

Gene Expression Profiling by DNA Microarrays and Metabolic Fluxes in *Escherichia coli*

Min-Kyu Oh and James C. Liao*

Department of Chemical Engineering, University of California, Los Angeles, California 90095

DNA microarray technology was applied to detect differential transcription profiles of a subset of the *Escherichia coli* genome. A total of 111 *E. coli* genes, including those in central metabolism, key biosyntheses, and some regulatory functions, were cloned, amplified, and used as probes for detecting the level of transcripts. An *E. coli* strain was grown in glucose, acetate, and glycerol media, and the transcript levels of the selected genes were detected. Despite extensive studies on *E. coli* physiology, many new features were found in the regulation of these genes. For example, several genes were unexpectedly up-regulated, such as *pps*, *ilvG*, *aroF*, *secA*, and *dsbA* in acetate and *asnA* and *asnB* in glycerol, or down-regulated, such as *ackA*, *pta*, and *adhE* in acetate. These genes were regulated with no apparent reasons by unknown mechanisms. Meanwhile, many genes were regulated for apparent purposes but by unknown mechanisms. For example, the glucose transport genes (*ptsHI*, *ptsG*, *crr*) in both acetate and glycerol media were down-regulated, and the *ppc*, glycolytic, and biosynthetic genes in acetate were also down-regulated because of the reduced fluxes. However, their molecular mechanisms remain to be elucidated. Furthermore, a group of genes were regulated by known mechanisms, but the physiological roles of such regulation remain unclear. This group includes *pckA* and *aspA*, which are up-regulated in glycerol, and *gnd* and *aspA*, which are down- and up-regulated, respectively, in acetate. The DNA microarray technology demonstrated here is a powerful yet economical tool for characterizing gene regulation and will prove to be useful for strain improvement and bioprocess development.

Introduction

Understanding how genes work at the genomic scale is essential for biotechnological research and application. Achieving this goal involves genome sequencing and determining the role of each gene in the cell. Recent development of DNA or oligonucleotide microarray technology has accelerated the investigation of gene regulation. By hybridization with labeled mRNA, cDNA, or chromosomal DNA, arrayed PCR products or oligonucleotides on a substrate, such as membranes, glass slides, or silicon chips, have been successfully used for monitoring transcript levels (1), single nucleotide polymorphism (2), or genomic variations between different strains (3). One of the most significant applications of this technique is the gene expression profiling at the whole genomic scale. The expression levels of the whole 6400 genes in the *Saccharomyces cerevisiae* genome have been detected successfully with either DNA or oligonucleotide microarray technology (4–7). The technique has also been used to investigate physiological changes of human cells, such as the aging process (8), or the response of fibroblasts to serum stimulation (9).

The impact of this technology in biotechnology would be tremendous, because most problems in this field call for knowledge of global regulation, which can be studied efficiently using this technology. As illustrated in *S. cerevisiae*, cell metabolism adapts to the environment during growth in a chemostat culture (7). The transcript

levels of approximately 10% of the genes were changed more than 2-fold after 250 generations in the fermentor. In addition to analyzing the physiological changes, detecting the effects of gene structure changes using the microarray technique are expected to improve the speed of strain development (10).

Despite its potential, the application of microarray technology in *Escherichia coli*, one of the most common industrial microorganisms, has just begun (11). Because of the cost, we limit our attention to genes that are most relevant to metabolic engineering applications. In this work, we arrayed 111 *E. coli* genes, which include genes in glycolysis, pentose pathway, the tricarboxylic acid (TCA) cycle, key biosynthetic pathways, and selected regulatory and protein processing genes. These genes were cloned to a common vector, and amplified by polymerase chain reaction (PCR) using a pair of vector-specific primers before arraying onto glass slides. The differential transcription levels of these genes in a common *E. coli* strain (MC4100) were monitored during the exponential growth phase in a minimal medium using glucose, glycerol, or acetate as the carbon source. Results show many new regulatory features within the subset of the genome and support the utility of DNA microarrays for sub-genomic applications.

Materials and Methods

Strains. *E. coli* MC4100 (F⁻ *araD139* (*argF-lac*) U169 *rpsL150 relA1 flb5301 deoC1 ptsF25 rbsR*) was used for monitoring the transcription levels. XL1-Blue (Stratagene, La Jolla, CA) was used for cloning of *E. coli* genes

* Phone: (310) 825–1656. Fax: (310) 206-1642. Email: liaoj@ucla.edu.

in pBluescript II KS+ (Stratagene). The strain MC4100 was grown in β -galactosidase minimal medium (contained per liter: KH_2PO_4 , 5.44 g; K_2HPO_4 , 10.49 g; $(\text{NH}_4)_2\text{SO}_4$, 2 g; $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.05 g; $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 5 mg; $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 0.125 mg; CaCl_2 , 0.5 mg) with a concentration of 40 mM of the specified carbon source prepared as previously described (12).

Cloning and Arraying of *E. coli* Genes. A total of 111 genes (shown in Figure 3 and legend) in glycolysis, TCA cycle, pentose phosphate pathway, fermentation, heat shock regulon, and selected biosynthetic pathways were amplified from the *E. coli* chromosome with the predesigned *E. coli* PCR primer pairs (Genosys, The Woodlands, TX) or with custom-designed primers (synthesized by Genosys) by *taq* polymerase from Promega (Madison, WI). Unpurified *E. coli* chromosome from MC4100 was used as the template for PCR. The products were cloned, directly or after purification using QIAquick PCR purification kit (Qiagen, Valencia, CA), to pBluescript II KS+ (Stratagene) by the T-A cloning method (13). The T-A cloning vector was made by digesting pBluescript II KS+ with EcoRV followed by adding T to the blunt ends using *taq* polymerase and dTTP. Plasmids containing the desired insert were selected by blue/white screening on 5-bromo-4-chloro-3-indolyl- β -D-galactosidase (X-gal) plates and by checking plasmid sizes with the Quick-Check method (14). The resulting plasmids were used as templates for further PCR amplification of the inserted genes. The PCR was conducted in a volume of 100 μL with pBluescript II KS+ specific primers (Genosys), 5'-CAATTAACCCTCACTAAAGGGAACA-3' and 5'-GTAATACGACTCACTATAGGGCGAATTGG-3'. These amplified DNA fragments were checked by size using agarose gels, precipitated with ethanol, and dissolved in 15 μL of 350 mM sodium bicarbonate/carbonate buffer (pH = 9.0). The resulting gene probes were arrayed on silylated slides (PGC Scientific, Gaithersburg, MD) using a robotic arrayer according to the protocol described previously (15). Three genes, *fadA* and *mvaA* from *Archaeoglobus fulgidus* and cDNA of HMG-CoA synthase from *Blattella germanica*, were cloned in the same manner and arrayed as controls. The distances between spots were 0.4 mm, and each gene probe was duplicated on the same slide.

RNA Purification and Labeling. *E. coli* total RNA was purified from 1×10^9 cells by RNeasy Midi Kit from Qiagen. The purified RNA solution was incubated at 37 °C with 100 units of DNase (Gibco BRL, Rockville, MD) and 40 units of RNasin (Promega) for 30 min, extracted with phenol/chloroform, and then precipitated with ethanol. After being dissolved in 10–20 μL of RNase-free water, 80 μg of total RNA was subjected to labeling, which occurred during reverse transcription with 200 units Superscript RNase H⁻ reverse transcriptase (Gibco BRL). The mixture of 111 C-terminal *E. coli* primers were used as primers and 0.5 mM dATP, dTTP, dGTP, 0.2 mM dCTP, and 0.1 mM Cy-3 or Cy-5 dCTP (Amersham Pharmacia, Piscataway, NJ) were used as monomers for reverse transcription. After reverse transcription, RNA was degraded by incubation at 65 °C for 40 min with 10 μL of 1 N NaOH. The resulting labeled cDNA was diluted with 60 μL of TE buffer (pH = 8.0). The Cy-3 and Cy-5 labeled cDNA solutions were mixed and then concentrated to 1–2 μL by Micron-50 from Millipore (Bedford, MA).

Hybridization and Scanning. Before hybridization, the slide was hydrated over 55 °C water for 5 s, snap dried on a 100 °C heating block, and cross-linked by UV light using the Stratalinker (Stratagene, La Jolla, CA).

The slide was washed in 0.1% SDS solution and then in distilled water. The PCR products on the slide were denaturated in 95 °C water for 1 min, followed by quick desiccation in –20 °C ethanol, and air-dried. The slide was prehybridized with 10 μL of 50% formamide, 5X SSC, 0.1% SDS, 1X Denhardt's solution, and 0.2 mg/mL salmon sperm DNA.

The concentrated Cy-3 and Cy-5 cDNA was resuspended in 10 μL of 50% formamide, 3X SSC, 1% SDS, 5X Denhardt's solution, 6% dextran sulfate, 0.15 mg/mL salmon sperm DNA and 0.075 mg/mL yeast total RNA. The labeled cDNA was denaturated at 95 °C for 3 min and then quickly chilled on ice. The cDNA was placed on the slide and covered by HYBRISLIP (PGC Scientific). The slide was assembled in the hybridization chamber (Corning, Charlotte, NC), and hybridization occurred for 14–20 h at 42 °C. The hybridized slide was washed in 2X SSC, 0.1% SDS for 5 min at room temperature, then 0.2X SSC for 5 min prior to scanning.

Hybridized slides were dried by spinning and scanned with a custom-built two-color laser scanning fluorimeter, and the scanned images were analyzed with the program Image (Biodiscovery, Santa Monica, CA). The brightness of each spot was normalized by the ratio of total intensities of Cy-3 and Cy-5 signals.

Results and Discussion

Cloning of Metabolic Genes. A total of 111 *E. coli* genes involved in central catabolic pathways, main biosynthesis, and key regulatory mechanisms were chosen for studying differential gene expression in different carbon sources. Initially, the genes were PCR-amplified from the *E. coli* chromosome using commercially available primer pairs. However, about 10% of the primers did not yield the desired products, while others yielded uneven amounts of PCR products, which may skew the result of hybridization. To alleviate this problem and to reduce the future cost of PCR primers for individual genes, we cloned all of these genes to a common cloning vector, pBluescript II KS+ (Stratagene). To do so, we used the successful commercial primers and redesigned the unsuccessful primers to amplify these genes. After purification, these PCR products were ligated to the EcoRV site of the cloning vector. The genes were then reamplified from the plasmids using a pair of plasmid-specific primers. The resulting PCR products were uniform in quantity and quality. These PCR products were arrayed on glass slides for hybridization experiments.

Gene Expression Profiling of *E. coli* in Different Carbon Sources. The expression levels of metabolic genes were monitored when *E. coli* MC4100 was grown in minimal media with three different carbon sources, glucose, glycerol, and acetate. The glucose culture was used as a reference for comparison. The organism was cultured to the mid-exponential phase (OD 0.5–0.7 at 550 nm) in shake flasks at 37 °C for several generations, quickly chilled on an ethanol/dry ice bath, and centrifuged for RNA purification. The RNA of the cells cultured in glycerol and acetate medium were labeled with Cy-3 dCTP during reverse transcription, while RNA from the glucose culture was labeled with Cy-5 dCTP. Each Cy-3 labeled cDNA was cohybridized with Cy-5 labeled cDNA on the same slide. Images for the microarray were scanned at two different wavelengths, 570 nm for Cy-3 fluorescence and 670 nm for Cy-5. The ratio of Cy-3 and Cy-5 intensities for a spot was normalized with the ratio between the total Cy-3 and Cy-5 intensities on the slide. The final results represent the ratios of transcript

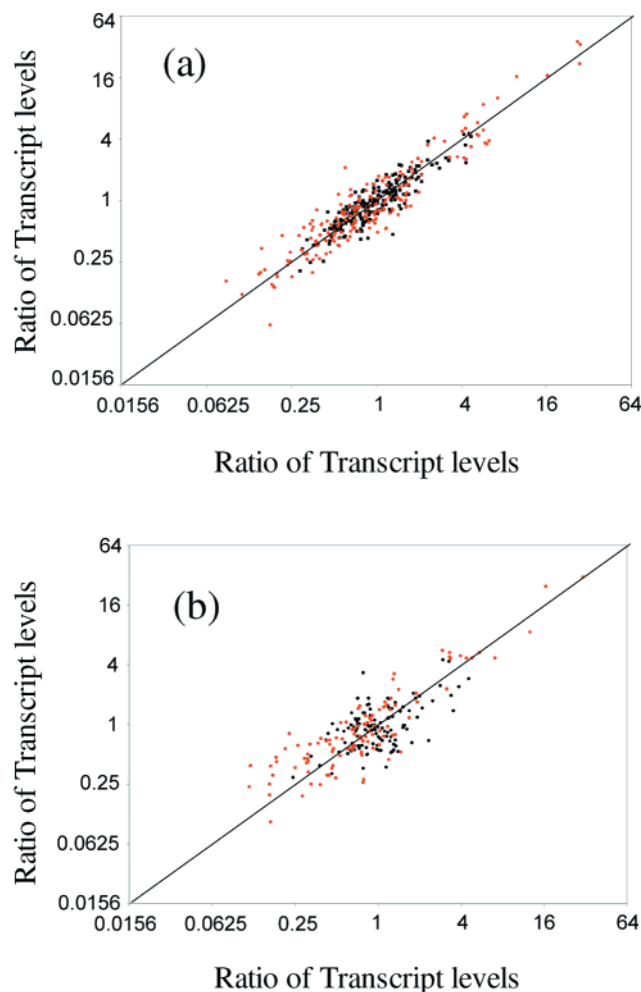


Figure 1. Comparison of the transcript ratios (a) detected by the duplicated spots on the same slide or (b) detected in different slides. The black dots are the results of the comparison between cells cultured in glycerol and glucose media, and the red dots are comparison between cells in acetate and glucose media. Each pair of data was plotted on the figure. If the data were completely reproducible, they will fall on the 45° line.

abundance in glycerol (or acetate) medium relative to the glucose medium.

Each experiment was repeated twice, and each slide had duplicated spots for each gene. Therefore four data sets were generated for each condition. Figure 1 shows comparisons of the repeated or duplicated experimental data. The consistency of these results was very good within the same slide (Figure 1a), and slightly degenerated between slides (Figure 1b). The four data sets were averaged geometrically to compare the expression levels of genes in different carbon sources.

The histograms of transcript level changes (Figure 2) demonstrated that in general the gene expression profile changed more significantly from glucose to acetate media and changed moderately from glucose to glycerol. Compared to the transcript abundance in glucose, 8 out of 111 genes in glycerol were up-regulated more than 2-fold, while 4 were down-regulated more than 2-fold. In comparison, in the acetate medium, 14 genes were up-regulated more than 2-fold, whereas 32 genes were down-regulated more than 2-fold. Overall, most biosynthetic genes were down-regulated in both glycerol and acetate media, while TCA cycle genes were up-regulated in both media (Figures 2a and 2b). The glycolytic genes were roughly unchanged in the glycerol medium but decreased in the acetate medium (Figures 2a and 2b).

Glycolysis. Most glycolytic genes were thought to be unregulated, with expression levels varying less than 2-fold (16). However, the results showed that the transcript abundance of several glycolytic genes (*pgi*, *pfkA*, *fba*, *tpiA*, *gapA*, *pgk*, *eno*, *pykF*, and *ppc*) were decreased 2- to 6-fold in the acetate medium compared to those in the glucose medium (Figure 3a). From a flux regulation standpoint, the decreases were not surprising, because these genes were not used as much in acetate. The first seven of these genes catalyze reversible reactions. Their down-regulation may reduce the flux in the gluconeogenic direction, which is much smaller than the glycolytic flux during growth in glucose. The down-regulation of *pykF*, but not *pykA* (Figure 3a), suggests that these two isoenzymes were differentially regulated. Indeed, *pykF*, but not *pykA*, has been suggested to be regulated by FruR (17), a global regulatory gene.

The *ppc* gene was down-regulated almost 3-fold in the acetate medium (Figure 3a). Although the regulation of this gene has not been reported, this result is consistent with the direction of the metabolic flux. During growth in acetate, the carbon flux is from oxaloacetate (OAA) to phosphoenolpyruvate (PEP), which is catalyzed by *pckA*. In fact, the *pckA* gene was up-regulated more than 4-fold. High-level expression of both *ppc* and *pckA* led to futile cycling (18), which results in an energy waste. It is thus consistent with down-regulated *ppc* during growth in acetate. The underlying molecular mechanism, however, remains to be determined.

In contrast, the glycolytic genes were not affected significantly during growth in glycerol. The only gene that was significantly affected was *pykA* (Figure 3c). When *E. coli* is cultured in glycerol, the flux from PEP to pyruvate is mediated by pyruvate kinases rather than the phosphotransferase system (PTS). Therefore, it is reasonable to expect that one of the pyruvate kinase genes is up-regulated to satisfy the greatly increased demand for pyruvate kinase. The *pykA* gene, rather than *pykF*, apparently serves this role. PykF is activated at the protein level by 1,6-fructose diphosphate (F1,6P), whose concentration is present at a relatively low level (19) and is not a good indicator for the metabolic state during growth in glycerol. Therefore, *pykF* may not be very active. PykA, on the hand, is activated by AMP, and its activity can still be regulated at the protein level. Therefore, *pykA* is a better choice during growth in glycerol.

The glucose transport genes *ptsHI*, *crr*, and *ptsG* were all down-regulated by 2- to 4-fold during growth in either acetate or glycerol, compared to glucose. This was consistent with previous observations (20). Even though *ptsHI* and *crr* are in the same operon, they are differentially regulated through different promoters (21).

Pentose Pathway. The pentose pathway genes were not regulated significantly in the media tested. An exception is the *gnd* (coding for 6-phosphogluconate dehydrogenase) gene, which is known to be under the control of growth rate (22). Thus, their transcription rates relative to other genes are proportional to the growth rate supported by the media. As expected, we observed that the gene was down-regulated more than 3-fold during growth in acetate (Figure 3a) in accordance with the growth rate decrease from 0.99 to 0.30 h⁻¹. However, the regulation was less significant in the glycerol medium because the growth rate in glycerol was only 30% lower (0.69 h⁻¹) than that in glucose.

Gluconeogenic Genes. The *pckA* gene was up-regulated 4- and 3-fold in acetate and glycerol, respectively, compared to that in glucose. This up-regulation

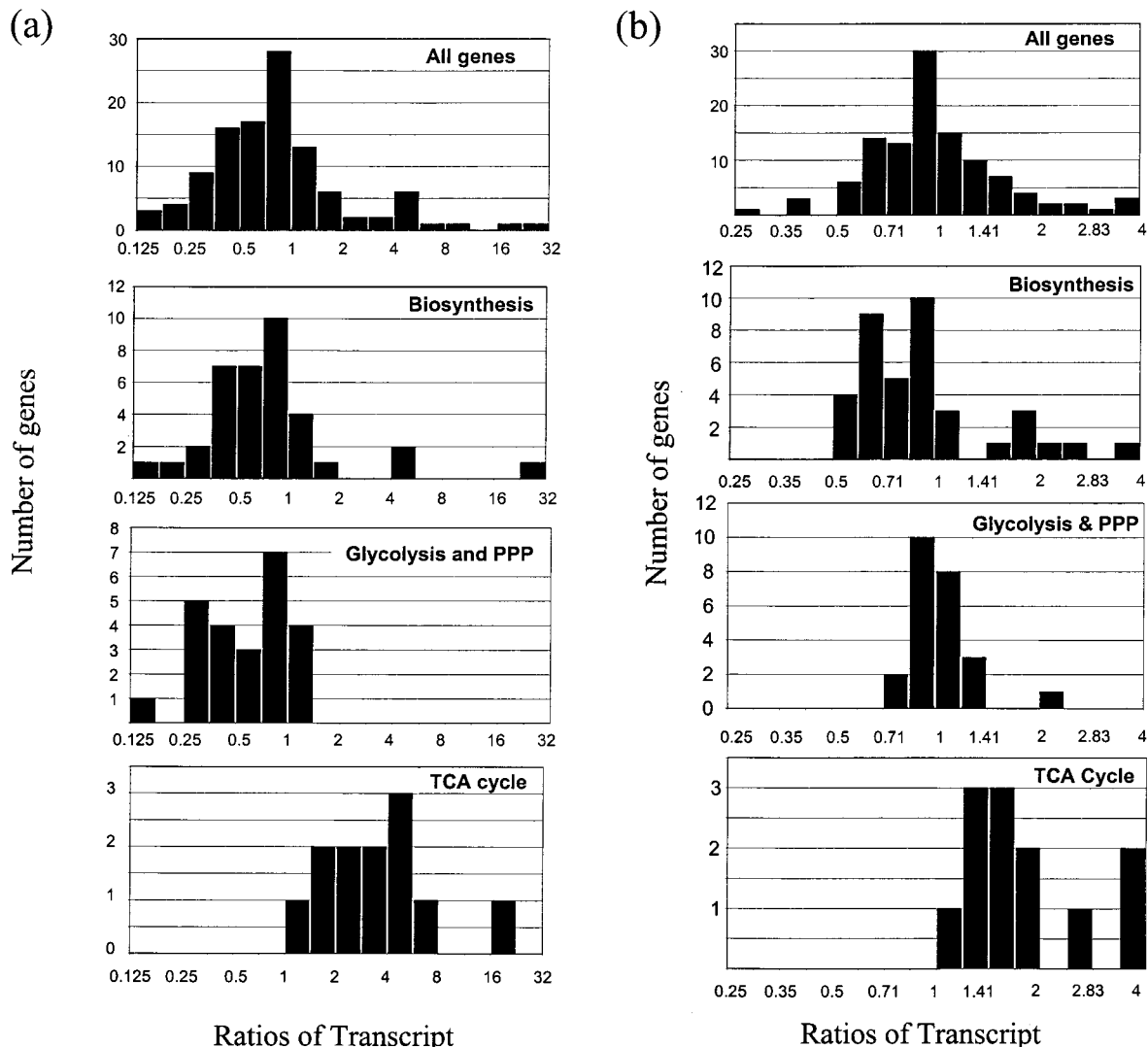


Figure 2. The histograms of the expression level ratios of the genes tested in (a) acetate and (b) glycerol media compared to those in glucose medium. These figures show the number of genes that exhibited a certain level of transcriptional regulation. The culture in glucose was used as a reference (the denominator in the transcript ratio). The histograms were plotted for all 111 genes together and for each individual pathway. PPP stands for pentose phosphate pathway.

was expected in acetate because of the need for metabolic flux from oxaloacetate (OAA) to phosphoenolpyruvate (PEP). However, the up-regulation in glycerol was unexpected, as the flux in the Pck-mediated reaction is not needed. It is possible that *pckA* is induced gratuitously by the increased cAMP (23) during growth in glycerol.

The *pps* gene was up-regulated 10-fold during growth in acetate but was almost unchanged during growth in glycerol. The up-regulation in acetate was surprising, as the *pps* gene was not required for growth under this condition. The gluconeogenic flux was expected to be mediated by PckA (24). Therefore, it is possible that at least one of the malic enzymes, which catalyze the transformation from malate to pyruvate, was also up-regulated, so that part of the gluconeogenic flux was carried by the malic enzyme-Pps path.

Fermentative Genes. Among the fermentation genes, the regulation of *pta* and *ackA* was most interesting. These genes were thought to be involved in both acetate synthesis and consumption and constitutively expressed in the cell (25). They are present in the same operon but are regulated differentially through different promoters (26). During growth in glycerol, *ackA* was down-regulated by almost 3-fold. This result is consistent with the observation that acetate production during growth in

glycerol is much lower than in glucose (24). However, during growth in acetate, both *pta* and *ackA* were down-regulated. This result suggests that other gene(s) may be responsible for acetate uptake and consumption during growth in acetate. Indeed, the *acs* gene (coding for acetyl-coA synthetase) may play such a role.

The intermediate of this pathway, acetyl phosphate, is known as an important metabolite involved in global regulation of metabolism (27). The 3-fold down-regulation of *ackA* in the glycerol medium was in stark contrast to the 30% down-regulation of the *pta* gene. The relatively small change of the latter may ensure the production of acetyl phosphate for regulatory purposes.

The *pfIBD* (encoding for pyruvate formate-lyase I and II, respectively) transcripts were not changed much in either media, although this fermentative pathway was not used when growing in acetate aerobically (28). This result suggests that the main control for switching between growth in glucose and acetate media may be at the posttranscriptional level. Indeed, pyruvate formate-lyase I is known to be interconverted between the active and inactive forms by an activase (29) and deactivase (30). These regulations may serve as the main control

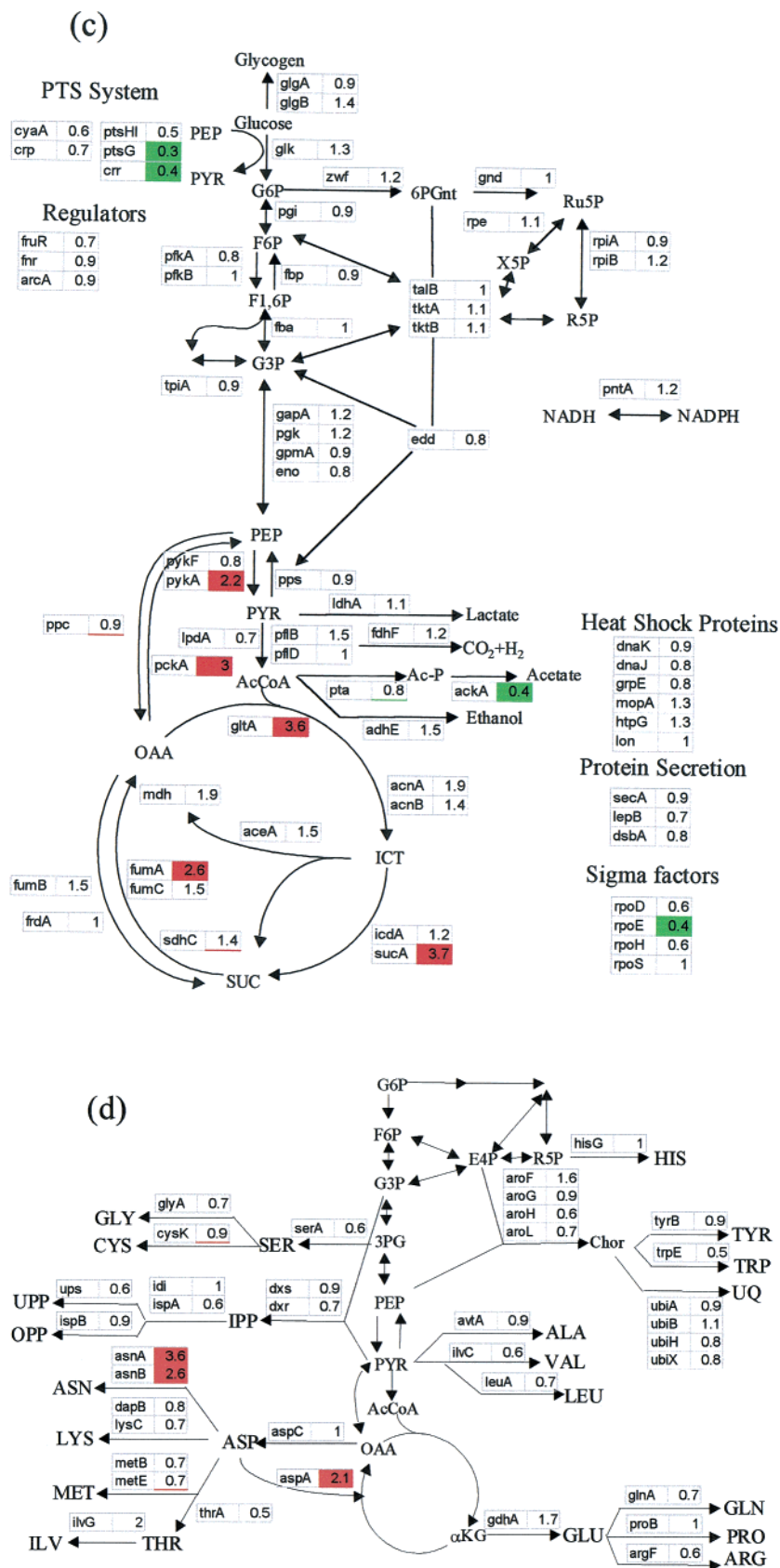


Figure 3. (continued) phosphate (Ac-P), isocitrate (ICT), α -ketoglutarate (KG), succinate (SUC), oxaloacetate (OAA), 6-phosphogluconate (6PGnt), ribulose-5-phosphate (Ru5P), ribose-5-phosphate (R5P), xylulose-5-phosphate (X5P), erythrose-4-phosphate (E4P), chorismate (Chor), ubiquinone (UQ), isopentenyl diphosphate (IPP), *trans*-octaprenyl diphosphate (OPP), and undecaprenyl diphosphate (UPP). Other metabolite symbols are three letter symbols for amino acids. Symbols for genes inside the boxes follow the *E. coli* K-12 linkage map. Red color represents increases greater than 2-fold, and green represents decreases greater than 2-fold.

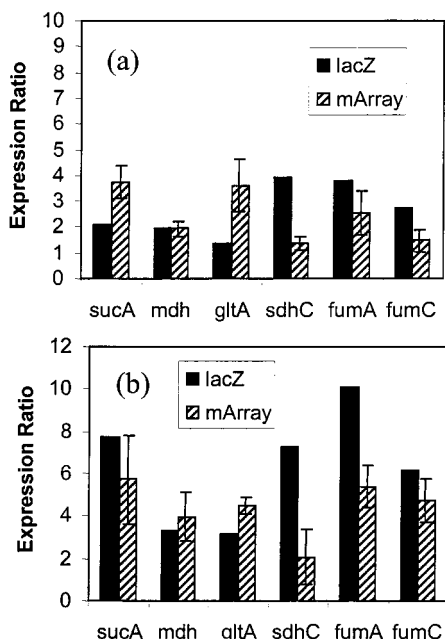


Figure 4. The comparison of transcript level changes detected by DNA microarrays and by *lacZ* operon fusions. The DNA microarray data are the geometric averages of four data for each gene with standard deviations. The increases of the gene expression levels in (a) acetate and (b) glycerol are compared to those in glucose. The *lacZ* fusion data were taken from the literature (12, 33–36), and the ratios were calculated from these data.

during the carbon source switching. Thus, it is not surprising that the transcriptional regulation was insignificant.

TCA Cycle. Most TCA cycle genes were up-regulated by 2- to 6-fold in the acetate medium. This up-regulation was expected since growth in acetate must be supported by aerobic metabolism. In addition, the glyoxylate shunt was increased by 20-fold, as this pathway was absolutely essential for growth in acetate and is highly regulated (31). The TCA cycle genes were also up-regulated in glycerol compared with cells cultured in glucose. These genes are known to be regulated by several regulators, such as cAMP-CRP, ArcA, and Fnr (32). Among these regulators, it is likely that the observed up-regulation in acetate and glycerol was mediated by cAMP-CRP, as cAMP level should be increased in nonglucose media.

To check the reliability of the microarray data, the transcript levels of six TCA cycle genes (*sucA*, *mdh*, *gltA*, *sdhC*, *fumA*, *fumC*) were compared with literature data that used *lacZ* transcriptional fusion for monitoring expression levels (Figure 4) (12, 33–36). Results showed that the microarray data were reasonably consistent with the data from the reporter genes, considering that the two approaches reported different information. The microarray data reflects the relative abundance of the transcript, which is influenced by both the transcriptional activity and the mRNA stability. The *lacZ*-reporter system, however, detects the integral of the transcriptional activity, assuming that the β -galactosidase protein is stable and the *lacZ* mRNA stability is constant. It should be noted that the reporter gene system cannot reflect the stability of the target mRNA. The largest deviation was seen in *sdhC*, which may suggest differential transcript stability between the reporter system and the native *sdhC* transcript.

Major Biosynthetic Pathways. The most striking result was the 30-fold increase in the *ilvG* transcript

abundance in acetate compared with that in glucose (Figure 2b). The *ilvG* gene, which is responsible for the synthesis of isoleucine, is regulated by many factors, such as IHF, ppGpp, isoleucine, and valine (37). The reason for this up-regulation was not immediately apparent. This gene was only mildly up-regulated (by about 2-fold) in glycerol media.

Most other biosynthetic genes were either decreased or insignificantly changed. The expression levels of 50% (18/36) of biosynthetic genes were decreased more than 30% in the glycerol medium (Figure 2b), reflecting the low growth rate which requires less fluxes to the biosynthetic monomers. Similarly, 80% (28/35) of biosynthetic genes were down-regulated to the same level in acetate media (Figure 2a). However, the expression levels of *asnA* (coding for the ammonia-dependent asparagine synthetase) and *asnB* (coding for the glutamine-dependent asparagine synthetase) were increased 3- and 2-fold, respectively, in the glycerol medium (Figure 3d) but were decreased 3- and 2-fold, respectively, in the acetate medium (Figure 3b). The transcript levels of these genes were changed in the similar way, although they are not in the same operon. This result suggests that they may serve the same purpose under these conditions. Since both genes can be induced by limitation of asparagine in the medium, it was possible that asparagine was limiting when the cells were grown in glycerol. However, supplementing asparagine in the glycerol medium did not increase the growth rate of the organism (data not shown). Therefore, the asparagine limitation hypothesis was not sustained.

The *aspA* gene was up-regulated in both glycerol and acetate. This result is consistent with the finding that *aspA* is under the control of catabolite repression (38). Therefore, it is de-repressed in media without glucose. However, the metabolic role of this gene in glycerol and acetate media is unclear.

The *aroF* (encoding 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase) gene increased slightly in the glycerol medium but significantly in the acetate medium. Both *aroF* and *tyrB* are known to be induced during tyrosine starvation (39). However, *tyrB* is either down-regulated slightly or insignificantly changed. These results suggest that additional regulation exists for *aroF* transcription.

Regulatory Genes. The transcript levels of *rpoE* and *rpoH* (encoding σ^E and σ^{32} , respectively) were decreased 2- to 3-fold in either acetate or glycerol. However, the transcript levels of heat shock genes (*dnaKJ*, *grpE*, *mopA*, *hspG*, *lon*) were unchanged in both media. Since the heat shock gene expression is largely proportional to the amount of σ^{32} , the constant heat shock gene expression implies that the level of σ^{32} activity remains constant despite its down-regulation at the transcript level. The level of σ^{32} is known to be regulated at multiple stages, including translational efficiency and protein stability (40, 41). It is likely that such posttranscriptional regulations of σ^{32} are significant in these conditions. The *lon* gene, coding for a heat-shock protease, was not significantly changed in either glycerol or acetate media, suggesting that protein degradation rate may remain unchanged in these media.

The expression levels of the genes for protein secretion and disulfide bond formation, *secA*, *lepB*, and *dsbA*, were decreased 2- to 3-fold in acetate. This result suggests that acetate accumulation, which is commonly observed in protein production in *E. coli*, may reduce the activity for protein secretion and disulfide bond formation.

Table 1. Comparison between Fluxes and Transcript Levels of *E. coli* Cultured in Glycerol and Acetate Compared to Those in Glucose^a

gene	glycerol vs glucose		acetate vs glucose	
	flux ratio	transcript ratio	flux ratio	transcript ratio
<i>pgi</i>	-0.08	0.9	-0.05	0.5
<i>pfkA</i>	-0.08	0.9	-0.05	0.3
<i>pfkB</i>	-0.08	1.0	-0.05	1.0
<i>fbp</i>	<i>b</i>	0.9	<i>b</i>	2.0
<i>fba</i>	-0.08	1.0	-0.05	0.3
<i>tpiA</i>	1.60	0.9	-0.05	0.3
<i>gap</i>	0.66	1.2	-0.09	0.4
<i>pgk</i>	0.66	1.2	-0.09	0.3
<i>gpmA</i>	0.65	0.9	-0.14	0.8
<i>eno</i>	0.65	0.8	-0.14	0.2
<i>pykA</i>	3.48	2.1	0.00	1.0
<i>pykF</i>	3.48	0.8	0.00	0.5
<i>pps</i>	<i>b</i>	0.9	0.00	10
<i>pflB</i>	0.61	1.5	0.00	1.1
<i>pflD</i>	0.61	1.0	0.00	0.9
<i>lpdA</i>	0.61	0.7	0.00	0.6
<i>ppc</i>	0.74	0.9	<i>b</i>	0.4
<i>pckA</i>	<i>b</i>	3.0	0.93	4.2
<i>gltA</i>	1.39	3.6	4.30	4.5
<i>acnA</i>	1.39	1.9	4.30	1.9
<i>acnB</i>	1.39	1.4	4.30	3.9
<i>icd</i>	1.39	1.2	2.93	1.4
<i>sucA</i>	1.66	3.8	3.96	5.7
<i>sdhC</i>	1.66	1.4	5.90	2.1
<i>fumA</i>	1.66	2.6	5.90	5.4
<i>fumC</i>	1.66	1.5	5.90	4.7
<i>mdh</i>	1.66	1.9	5.90	4.0

^a Metabolic fluxes were taken from the literature (24) and expressed as the ratio between the glycerol (or acetate) culture and the glucose culture. The negative flux ratios represent fluxes occur in opposite directions. ^b Flux cannot be separated from the reverse reaction catalyzed by another enzyme.

Comparison between Transcript Levels and Fluxes. It is known that metabolic flux ratios change significantly between cells grown in different carbon sources, particularly between glycolytic and gluconeogenic substrates. Although DNA microarray technology does not detect any regulation at the translational or posttranslational level, with a few well-characterized exceptions, the transcriptional regulation is the main mode of regulation in *E. coli*. Comparing the transcript profiles and the flux distribution showed that there was only a qualitative correlation between changes in fluxes and transcript levels, and many exceptions existed (Table 1). In glycerol, the TCA cycle fluxes increased by 30–70% (24), and similar changes in transcript levels were observed, except that *gltA*, *sucA*, and *fumA* increased much more (about 3-fold). The flux catalyzed by the pyruvate kinases was increased more than 3-fold (24), which was supported by the *pykA* up-regulation. The glycolytic fluxes decreased by about 30%, but the transcript levels were almost unchanged.

In acetate, most glycolytic genes catalyze the flux in the gluconeogenic direction. This gluconeogenic flux is much smaller than the glycolytic flux when the cells were grown on glucose. The transcript levels of these genes did qualitatively reflect such changes. The TCA cycle flux increased significantly, and so did the transcript levels of the TCA cycle genes. As discussed above, the *pps* gene was not known to support *E. coli* growth in acetate, and thus the flux mediated by this gene was calculated to be zero. However, its transcript level increased by 10-fold, suggesting an unknown role of this gene during growth in acetate.

This comparison shows that transcript levels cannot be used to infer fluxes in a quantitative manner. How-

Table 2. Genes Differentially Regulated in Acetate and Glycerol Compared to Glucose^a

(a) Acetate/Glucose		
function	mechanism	
	known	unknown
known	TCA cycle, <i>pckA</i>	PTS, <i>ppc</i> , glycolysis, biosynthesis genes
unknown	<i>aspA</i> , <i>gnd</i>	<i>pta.ackA</i> , <i>adhE</i> , <i>pps</i> , <i>ilvG</i> , <i>aroF</i> , <i>secA</i> , <i>dsbA</i>
(b) Glycerol/Glucose		
function	mechanism	
	known	unknown
known	TCA cycle (<i>gltA</i> , <i>sucA</i> , <i>fumA</i>), <i>pykA</i>	<i>ptsG</i> , <i>crr</i> , <i>ackA</i>
unknown	<i>pckA</i> , <i>aspA</i>	<i>asnA</i> , <i>asnB</i>

^a The genes are divided into four groups according to the knowledge of their mechanisms and functions.

ever, changes in transcript levels reveal the existence of significant regulation and suggest possible strategies of genetic modification.

Conclusion

This work demonstrates the utility of the DNA microarray technology in studying regulation of gene expression. Complete analysis of the data requires a statistical model, which is currently under development. Therefore, relatively large (2-fold) and consistent changes were used as criteria for analysis. As discussed above, many genes are regulated either by unknown mechanisms or for unknown functions. Table 2 categorizes the genes that are differentially regulated in different media according to the knowledge of their mechanisms and functions. Surprisingly, despite the extensive work on *E. coli* physiology in different carbon sources, still many genes are regulated for unknown purposes by unknown mechanisms. For example, several genes were unexpectedly up-regulated, such as *pps*, *ilvG*, *aroF*, *secA*, and *dsbA* in acetate and *asnA* and *asnB* in glycerol, or down-regulated, such as *ackA*, *pta*, and *adhE* in acetate, without apparent reasons. Meanwhile, many genes are regulated for apparent purposes but by unknown mechanisms. For example, the down-regulations of PTS genes in both acetate and glycerol media and down-regulations of glycolysis and biosynthesis genes in acetate could be explainable by the flux requirement, but further investigation for the molecular regulatory mechanisms is needed. Furthermore, there exists a group of genes whose regulation can be attributed to known mechanisms, but the physiological roles of such regulation remain unclear. This group includes *pckA* and *aspA*, which are up-regulated in glycerol, and *gnd* and *aspA*, which are down- and up-regulated, respectively, in acetate.

Apparently, much more remains to be learned regarding *E. coli* physiology. Before complete understanding of regulation, flux analysis based only on pathway stoichiometry needs to proceed with caution. The technology demonstrated here provides a powerful yet economical tool for characterizing different strains under various conditions. It may prove to be useful for strain development, process diagnosis, and process monitoring in bioreactors.

Acknowledgment

This study was supported by NSF grant BES-9906758. The authors appreciate Mum dough Aldimassi and Stan Nelson of UCLA Medical School for generating DNA microarrays and other technical assistance. Also we

thank Robert Gunsalus for providing the MC4100 strain and useful suggestions for the experiments.

References and Notes

- (1) Schena, M.; Shalon, D.; Davis, R. W.; Brown, P. O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **1995**, *270*, 467–470.
- (2) Wang, D. G. et al. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* **1998**, *280*, 1077–1082.
- (3) Behr, M. A.; Wilson, M. A.; Gill, W. P.; Salamon, H.; Schoolnik, G. K.; Rane, S.; Small, P. M. Comparative genomics of BCG vaccines by whole genome DNA microarray. *Science* **1999**, *284*, 1520–1523.
- (4) DeRisi, J. L.; Iyer, V. R.; Brown, P. O. Exploring the Metabolic and Genetic Control of Gene Expression on a Genomic Scale. *Science* **1997**, *278*, 680–686.
- (5) Wodicka, L.; Dong, H.; Mittmann, M.; Ho, M. H.; Lockhart, D. J. Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nat. Biotechnol.* **1997**, *15*, 1359–1367.
- (6) Chu, S.; DeRisi, J. L.; Eisen, M. B.; Mulholland, J.; Botstein, D.; Brown, P. O.; Herskowitz, I. The transcriptional program of sporulation in budding yeast. *Science* **1998**, *282*, 699–705.
- (7) Ferea, T. L.; Botstein, D.; Brown, P. O.; Rosenzweig, R. F. Systematic changes in gene expression patterns following adaptive evolution in yeast. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9721–9726.
- (8) Lee, C. K.; Klopp, R. G.; Weindruch, R.; Prolla, T. A. Gene expression profile of aging and its retardation by caloric restriction. *Science* **1999**, *285*, 1390–1393.
- (9) Iyer, V. R.; Eisen, M. B.; Ross, D. T.; Schuler, G.; Moor, T.; Lee, J. C. F.; Trent, J. M.; Staudt, L. M.; Hudson, J., Jr.; Boguski, M. S.; Lashkari, D.; Shalon, D.; Botstein, D.; Brown, P. O. The transcriptional program in the response of human fibroblasts to serum. *Science* **1999**, *283*, 83–87.
- (10) Kao, C. M. Functional genomic technologies: creating new paradigms for fundamental and applied biology. *Biotechnol. Prog.* **1999**, *15*, 304–311.
- (11) Richmond, C. S.; Glasner, J. D.; Mau, R.; Jin, H.; Blattner, F. R. Genome-wide expression profiling in *Escherichia coli* K-12. *Nucleic Acids Res.* **1999**, *27*, 3821–3835.
- (12) Park, S. J.; McCabe, J.; Turna, J.; Gunsalus, R. P. Regulation of the citrate synthase (*glcA*) gene of *Escherichia coli* in response to anaerobiosis and carbon supply: role of the *arcA* gene product. *J. Bacteriol.* **1994**, *176*, 5086–5092.
- (13) Ausubel, F. M. *Current Protocols in Molecular Biology*; John Wiley and Sons: 1994.
- (14) Akada, R. Quick-check method to test the size of *Escherichia coli* plasmids. *Biotechnique* **1994**, *17*, 58.
- (15) Welford, S. M.; Gregg, J.; Chen, E.; Garrison, D.; Sorensen, P. H.; Denny, C. Y.; Nelson, S. F. Detection of differentially expressed genes in primary tumor tissues using representational difference analysis coupled to microarray hybridization. *Nucleic Acids Res.* **1998**, *26*, 3059–3065.
- (16) Fraenkel, D. G. Glycolysis. In *Escherichia coli and Salmonella*; Neidhardt, F. C., Ed.; ASM Press: Washington, D.C., 1996; pp 189–198.
- (17) Saier, M. H., Jr.; Ramseier, T. M. The catabolite repressor/activator (Cra) protein of enteric bacteria. *J. Bacteriol.* **1996**, *178*, 3411–3417.
- (18) Chao, Y. P.; Liao, J. C. Metabolic responses to substrate futile cycling in *Escherichia coli*. *J. Biol. Chem.* **1994**, *269*, 5122–5126.
- (19) Lowry, O. H.; Carter, J.; Ward, J. B.; Glaser, L. The effect of carbon and nitrogen sources on the level of metabolic intermediates in *Escherichia coli*. *J. Biol. Chem.* **1971**, *246*, 6511–6521.
- (20) De Reuse, H.; Danchin, A. The *ptsH*, *ptsI*, and *crr* genes of the *Escherichia coli* phosphoenolpyruvate dependent phosphotransferase system: a complex operon with several modes of transcription. *J. Bacteriol.* **1988**, *170*, 3827–3837.
- (21) Fox, D. K.; Presper, K. A.; Adhya, S.; Roseman, S.; Garges, S. Evidence for two promoters upstream of the *pts* operon: Regulation by the cAMP receptor protein regulatory complex. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 7056–7059.
- (22) Pearse, A. J.; Wolf, R. E. Determination of the growth rate-regulated steps in expression of the *Escherichia coli* K-12 *gnd* gene. *J. Bacteriol.* **1994**, *176*, 115–122.
- (23) Goldie, H.; Saier, M. H. J. Binding of cyclic AMP receptor protein to the *pckA* gene of *Escherichia coli*. *Abstracts of the General Meeting of the American Society for Microbiology* **1992**, *92*, 235.
- (24) Holms, H. Flux analysis and control of the central metabolic pathways in *Escherichia coli*. *FEMS Microbiol. Rev.* **1996**, *19*, 85–116.
- (25) Brown, T. D.; Jones-Mortimer, M. C.; Kornberg, H. L. The enzymic interconversion of acetate and acetyl-coenzyme A in *Escherichia coli*. *J. Gen. Microbiol.* **1977**, *102*, 327–336.
- (26) Kakuda, H.; Hosono, K.; Shiroishi, K.; Ichihara, S. Identification and characterization of the *ackA-pta* operon and complementation of the acetate utilization by an *ack-pta* deletion mutant of *Escherichia coli*. *J. Biochem.* **1994**, *116*, 916–922.
- (27) McCleary, W. R.; Stock, J. B.; Ninfa, A. J. Is acetyl phosphate a global signal in *Escherichia coli*? *J. Bacteriol.* **1993**, *175*, 2793–2798.
- (28) Sawers, G.; Suppmann, B. Anaerobic induction of pyruvate formate-lyase gene expression is mediated by the ArcA and Fnr proteins. *J. Bacteriol.* **1992**, *174*, 3474–3478.
- (29) Sauter, M.; Sawers, R. G. Transcriptional analysis of the gene encoding pyruvate formate-lyase activating enzyme of *Escherichia coli*. *Mol. Microbiol.* **1990**, *4*, 355–363.
- (30) Knappe, J.; Sawers, G. A radical-chemical route to acetyl-CoA: the anaerobically induced pyruvate formate-lyase system of *Escherichia coli*. *FEMS Microbiol. Rev.* **1990**, *75*, 383–398.
- (31) Chung, T.; Klumpp, D. J.; LaPorte, D. C. Glyoxylate bypass operon of *Escherichia coli*: cloning and determination of the functional map. *J. Bacteriol.* **1988**, *170*, 386–392.
- (32) Cronan, J. E., Jr.; Laporte, D. Tricarboxylic acid cycle and glyoxylate bypass. In *Escherichia coli and Salmonella*; Neidhardt, F. C. Ed.; ASM Press: Washington, D.C., 1999; pp 206–216.
- (33) Park, S. J.; Tseng, C. P.; Gunsalus, R. P. Regulation of succinate dehydrogenase (*sdhCDAB*) operon expression in *Escherichia coli* in response to carbon supply and anaerobiosis: role of ArcA and Fnr. *Mol. Microbiol.* **1995**, *15*, 473–482.
- (34) Park, S. J.; Cotter, P. A.; Gunsalus, R. P. Regulation of malate dehydrogenase (*mdh*) gene expression in *Escherichia coli* in response to oxygen, carbon, and heme availability. *J. Bacteriol.* **1995**, *177*, 6652–6656.
- (35) Park, S. J.; Chao, G.; Gunsalus, R. P. Aerobic regulation of the *sucABCD* genes in *Escherichia coli*, which encode α -ketoglutarate dehydrogenase and succinyl coenzyme A synthetase: role of ArcA, Fnr, and the upstream *sdhCDAB* promoter. *J. Bacteriol.* **1997**, *179*, 4138–4142.
- (36) Park, S. J.; Gunsalus, R. P. Oxygen, iron, carbon, and superoxide control of the fumarase *fumA* and *fumC* genes of *Escherichia coli*: role of the *arcA*, *fnr*, and *soxR* gene products. *J. Bacteriol.* **1995**, *177*, 6255–6262.
- (37) Umbarger, H. E. Biosynthesis of the branched-chain amino acids. In *Escherichia coli and Salmonella*; Neidhardt, F. C. Ed.; ASM Press: Washington, D.C., 1996; pp 442–457.
- (38) Woods, S. A.; Guest, J. R. Differential roles of the *Escherichia coli* fumarases and *fnr*-dependent expression of fumarase B and aspartase. *FEMS Microbiol. Lett.* **1987**, *48*, 219–224.
- (39) Pittard, A. J. Biosynthesis of the aromatic amino acids. In *Escherichia coli and Salmonella*; Neidhardt, F. C. Ed.; ASM Press: Washington, D.C., 1996; pp 458–484.
- (40) Straus, D. B.; Walter, W. A.; Gross, C. A. The heat shock response of *Escherichia coli* is regulated by changes in the concentration of σ -32. *Nature* **1987**, *329*, 348–351.
- (41) Kanemori, M.; Mori, H.; Yura, T. Induction of heat shock proteins by abnormal proteins results from stabilization and not increased synthesis of σ -32 in *Escherichia coli*. *J. Bacteriol.* **1994**, *176*, 5648–5653.

Accepted for publication January 20, 2000.

BP000002N