Global Expression Profiling of Acetate-grown Escherichia coli*

Min-Kyu Oh, Lars Rohlin, Katy C. Kao, and James C. Liao‡
From the Department of Chemical Engineering, UCLA, Los Angeles, California 90095

This study characterized the transcript profile of Escherichia coli in acetate cultures using DNA microarray on glass slides. Glucose-grown cultures were used as a reference. At the 95% confidence level, 354 genes were up-regulated in acetate, while 370 genes were down-regulated compared with the glucose-grown culture. Generally, more metabolic genes were up-regulated in acetate than other gene groups, while genes involved in cell replication, transcription, and translation machinery tended to be down-regulated. It appears that E. coli commits more resources to metabolism at the expense of growth when cultured in the poor carbon source. The expression profile confirms many known features in acetate metabolism such as the induction of the glyoxylate pathway, tricarboxylic acid cycle, and gluconeogenic genes. It also provided many previously unknown features, including induction of malic enzymes, ppsA, and the glycolate pathway and repression of glycolytic and glucose phosphotransferase genes in acetate. The carbon flux derived from the malic enzymes and PpsA in acetate was further confirmed by deletion mutations. In general, the gene expression profiles qualitatively agree with the metabolic flux changes and may serve as a predictor for gene function and metabolic flux distribution.

Physiological characteristics of Escherichia coli using acetate or glucose as a sole carbon and energy source have been studied for more than three decades (1, 2). Briefly, E. coli uptakes glucose using the phosphotransferase system that converts extracellular glucose into intracellular glucose 6-phosphate, which can be further metabolized by the glycolytic pathway to produce energy and biosynthetic precursors. In the presence of glucose, the adenylate cyclase is inactive, and the cAMP level is low. In the absence of glucose, the adenylate cyclase is activated to produce cAMP, which when binding to the cAMP receptor protein activates the expression of a large set of catabolite derepression genes (2, 3). On the other hand, acetate is transported into the cell and converted to acetyl-CoA, which is further metabolized through the tricarboxylic acid cycle. The acetate-metabolizing genes are typically repressed in the presence of glucose. The induction and regulation of acetate-metabolizing genes have been studied extensively (4). Since the two carbon sources, glucose and acetate, are utilized by distinct metabolic pathways, the metabolic flux distribution differs significantly in these two carbon sources (5). Understanding global gene expression profiles in different carbon sources is important to the investigation of E. coli growth in natural environment, where the availability of carbon sources changes dynamically. Acetate-metabolizing culture is particularly relevant to the biotechnology industry, since the accumulation of acetate in bioreactor is commonly observed and often poses as an obstacle to high cell density cultivation (6).

The recent advent of microarray technology allows a thorough analysis of gene expression patterns in different environmental conditions (7, 8). In this approach, individual DNA probes are arrayed on a small glass surface, and labeled first strand cDNA from specific tissue or cell sources is hybridized onto the array. The amount of fluorescence at each DNA probe spot correlates with the abundance of specific mRNA transcript in the cell. This approach enables the characterization of transcriptionally regulated pathways at a genomic scale. In particular, the genome of E. coli has been arrayed and used for the comprehensive analysis of the expression level in various physiological states (9–21).

In this paper, we investigated the transcript profile of E. coli grown in acetate as the sole carbon source and compared it with the culture grown in glucose. We employed the fluorescence-based microarray system by spotting 96% E. coli open reading frames on glass slides. The expression levels of each gene were monitored by fluorescence-labeled cDNA using the printed DNA microarray. A rigorous statistical method was used to access the confidence interval of expression ratio of each gene by taking into account slide-to-slide and culture-to-culture variations. The gene expression profile in acetate was used to assess the metabolic flux distribution in key pathways.

EXPERIMENTAL PROCEDURES
Preparation of DNA Array—To make the E. coli cDNA microarrays, PCRs were performed in 96-well plates using Genosys E. coli ORFmers (Sigma) as primers and E. coli MG1655 (E. coli Genetic Stock Center, Yale University) chromosome as a template. Eppendorf Mastercycler Kit (Westbury, NY) was used for a 100-μl volume PCR. After PCR, 3 μl of PCR solutions were run in 1% (w/v) agarose gel to examine the quality of PCR products. Among 4290 primer pairs, 161 pairs failed to yield the desired PCR products. The remaining 4129 PCR products representing 96% of the E. coli open reading frames were precipitated by mixing with 10 μl of 3 M sodium acetate (pH 5.3) and 66 μl of isopropanol alcohol and centrifuging at 4000 rpm for 45 min. The precipitants were dissolved in 10 μl of 350 mM sodium bicarbonate/carbonate buffer (pH 9.0) and printed on a poly-l-lysine-coated glass slide using a robotic spotter. The diameter of each spot was about 150 μm, and the distance between the centers of the spots was about 250 μm.

The slide was hydrated over 95 °C water for 5 s and snap dried on a 100 °C heating block. The probes were cross-linked to the slide by UV light using the Stratalinker (Stratagene, La Jolla, CA) at 400 mJ. The free lysine groups on the slide surface were blocked by soaking slides in the mixture of 315 ml of m-methylpyrrolidinone with 5 g of succinic anhydride and 35 ml of sodium borate solution (0.2 M, pH 8.0) for 15 min. The slides were then washed with 95 °C water for 2 min.
and transferred to 95% ethanol at room temperature for 1 min. The slides were dried by centrifugation (22).

Strain and Culture Conditions—E. coli MC4100 (F araD139 (argF- lac) U169 rpsL150 relA1 hisS301 deoC1 ptsP25 rbsR) was cultured in M9 minimal medium (23) containing either 0.5% (w/v) glucose or 0.25% (w/v) sodium succinate as carbon sources for exponential growth phase (A600 = 0.3–0.8) in M9 minimal medium with glucose or acetate. Fluorescence was excited at a wavelength of 395 nm, and emission was measured at a wavelength of 509 nm. The fluorescence intensity for each culture was normalized by dividing the fluorescence by cell density at A600.

Deletion Mutation of pckA, ppsA, sfcA, and maeB—Each gene was disrupted by the method developed by Datsenko and Wanner (26). Briefly, primers (listed in Table I) for deletion were used to amplify the chloramphenicol-resistant gene from pKD3 (26). Strains harboring pKD46 (26) were grown in SOB medium containing 200 mg/liter ampicillin and 1 mML -arabinose and transformed with the PCR products using an electroporator (Eppendorf). Chloramphenicol-resistant strains were selected on agar plate, and the chloramphenicol-resistant gene was eliminated from the strain by transforming pCP20 (26) and colony-purified at 42 °C. The chloramphenicol-resistant gene and pCP20 were popped out during this process. The mutation was confirmed by PCR with primers for confirmation (Table I). Multiple deletion mutations were performed sequentially.

RESULTS

Transcript Profiling Using DNA Microarray—E. coli transcript profiles in acetate and glucose minimal media were compared using six sets of DNA microarray data generated from three independent experiments. To achieve a balanced physiological state, the E. coli strain was subcultured at least twice after initial inoculation from an LB agar plate and harvested in the midlog phase. Total RNA was purified and labeled with Cy3 or Cy5 dCTP during reverse transcription. In the first two experiments, we labeled the RNA from the glucose culture with Cy5 dyes and the RNA from the acetate culture with Cy3. The labels were reversed in the third experiment. In each experiment, the labeled cDNAs were mixed and then hybridized on two microarray slides. For calibration experiments, RNA from glucose culture was divided into two tubes and labeled with Cy3 and Cy5 dCTP, respectively, and hybridized to two slides. The same process was repeated with RNA purified from an acetate culture. Four data sets of calibration experiments were used to provide the statistical parameter for slide-dependent variation of each gene. This information was incorporated to evaluate the confidence intervals in the MCMC method.

The arraying solutions without DNA were arrayed 24 times on each slide in four different positions and used as negative controls. The average plus two S.D. values of intensities of the negative controls was computed in every slide and used as a threshold for data filtering. The spots showing intensity lower than the threshold were filtered out. These spots were attributed to either improper probe arraying or low expression. The spots whose intensities exceeded the detection range of the scanner were also excluded. After filtering the outliers, the rank-invariant genes were selected as a basis for normalization, and
then the logarithmic ratio of expression levels were calculated as described above. The four sets of calibration data and six sets of acetate/glucose comparison data were used to access the expression ratios using the MCMC method. The expression levels of 3649 genes, which passed the threshold filtering, were computed successfully. Among them, 354 genes were up-regulated, and 370 genes were down-regulated with 95% confidence in acetate medium compared with those in glucose medium. Because of the high number of simultaneous statistical tests used, the error rate at 5% for 4000 genes will generate 200 false positives when discussing the genome as a whole. Therefore, a more stringent confidence level was also used. With 99% confidence, 185 and 177 genes were up- and down-regulated, respectively. At this confidence level, the error rate of 1% will produce 40 false positives. However, when discussing individual genes, a 95% confidence level was appropriate for most purposes discussed here unless specified otherwise.

In Fig. 1, the logarithmic ratio of each gene was plotted against the average log intensity of that gene. The red and green dots represent the genes that were up- or down-regulated, respectively, with 95% or 99% confidence level, whereas the black dots represent the genes that were not differentially regulated based on the statistical judgment. Fig. 1 shows that the -fold difference does not necessarily correlate with statistical significance, as previously pointed out (10). For example, rpmF was down-regulated more than 4-fold in acetate, but it was statistically insignificant because of its high variances of the data. Meanwhile, the dnaN level was repressed only 1.3-fold with very small variance, and thus this gene was determined to be down-regulated with a 99% confidence level. The source of the variance can be severalfold, including poor quality or quantity of PCR products, cross-contamination by nonspecific binding, and high background on the slides. Therefore, a vigorous statistical analysis and repeats of the experiment are essential to obtain reliable data. The expression ratios and confidence intervals of all the genes are available on the World Wide Web at www.seas.ucla.edu/~liao/glecac.html.

### Functional Groups of Differentially Regulated Genes

The differently expressed genes with 95 and 99% confidence intervals were further classified into 23 groups by their functions (Table II). As expected, the most significant difference of growing in different carbon sources occurred among the central intermediary metabolic genes. 27 and 14% of the 161 genes in this group were significantly induced or repressed, respectively, in acetate medium compared with those in glucose medium. In the functional groups, such as carbon compound...
catabolism and fatty acid metabolism, the numbers of up-regulated genes surpassed those of down-regulated genes significantly, suggesting that these groups of genes were generally repressed by glucose or induced by acetate. Apparently, E. coli up-regulates the genes that allow the utilization of any possible carbon sources such as fatty acids when growing in a relatively poor carbon source such as acetate.

On the other hand, the number of down-regulated genes was much higher than up-regulated genes in many functional groups. For example, amino acid biosynthesis and nucleotide biosynthesis genes were generally down-regulated. The number of repressed genes exceeds the induced genes by 25 to 9 and 10 to 3 in these two groups, respectively, suggesting that E. coli turned down the expression of biosynthetic genes to match the low growth rate and save energy in the poor carbon source. In addition, in the functional groups such as cell structure, DNA replication, transcription, and translation, the number of down-regulated genes surpasses the up-regulated genes by 21 to 7, 22 to 4, 9 to 3, and 47 to 2, respectively. This result can be explained by the fact that the growth rate was much lower in acetate compared with glucose as a carbon source. Indeed, many of the genes belonging to these categories were known to be correlated with the growth rate of the cell (27).

Central Metabolic Genes Involved in Acetate Metabolism—When acetate is metabolized as a sole carbon and energy source in E. coli, it is first activated to acetyl-CoA and then metabolized in the tricarboxylic acid cycle and glyoxylate shunt (4). Two pathways were responsible for the acetate activation, pta-ackA and ackA. Interestingly, pta and ackA were both down-regulated by ~2-fold in acetate. On the other hand, ackA was induced more than 8-fold in acetate, one of the most significantly up-regulated genes. This result suggests that Acetate-CoA is the major enzyme for acetate uptake and activation. Acetyl-CoA is converted to malate through the glyoxylate shunt. The induction mechanism of the glyoxylate shunt genes, aceB and aceA, in acetate has been well characterized (4). These two genes are located in the same operon with aceK (encoding isocitrate dehydrogenase kinase/phosphatase). The expression of aceB gene was not monitored in this experiment because of failed PCR amplification. However, the other genes (aceA and aceK) in the same operon showed more than 10-fold up-regulation. Not only the aceBAK operon but also the gldDGB operon was induced significantly. The last gene in the gld operon is the secondary malate synthase (gldC), which can replace the malate synthase A (aceB) in acetate (28) when aceB is mutated. This operon could be induced by acetate (29). Meanwhile, most of the tri-carboxylic acid cycle genes (glta, acnA, acnB, icdA, sucA, BCD operon, sdhCDAB operon, fumA, fumC, fumB, and mdh) were up-regulated. In particular, the genes involved in the glyoxylate pathway, mdh, gltA, and acnB, were highly up-regulated more than 4-fold. Together with induction of aceBAK operon, these results confirmed that metabolic flux in the glyoxylate cycle (4) is very high in acetate (Fig. 2 and Table III).

The gluconogenic enzyme, phosphoenolpyruvate carboxykinase (coded by pckA), is known to be responsible for delivering the carbon flux from the tricarboxylic acid cycle to the gluconogenic pathways in acetate (4, 5). Indeed, pckA was 5–14-fold induced in acetate. Surprisingly, ppsA (coded for phosphoenolpyruvate synthase) was also induced 9–21-fold in acetate, although the gene product was nonessential for gluconeogenesis during growth on acetate (5). This result suggested that PpsA may play an important role for gluconeogenic flux when E. coli is grown in acetate. Together with the malic enzymes, PspA could serve the same function as PckA. Indeed, both NAD-dependent (sfcA) and NADP-dependent (maeB) malic enzymes were also induced. The functions of these gluconeogenic genes in acetate were further verified with deletion mutations, which will be discussed shortly.

Other Carbon and Energy Metabolism Genes—Many glycolytic genes (pfkB, fbo, gapA, epd, pgk, eno, pykF, and ppc) were down-regulated in acetate. In addition, the first two pentose pathway genes (zuf and gnd) and pyruvate dehydrogenase (aceEF operon) were also significantly down-regulated. These data correlate with the reduced metabolic flux in these pathway genes. The expression levels of many carbon transport genes were also affected seriously by different carbon sources. The genes involved in glucose transport, the ptsH-crR operon and ptsG, were repressed significantly in acetate, 1.5–2-fold and 2–4-fold, respectively. These operons were known to be

| Functional groups | Up 95% confidence | Down 95% confidence | Total no. of genes |
|-------------------|------------------|---------------------|-------------------|
| Amino acid biosynthesis and metabolism | 9 (7.8%) | 28 (24.5%) | 37 (3.1%) |
| Biosynthesis of cofactors, prosthetic groups, and carriers | 9 (10.0%) | 8 (8.9%) | 17 (1.4%) |
| Carbon compound catabolism | 22 (18.8%) | 4 (3.4%) | 26 (2.1%) |
| Cell processes (including adaptation, protection) | 6 (3.7%) | 13 (8.0%) | 19 (1.5%) |
| Cell structure | 7 (4.0%) | 21 (11.9%) | 28 (2.2%) |
| Central intermediary metabolism | 43 (26.7%) | 23 (14.3%) | 66 (5.2%) |
| DNA replication, recombination, modification, and repair | 4 (3.7%) | 22 (20.2%) | 26 (2.1%) |
| Energy metabolism | 30 (14.1%) | 24 (11.3%) | 54 (4.3%) |
| Fatty acid and phospholipid metabolism | 10 (20.5%) | 3 (6.3%) | 13 (1.0%) |
| Hypothetical, unclassified, unknown | 127 (10.0%) | 85 (6.7%) | 212 (16.7%) |
| Membrane proteins | 9 (0%) | 1 (0.3%) | 10 (0.8%) |
| Nucleotide biosynthesis and metabolism | 3 (6.3%) | 10 (20.8%) | 13 (1.0%) |
| Other known genes | 1 (4.2%) | 2 (4.2%) | 3 (0.2%) |
| Phage, transposon, or plasmid | 14 (17.7%) | 2 (2.5%) | 16 (1.3%) |
| Putative chaperones | 0 (0%) | 1 (0.3%) | 1 (0.1%) |
| Putative enzymes | 28 (12.6%) | 15 (6.8%) | 43 (3.4%) |
| Putative regulatory proteins | 4 (3.3%) | 4 (3.3%) | 8 (0.6%) |
| Putative transport proteins | 4 (3.0%) | 12 (9.1%) | 16 (1.3%) |
| Regulatory function | 1 (2.3%) | 6 (14.0%) | 7 (0.6%) |
| Structural proteins | 1 (2.5%) | 4 (10.0%) | 5 (0.4%) |
| Transcription, RNA processing, and degradation | 3 (5.8%) | 9 (17.3%) | 12 (0.9%) |
| Translation, post-translational modification | 3 (1.3%) | 47 (51.1%) | 50 (4.0%) |
| Transport and binding proteins | 26 (10.0%) | 28 (10.8%) | 54 (4.3%) |
| Total | 354 (9.7%) | 370 (10.1%) | 724 (5.7%) |

Global Expression Profiling of Acetate-grown E. coli

Table II
Numbers and percentages of up- or down-regulated genes in each functional group

Annotations follow Blattner et al (50).
regulated by Mlc (30). In the absence of glucose, Mlc represses the operons. In the presence of extracellular glucose, the conformation of EIIBCglc protein is changed and bound strongly with Mlc, which no longer represses the operons. On the other hand, transport genes for other carbon sources were induced significantly by catabolite derepression. Examples are the galactose ABC transporter operon (mglBAC, 4–8-fold), the ribose uptake gene operon (rbsD and rbsABC operon, 2–8-fold and 3–5-fold, respectively), the N-acetyl-D-glucosamine transport subunit (nagE, 2–4-fold), the arginine ABC transport gene (argT, 2–6-fold), the C4 dicarboxylate transporter gene (dctA, 3–6-fold) (31), tagatose metaoblic genes (gatYZ operon, 3–10-fold), and the maltose translocating gene (lkmB, 2–4-fold).

Interestingly, not only the genes involved in glyoxylate shunt pathway (ace and glc operons), but also those involved in other glyoxylate-related metabolic pathways, such as glycolate and allatoin metabolism, were all up-regulated (Fig. 3 and Table IV). These genes are located close together on the chromosome and expressed by four different operons (32). However, the role of these genes in acetate metabolism is unknown.

Genes Involved in the Cell Machinery—Among the genes involved in the cell structure, DNA replication, transcription, and translation, a total of 16 genes (3.3% of the total) were up-regulated, whereas 99 genes (20%) were down-regulated. Generally, the expression data of these groups of genes do not vary much from experiment to experiment compared with those of metabolic genes. Therefore, the accuracy of measurements was higher. Among 60 ribosomal proteins, including S1–S21, L1–L25, L27–L36, two EF-Tu subunits, EF-Ts, and EF-G, the expression levels of 40 genes were successfully monitored. Among them, more than 70% (29 of 40) were down-regulated in acetate at the 95% confidence level. The down-regulation of these genes was attributed to the growth rate-dependent regulation (27).

Roles of Gluconeogenic Genes in Acetate Growth—Although ppsA was known to be nonessential for growth in acetate, it was up-regulated 9–21-fold in acetate (Fig. 2 and Table III). To verify the induction of ppsA transcript in acetate, we constructed a transcriptional fusion between the ppsA promoter and the green fluorescence protein. The E. coli MC4100 harboring the reporter plasmid (pKK1) was cultured in either glucose or acetate minimal media, and the fluorescence level
Global Expression Profiling of Acetate-grown *E. coli*

The -fold changes of central metabolic gene expression in acetate compared with glucose

The expected transcript level changes in acetate were between upper and lower bounds when the microarray data were analyzed with 95% and 99% confidence intervals using the MCMC method described under "Experimental Procedures." The -fold changes of operon are the averages of the -fold changes of the genes belonging to the operon.

| Gene or operon names | 95% confidence | 99% confidence | Gene functions                                      |
|---------------------|----------------|----------------|----------------------------------------------------|
|                      | Lower bound | Upper bound    | Lower bound | Upper bound |
| ptsH-cri                  | 0.37         | 0.7            | 0.33        | 0.78        | PTS family Hpr, enzyme I and II |
| ptsG                  | 0.21         | 0.44           | 0.17        | 0.52        | PTS family enzyme BC (glucose) |
| gdh                  | 0.74         | 1.6            | 0.61        | 2.1         | Gluconokinase                  |
| pgi                  | 0.57         | 1.4            | 0.41        | 1.7         | Glucosephosphate isomerase     |
| pfpA                | 0.45         | 0.78           | 0.38        | 0.96        | 6-Phosphofructokinase I        |
| pfpB                | 0.31         | 1.9            | 0.65        | 2.2         | 6-Phosphofructokinase II       |
| fbp                  | 2.5          | 4.3            | 3.1         | 5.2         | Fructose-bisphosphatase        |
| fda                 | 0.4          | 0.62           | 0.33        | 0.76        | Fructose-bisphosphate aldolase |
| tpiA                | 0.65         | 1.4            | 0.54        | 1.4         | Triose-phosphate isomerase     |
| gapA                 | 0.3          | 0.69           | 0.24        | 1.0         | Glyceraldehyde-3-phosphate     |
|                      |              |                |            |             | dehydrogenase A                |
| epd                  | 0.15         | 0.31           | 0.14        | 0.43        | β-Erythrose-4-phosphate        |
|                      |              |                |            |             | dehydrogenase                   |
| pgk                  | 0.46         | 0.77           | 0.41        | 0.85        | Phosphoglycerate kinase        |
| gpmA                | 0.84         | 4              | 0.35        | 6.6         | Phosphoglyceromutase 1          |
| gpmB                | 0.78         | 5.1            | 0.75        | 6.8         | Phosphoglyceromutase 2          |
| eno                  | 0.34         | 0.96           | 0.18        | 0.99        | Enolase                          |
| pykA                | 0.98         | 1.6            | 0.96        | 1.8         | Pyruvate kinase II              |
| pykF                | 0.19         | 0.52           | 0.14        | 0.74        | Pyruvate kinase 1               |
| ppsA                | 8.6          | 21.0           | 6.6         | 24.7        | Phosphoenolpyruvate synthase    |
| ppc                 | 0.17         | 0.45           | 0.24        | 0.53        | Phosphoenolpyruvate carboxylase |
| pckA                | 4.9          | 14.4           | 4.3         | 15.9        | Phosphoenolpyruvate carboxykinase |
| aceEF               | 0.27         | 0.64           | 0.24        | 0.77        | Pyruvate dehydrogenase          |
| pta-ackA            | 0.41         | 0.82           | 0.34        | 0.96        | Phosphotransacylase and acetate |
|                      |              |                |            |             | kinase A                        |
| acs                  | 7.8          | 11.6           | 8.6         | 12.6        | Acetyl-CoA synthetase           |
| gltA                 | 3.5          | 6.8            | 2.9         | 9.0         | Citrate synthase                |
| acnA                | 1.1          | 2.1            | 0.95        | 2.4         | Aconitate hydratase 1           |
| acnB                | 3.6          | 13.1           | 2.1         | 13.8        | Aconitate hydratase 2           |
| icdA                | 1.3          | 2.5            | 1.2         | 3.2         | Isocitrate dehydrogenase        |
| sucAB               | 1.2          | 3.2            | 0.89        | 4.1         | 2-Oxoglutarate dehydrogenase    |
| succCD              | 1.8          | 4.7            | 1.4         | 5.3         | Succinyl-CoA synthetase         |
| sdhCDAB             | 1.2          | 2.5            | 0.88        | 3.4         | Succinate dehydrogenase         |
| fabABCD             | 0.77         | 2.2            | 0.57        | 2.6         | Fumarate reductase              |
| fumA                | 2.2          | 5.7            | 2.0         | 8.1         | Fumarase A                      |
| fumB                | 1.2          | 2.7            | 0.9         | 3.1         | Fumarase B                      |
| fumC                | 1.3          | 3.5            | 0.78        | 3.6         | Fumarase C                      |
| mdh                 | 2.4          | 6.2            | 2.0         | 9.3         | Malate dehydrogenase, NAD(P)‐binding |
| aceAK               | 17.4         | 32.8           | 16.1        | 36.2        | Isocitrate lyase and icdA kinase/phosphatase |
| sfcA                | 1.2          | 2.4            | 0.92        | 2.7         | NAD-linked malate dehydrogenase |
| maeB                | 4.2          | 6.4            | 4.1         | 7.4         | Putative malate dehydrogenase   |
| zwf                 | 0.41         | 0.77           | 0.3         | 0.99        | Glucose-6-phosphate dehydrogenase |
| gnd                 | 0.43         | 0.84           | 0.41        | 0.9         | Gluconate-6-phosphate dehydrogenase |
| rpe                 | 0.78         | 1.4            | 0.68        | 1.5         | β-Ribulose-5-phosphate 3-epimerase |
| rpiA                | 0.71         | 1.4            | 0.71        | 1.8         | Ribosephosphate isomerase       |
| rpiB                | 0.8          | 1.4            | 0.57        | 2.2         | Ribose 5-phosphate isomerase B  |
| tnaA                | 1.1          | 5.2            | 1.0         | 7.0         | Transaldolase A                 |
| tnaB                 | 0.69        | 1.3            | 0.57        | 1.5         | Transaldolase B                 |
| tktA                | 0.95        | 1.9            | 0.78        | 2.5         | Transketolase 1                 |
| tktB                | 0.84        | 2.1            | 0.67        | 2.9         | Transketolase 2                 |

was monitored in the midlog phase. In two independent experiments, the fluorescence level of the cells grown in acetate was detected at least 4-fold higher than the level in glucose (Fig. 4). This result supports the data obtained by DNA microarray analysis. Although previously undetermined, the role of PpsA in acetate-grown *E. coli* may be the delivery of metabolites from the tricarboxylic acid cycle to the Embden-Meyerhoff pathway. This gluconeogenic function requires the malic enzymes (coded by *sfcA* and *maeB*), which were also up-regulated in acetate. The malic enzyme-PNPH pathway can theoretically serve as an alternative to PckA, which is known to be the main gluconeogenic pathway in acetate-grown *E. coli* and is also up-regulated significantly. To determine the roles of these genes in *E. coli* growth in acetate, a set of deletion mutants was constructed as described. The deletions of the genes on the chromosome were confirmed by PCR using the primers outside of the genes. The growth rates of those mutants were measured in acetate (Table V). Almost no growth inhibition was detected in either *pckA* or *ppsA* single mutants, while the growth of the *pckA* *ppsA* double mutant was abolished in acetate (Table V). These results showed that PckA and PpsA-malic enzyme can serve as an alternative pathway to each other and that these are the only two gluconeogenic enzymes to provide the phosphoenolpyruvate pool. We further examined the functions of the malic enzymes, which are required to supply pyruvate, the substrate of PpsA. The deletion mutation of one malic enzyme, either *sfca* or *maeB*, in the *pckA* background did not reduce the growth rate in acetate. However, the deletion of both malic enzymes with *pckA* mutation abolished the growth in acetate (Table V). The series of mutation studies proved that PckA and PpsA-malic enzymes are the two substitutable pathways that can deliver carbon flux from the tricarboxylic acid cycle to the Embden-Meyerhoff pathway.
why the pta ackA mutant did not grow well in acetate.

Among the genes that are highly up-regulated in acetate, ppsA was unexpected, since this gene was dispensable during growth on acetate. The ppsA induction was confirmed independently by a promoter fusion experiment. The role of ppsA during growth on acetate was demonstrated using mutation analysis. In acetate, PpsA and the two malic enzymes form an alternative pathway to PckA. Deletion mutation of either one of the pathways has no effect on growth rate in acetate, suggesting that either one has sufficient capacity to deliver carbon flux to the Embden-Meyerhoff pathway. The malic enzyme pathway was not clearly defined in E. coli previously. It has been known for more than 30 years that two malic enzymes existed in E. coli (36). However, the genes coding those enzymes were not mapped on the genetic level until recently. NAD-dependent malic enzyme, sfcA, was cloned and characterized recently (37), whereas NADP-dependent malic enzyme, maeB, was only predicted with sequence similarity. Interestingly, both sfcA and maeB were highly up-regulated in acetate. In addition, mutation analysis showed that these genes can compensate each other in acetate cultures.

The most striking difference between acetate-grown and glucose-grown E. coli cultures is the up-regulation of metabolic genes such as acs, pckA, ppsA, and aceBAK. These genes are either essential or carry significant metabolic flux. In general, results from the microarray analysis summarized in Fig. 2 show a dramatic agreement with the deduced metabolic pathway for acetate growth. All of the essential or flux-carrying pathways (gluoxylate shunt, tricarboxylic acid cycle) were up-regulated, and most of the unused pathways (glycolysis) were down-regulated. It is tempting to suggest that the significantly up-regulated genes delineate the functional pathways. The PpsA-malic enzymes pathway, the PckA pathway, and the Acs pathway are good supports for this argument. If this theory holds, one can predict functional pathways through gene expression profiling.

Regulatory Systems—Although many global regulators have been characterized, it is not straightforward to identify the regulators involved in the change of expression profiles for glucose to acetate cultures. For example, the tricarboxylic acid cycle genes were up-regulated in acetate partly by ArcA and Fnr (38). However, the effects of those regulators were not observed in other ArcA or Fnr-regulated genes, such as cydAB, cyoABCDE, dmsABC, and sodA (38). It is possible that multiple regulators co-control these genes to manifest a complex response to the environment. The only regulatory system showing ubiquitous effects on many different genes is the catabolite derepression by cAMP-CAMP receptor protein. This regulator activates many genes in acetate. For example, the tricarboxylic acid cycle and gluconeogenic pathway and many carbon transport genes were fully or partially induced by catabolite derepression. Among the biosynthetic genes, which were generally down-regulated, a few genes, such as tnaA, ilvBN, dadX, and aspA, were strongly up-regulated, possibly attributable to this mechanism (39–41). Catabolite derepression also induces fatty acid degradation genes, fadD and fadBA (42), and flagella regulatory genes (flhCD) (43) in acetate growth. Finally, the 11–20-fold and 3–5-fold inductions of catA and asaA, of which the functions are not clear, were also attributed to this mechanism (44, 45).

Global Expression Profiling of Acetate-grown E. coli

### Figure 3

**The expression profiles of gluoxylate metabolic genes in acetate compared with in glucose.** The numbers, arrows, and abbreviations are as described in the legend to Fig. 2.

**Discussion**

Statistics—Although DNA microarray has gained momentum in genomic scale expression profiling, methods for assessing statistical significance have just begun to be developed. The initial approach of assigning genes that shows a more than 2-fold expression ratio as “significant” has proven unsatisfactory (10). Because the variations among the signals on the array are gene-dependent, a common threshold for all of the genes is inadequate. Our data here again show the same conclusion. Long et al. (33) proposed a t test using the Bayesian inference of variance. This approach was developed for the one-channel membrane arrays and used the intensity data of the two treatments as variables for comparison. This method is not optimal for the two-channel glass arrays, which allow the direct comparison of two treatments (acetate and glucose) simultaneously on the same slide. Such two-channel systems reduce the slide-to-slide variation, which can be significant in both glass and membrane arrays. The method used here (24) directly compares the two treatments by taking the log ratio of the normalized intensity. The gene-dependent, slide-to-slide and experiment-to-experiment variations were accounted for in a hierarchical Bayesian model. The gene-specific confidence intervals as a result of multiple experiments and multiple slides were calculated. Because of the large number of statistical tests, we calculated both 95 and 99% confidence levels. At these two confidence levels, about 200 and 40 false positives, respectively, will occur among 4000 genes, when the whole genome is considered simultaneously.

Metabolic Genes—In general, the expression profiles obtained from E. coli grown in glucose and acetate agree with the direction of intracellular carbon fluxes. In acetate cultures, the phosphotransferase systems for glucose uptake and glycolytic genes were highly down-regulated compared with those in glucose cultures. On the other hand, genes involved in acetate uptake (acs), the gluoxylate cycle (aceBAK and gleB), the tricarboxylic acid cycle (gltA, iedA, acnA, acnB, sucABC, sdhC-DAB, fumA, fumB, fumC, and mdh), and gluconeogenesis (pckA, ppsA, sfcA, and maeB) were all up-regulated in acetate. The only exception is the Pta-AckA pathway, which was down-regulated significantly in acetate. This result suggests that Acs is the main path for acetate uptake, whereas the Pta-AckA pathway is used for acetate excretion during growth on glucose. Because the mutation of both pta and ackA inhibited cell growth in acetate (34), it has been suspected that this pathway also delivers a significant amount of carbon flux into the cell. However, acs induction in acetate has been shown to be impaired in the pta ackA double mutant (35), which may explain
Global Expression Profiling of Acetate-grown E. coli

The fold changes of glycolate metabolic gene expression in acetate compared with glucose

The expected transcript level changes in acetate were between upper and lower bounds when the microarray data were analyzed with 95 and 99% confidence intervals using the MCMC method described under “Experimental Procedures.”

| Gene name | 95% confidence | 99% confidence |
|-----------|----------------|----------------|
|            | Lower bound   | Upper bound   | Lower bound | Upper bound |
| gcE        | 3.0           | 28.3          | 1.0         | 49.1        |
| gcR        | 2.3           | 14.9          | 2.1         | 16.9        |
| glsK       | 1.2           | 3.0           | 0.9         | 4.2         |
| fucO       | 1.6           | 4.4           | 1.3         | 4.8         |
| aldB       | 6.2           | 20.4          | 4.9         | 29.2        |
| allA       | 1.9           | 8.7           | 1.2         | 12.7        |
| allB       | 3.5           | 40.3          | 2.0         | 95.7        |
| allC       | 1.6           | 2.9           | 1.5         | 3.6         |
| allD       | 1.3           | 4.1           | 1.5         | 5.5         |
| allR       | 1.6           | 2.5           | 1.4         | 3.0         |

Table IV

The expected transcript level changes in acetate were between upper and lower bounds when the microarray data were analyzed with 95 and 99% confidence intervals using the MCMC method described under “Experimental Procedures.”

The expression levels of green fluorescence protein fused to the ppsA promoter in E. coli grown in glucose and acetate minimal medium. The y axis is the relative fluorescence level normalized to cell density and the average fluorescence level in glucose.

Figure 4

The expression levels of green fluorescence protein fused to the ppsA promoter in E. coli grown in glucose and acetate minimal medium. The y axis is the relative fluorescence level normalized to cell density and the average fluorescence level in glucose.

TABLE V

Growth rates of the mutants of gluconogenic pathway genes in acetate

The strains were grown in acetate minimal medium, and their growth rates were calculated from the cell density measured every hour.

| Strains | Growth rate (h⁻¹) |
|---------|------------------|
| MC4100  | 0.33 ± 0.05      |
| MC4100 ppsA | 0.26 ± 0.02    |
| MC4100 yppA | 0.32 ± 0.02    |
| MC4100 ppsA ΔppA | 0.0         |
| MC4100 ΔpsA | 0.29 ± 0.03    |
| MC4100 ΔpsA ΔpsA | 0.29 ± 0.02 |

FIG. 4. The expression levels of green fluorescence protein fused to the ppsA promoter in E. coli grown in glucose and acetate minimal medium. The y axis is the relative fluorescence level normalized to cell density and the average fluorescence level in glucose.

Acetate Shock—The effect of acetate shock induced by high concentrations of sodium acetate (100 mM) was previously studied with membrane arrays (13). Those experiments aimed to study the physiological response of E. coli to a sudden addition of a high concentration of acetate during growth in a glucose minimal medium. In contrast, we monitored the transcript profile of E. coli during balanced growth in an acetate medium with a nontoxic concentration (42 mM). As expected, the results from these experiments were not the same. Although there were some common phenomena, such as down-regulation of ribosomal proteins and up-regulation of katE, dhnA, talA, and pflB, most transcript profiles in this experiment were different from the acetate shock experiment. In particular, the effect of sE that was significant in acetate shock was not observed in acetate culture.

REFERENCES

1. Cronan, J. E., Jr., and Laporte, D. (1996) in Escherichia coli and Salmonella: Cellular and Molecular Biology (Neidhardt, F. C., Curtiss, R., III, Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., and Umbarger, H. E., eds) 2nd Ed., pp. 206–216, ASM Press, Washington, D. C.

2. Saier, M. H., J., Ramsier, T. M., and Reisser, J. (1996) in Escherichia coli and Salmonella: Cellular and Molecular Biology (Neidhardt, F. C., Curtiss, R., III, Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., and Umbarger, H. E., eds) pp. 1325–1343, ASM Press, Washington, D. C.

3. Postma, P. W., Lengeler, J. W., and Jacobson, G. R. (1996) in Escherichia coli and Salmonella: Cellular and Molecular Biology (Neidhardt, F. C., Curtiss, R., III, Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., and Umbarger, H. E., eds) pp. 1149–1174, ASM Press, Washington, D. C.

4. Cozzone, A. J. (1998) Annu. Rev. Microbiol. 52, 127–164

5. Holmes, H. (1996) FEMS Microbiol. Rev. 19, 85–116

6. Luli, G. W., and Strohl, W. R. (1990) Appl. Environ. Microbiol. 56, 1004–1011

7. Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. (1995) Science 270, 467–470

8. DellIsi, J., Iyer, V. R., and Brown, P. O. (1997) Science 278, 680–686

9. Richmond, C. S., Glasser, J. D., Mas, R., Jin, H., and Blattner, F. R. (1999) Nucleic Acids Res. 27, 3821–3835

10. Arfin, S. M., Long, A. D., Io, E. T., Tellier, L., Riehle, M. M., Puelle, E. S., and Hatfield, G. W. (2000) J. Biol. Chem. 275, 29672–29684

11. Khodursky, A. B., Peter, B. J., Cozzarelli, N. R., Botstein, D., Brown, P. O., and Yanofsky, C. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 12170–12175

12. Zimmer, D. P., Soupeine, E., Lee, H. L., Wendisch, V. P., Khodursky, A. B.,
Global Expression Profiling of Acetate-grown E. coli

13. DeLisa, M. P., Wu, C. F., Wang, L., Valdes, J. J., and Bentley, W. E. (2001) J. Bacteriol. 183, 4562–4570
14. Pruss, B. M., Liu, X., Hendrickson, W., and Matsumura, P. (2001) FEMS Microbiol. Lett. 197, 91–97
15. Zheng, M., Wang, X., Templeton, L. J., Smulski, D. R., LaRossa, R. A., and Baldom, L. (2001) J. Bacteriol. 183, 5239–5247
16. Peterso, P. J., Rennik, M. H., and Dempel, B. (2001) J. Bacteriol. 183, 3890–3902
17. Oh, M. K., and Liao, J. C. (2000) Biotechnol. Prog. 16, 278–286
18. Oh, M. K., and Liao, J. C. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 881–885
19. Fisher, J. P., Pansegraue, W., Frank, R., Blecker, H., Scholz, P., Bagdasarian, M., and Lanka, E. (1986) Gene (Amst.) 48, 119–131
20. Datsenko, K. A., and Wanner, B. L. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 6640–6645
21. Keener, J., and Nomura, M. (1996) in Escherichia coli and Salmonella: Cellular and Molecular Biology (Neidhardt, F. C., Curtiss, R., III, Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., and Umbarger, H. E., eds) 2nd Ed., pp. 123–407, ASM Press, Washington, D. C.
22. Verheugen, R., and De Vlieghere, M. (1968) Eur. J. Biochem. 5, 81–90
23. Pellicer, M. T., Fernandez, C., Badia, J., Aguilar, J., Lin, E. C., and Baldom, L. (1999) J. Biol. Chem. 274, 1745–1752
24. Davies, S. J., Golby, P., Omrani, D., Broad, S. A., Harrington, V. L., Guest, J. R., Kelly, J. D., and Andrews, S. (1999) J. Bacteriol. 181, 5624–5635
25. Cuss, E., Obradors, N., Baldoma, L., Badia, J., and Aguilar, J. (1999) J. Bacteriol. 181, 7479–7484
26. Long, A. D., Mangalam, H. J., Chan, B. Y., Tolleri, L., Hatfield, G. W., and Baldi, P. (2001) J. Bacteriol. 183, 19937–19944
27. Stols, L., and Donnelly, M. J. (1997) Appl. Environ. Microbiol. 63, 2695–2701
28. Lynch, A. S., and Lin, E. C. C. (1996) in Escherichia coli and Salmonella: Cellular and Molecular Biology (Neidhardt, F. C., Curtiss, R., III, Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., and Umbarger, H. E., eds) 2nd Ed., pp. 1526–1538, ASM Press, Washington, D. C.
29. Gong, F., and Yanofsky, C. (2001) J. Biol. Chem. 276, 1974–1983
30. Umbarger, H. E. (1996) in Escherichia coli and Salmonella: Cellular and Molecular Biology (Neidhardt, F. C., Curtiss, R., III, Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., and Umbarger, H. E., eds) 2nd Ed., pp. 442–457, ASM Press, Washington, D. C.
31. Reizer, J. (1996) in Escherichia coli and Salmonella: Cellular and Molecular Biology (Neidhardt, F. C., Curtiss, R., III, Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., and Umbarger, H. E., eds) 2nd Ed., pp. 123–145, ASM Press, Washington, D. C.
32. Schulz, J. E., and Matin, A. (1991) J. Mol. Biol. 218, 129–140
33. Fraley, C. D., Kim, J. H., McCann, M. P., and Matin, A. (1998) J. Bacteriol. 180, 4287–4296
34. Kakuda, H., Hosono, K., Shiroishi, K., and Ichihara, S. (1994) J. Biochem. (Tokyo) 116, 916–922
35. Kumari, S., Beatty, C. M., Browning, D. F., Bushy, S. J., Simel, E. J., Hovel-Miner, G., and Wolfe, A. J. (2000) J. Bacteriol. 182, 4173–4179
36. Murai, T., Tokushige, M., Nagai, J., and Katsuki, H. (1971) Biochem. Biophys. Res. Commun. 43, 875–881
37. Oh, M. K., and Liao, J. C. (2000) Biotechnol. Prog. 16, 278–286
38. Oh, M. K., and Liao, J. C. (2000) Metabol. Eng. 2, 201–209
39. Pruss, B. M., Liu, X., Hendrickson, W., and Matsumura, P. (2001) FEMS Microbiol. Lett. 197, 91–97
40. Reizer, L. J. (1996) in Escherichia coli and Salmonella: Cellular and Molecular Biology (Neidhardt, F. C., Curtiss, R., III, Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., and Umbarger, H. E., eds) 2nd Ed., pp. 123–145, ASM Press, Washington, D. C.