. The synonymous mutation rate should equal the expected mutation rate (36) because synonymous sites are not a selective constraint. In *E. coli*, the point mutation rate is very low if the mutations are selectively neutral (4.5 \times 10^{-9} mutation per synonymous site per year, 4.1 \times 10^{-10} mutation per bp, or 1.6 \times 10^{-10} mutation per bp per generation) (2, 17, 40). Our finding that only one synonymous substitution (*ylbE1*) was fixed during the evolution with 1,1,2-PDO for ~700 generations is consistent with a low neutral mutation rate. We identified two major genetic events that specifically targeted 1,1,2-PDO catabolism in the evolved population (Table 1). The first mutation that we found was the IS5 insertion between the *fucAO* and *fucPIKUR* operons, which resulted in constitutive activation of the *fucAO* operon (Fig. 5) (49, 53, 56). Mobile genetic elements, particularly insertion sequence elements, excise from and insert into DNA at a relatively high frequency. Movement of these elements is one of the most common types of genetic changes and is considered an important factor in evolution (25). The *E. coli* K-12 MG1655 genome contains seven different *IS* elements (a total of 26 copies), the most prevalent of which are IS7 and IS5 (46). In the case of the *bgl* operon, insertion of IS1 or IS5 in the *cis* regulatory region, *bglR*, can allow induction expression (42). The levels of activity of the *tdh* operon enzymes are too low to allow WT *E. coli* cells to utilize threonine as a carbon source. However, IS3 insertion activated the *tdh* promoter constitutively, resulting in *E. coli* cells that utilize threonine as a sole carbon source (1). The second event was a single nucleotide alteration (C22 \rightarrow A22) in the *fucO* gene that resulted in an N-terminal hydrophobic amino acid substitution (Leu8 \rightarrow Met9). It is well known that POR is irreversibly inactivated by a metal-catalyzed oxidation system (8). An *E. coli* mutant able to use 1,1,2-PDO aerobically was reported previously (48), and it expressed constitutively a POR mutant enzyme with resistance to MCO due to an Ilv2-to-Leu7 or Leu8-to-Val9 substitution (39). This suggests that the mutation found in this study also might confer MCO resistance with POR in the evolved population. The purified POR protein from cells expressing the mutant gene showed a 10-fold increase in the reaction rate compared with the protein of the ancestral GC strain, suggesting that rapid growth on 1,1,2-PDO can be partially attributed to the altered kinetic and regulatory properties of POR.

The *ylbGMEDA* operon of *E. coli* K-12 is an amino acid biosynthesis operon required for synthesis of the branched-amino acids isoleucine, leucine, and valine (51). The first two genes of this operon, *ylbG* and *ylbM*, encode the large and small subunits of AHAS II, respectively. AHAS catalyzes the first step in the biosynthesis of the branched amino acids. The reaction involves decarboxylation of pyruvate, followed by condensation with a second molecule of pyruvate or with 2-oxobutyrate. The *ylbG* gene in WT *E. coli* K-12 contains a frameshift mutation which results in termination of translation in the middle of the gene. Given that *E. coli* K-12 does not contain a functional version of AHAS II, the presence of excess valine represses synthesis of isoleucine even during isoleucine starvation (37). This phenomenon, termed valine toxicity, is relieved by depletion of valine to levels that enable isoleucine synthesis, presumably as a result of valine catabolism.

Recently, a *ylbE1* mutation in *E. coli* K-12 resulted in anaerobic resistance to nitric oxide (35). *E. coli* K-12 MG1655 populations generated by ionizing radiation were evolved, and analysis showed that there was a 1-nucleotide insert mutation in *ylbE1* that restored the full-length reading frame of an annotated pseudogene (27). Although the crystal structure of the *ylbE* product (oxidoreductase) from *Lactococcus lactis* (PDB accession code 3DQP) was determined recently, there is no similarity between *E. coli* *ylbE* and *L. lactis* *ylbE*. We tried to characterize the *ylbE* product of the evolved *E. coli*; however, there was no evidence that *ylbE* is involved in 1,1,2-PDO metabolism.

The IS5 insertion mutation appeared first when we determined the order of acquisition of mutations in the eBOP12 population. Although the *fucO* mutation or IS5 insertion, which was introduced into the chromosomes of the WT and
GC strains, resulted in no growth in L-1,2-PDO-supplemented minimal medium, two major mutations (fucO and IS5 insertion) involved in L-1,2-PDO catabolism appeared early during adaptive evolution, which corresponded to the increase in the growth rate during the first ~100 generations (250 to 350 generations) after the evolved population started to utilize L-1,2-PDO. The evolved E. coli showed stepwise increases in the growth rate during adaptation as additional mutations were acquired (Fig. 3). It was reasonable to expect to see simultaneous mutations (mutations in the early stages of acquiring new catabolic functions. Al-though we were able to relate some of the observed adaptive mutations (mutations in fucO and its promoter) to known regulatory networks that cause increased expression of an enzyme with marginal activity with the substrate important in the early stages of acquiring new catabolic functions. Al-
though we were able to relate some of the observed adaptive mutations (mutations in fucO and its promoter) to known regulatory networks that cause increased expression of an enzyme with marginal activity with the substrate important in the early stages of acquiring new catabolic functions. Al-
though we were able to relate some of the observed adaptive mutations (mutations in fucO and its promoter) to known regulatory networks that cause increased expression of an enzyme with marginal activity with the substrate important in the early stages of acquiring new catabolic functions. Al-

The mutations identified can be used to reconstruct the phenotypic changes. The reconstructed strain with mutations in fucO and its promoter grew in minimal medium containing L-1,2-PDO, suggesting that these mutations are a prerequisite for utilization of L-1,2-PDO and provide a great fitness advantage (53 to 61%). Several studies, beginning with the classic studies on β-galactosidase (28), on amidases (26), and on rib-


fucO and IS5 insertion when the evolved E. coli started to utilize L-1,2-PDO during laboratory evolution. One possible interpretation is that the fucO mutation was present at a very low frequency in that time point.


ACKNOWLEDGMENTS

We thank Adam Feist, Kenyon Applebee, Jeff Orth, and Vasily Portnoy for helpful discussions. We also thank Christian Barrett for the whole-genome comparison and Marc Abrams for editing the manuscript.

REFERENCES


5. Cabicul, E., J. Aguilar, and J. Ros. 1994. Metal-catalyzed oxidation of Fe2+-
dehydrogenases. Consensus target sequence between propanediol oxidore-

B.Ø.P. and UCSF have a financial interest in Genomatica, Inc.; although grant GM062791 was identified for conflict-of-interest management based on the overall scope of the project and its potential to benefit Genomatica, Inc., the research findings included in this paper are not necessarily directly related to the interests of Genomatica, Inc.


ERRATUM

Adaptive Evolution of *Escherichia coli* K-12 MG1655 during Growth on a Nonnative Carbon Source, L-1,2-Propanediol

Dae-Hee Lee and Bernhard Ø. Palsson

*Department of Bioengineering, University of California, San Diego, La Jolla, California 92093-0412*

Volume 76, no. 13, pages 4158–4168, 2010. Page 4159: In Fig. 1A, the labels “Aerobic condition” and “Anaerobic condition” should be transposed.