Development of DNA Microarrays for Metabolic Pathway and Bioprocess Monitoring

Gregory Stephanopoulos
Department of Chemical Engineering
Massachusetts Institute of Technology

<u>FINAL PROGRESS REPORT</u>

July 31, 2004

1.0   Use of full genome microarrays to assess strains of *Synechocystis* accumulating intracellular poly-β-hydroxybutyrate (PHB)

Transcriptional profiling experiments utilizing DNA microarrays to study the intracellular accumulation of PHB in *Synechocystis* has proved difficult in large part because strains that show significant differences in PHB which would justify global analysis of gene expression have not been isolated. Thus the first sections of this report detail efforts to design phenotypic screens for the isolation of strains for global transcription and flux analyses.

1.1   Bacterial Poly-3-hydroxybutyrate Screening

In order to assess large combinatorial libraries of *Synechocystis* PCC 6803 for accumulation of poly-3-hydroxybutyrate (PHB), a number of methods for performing single cell measurements were evaluated. The methods evaluated were: (a) dielectrophoresis field flow fractionation (DEP-FFF), (b) flow cytometry, and (c) inertial microfluidics. The different methods relied on various material properties related to PHB content in the cell including: electrical conductivity, fluorescent labeling, and density. These methods should be able to measure and isolate cells from a population of ~10^7 based on their intracellular PHB accumulation.

1.2   Dielectrophoresis Field Flow Fractionation

DEP-FFF is a chromatographic-like method for separating micron size particles on their electrical properties. DEP-FFF uses an electric force that is dependent on PHB content to levitate cells to different heights in a parabolic flow field. This microfluidic technique was developed and characterized for use in separating *Synechocystis* libraries. DEP-FFF has shown the most promising results of the systems characterized.
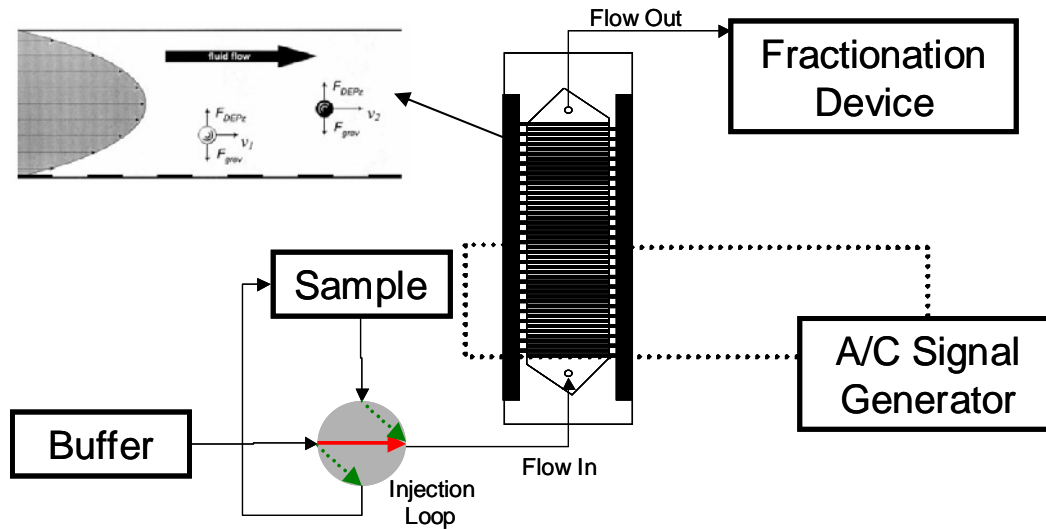
DEP forces are well suited for screening *Synechocystis* libraries. DEP is a charge-independent force that can be exploited in alternating current electric fields. The DEP force is small (~pN) and is effective on the micron length scale. This makes it appropriate for moving cells in microfluidic devices.

The DEP force is a strong function of the amount of polymer in the cell. The magnitude and direction of the DEP force is dependent on a variety of factors including the conductivity of the media, cytoplasm, the frequency of the AC electric field, and electric field gradients. The cytoplasmic conductivity of *Synechocystis* is a strong function of the amount of polymer accumulated. This is because the volume of PHB in the cell is nonconductive, and as this volume changes, so does the average conductivity of the cytoplasm. Cells with high PHB have a strong DEP force, and cells with little PHB have a weak force.

DEP-FFF takes advantage of these PHB-dependent DEP forces. DEP-FFF has been used to separate viable bacteria from dead bacteria, bacteria from yeast, and cancerous from non-cancerous cells (references here). The DEP-FFF device used in the work was

fabricated in collaboration with the Prof. Joel Voldman (Research Laboratory of Electronics, Massachusetts Institute of Technology).

Figure 1.1 gives a schematic representation of the device. A sample loop is used to load 10^7 cells into the front of the DEP-FFF chamber. Flow is stopped, and the cells are allowed to settle to the bottom of the chamber. At this time, the signal generator is turned on and the electric field is induced. The DEP force repels cells from the electrodes (see Figure 1.1 inset). The strength of the DEP force decays with distance from the electrodes. At some point, the cells equilibrate at a height where the DEP force and gravitational force are equal. Since the magnitude of the DEP force is a function of the PHB content of the cells, the equilibrium height will be higher for cells that contain more PHB.

Figure 1.1: Schematic Representation of DEP-FFF

After the cells have been equilibrated in the electric field, flow is induced. The laminar flow profile of the device causes cells that are near the centerline to travel quickly through the device and cells near the bottom to travel slowly (all cells are well below the half height to avoid convolution do to the upper half of the parabola). This causes high PHB cells to have a short residence time in the device, and cells collected in early fractions to have more PHB.

Experiments were carried out to determine the best design for discriminating cells based on PHB content. Parameters evaluated were solution conductivity, voltage, frequency, electrode width, and flow chamber height. Table 1.1 gives the values of the parameters used in the current DEP-FFF design. Electrode width was varied between 15 um and 50 um. Electrode width had no effect on levitation height. Solution conductivity was varied using

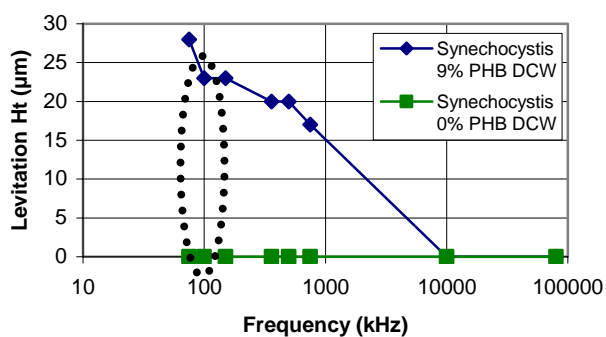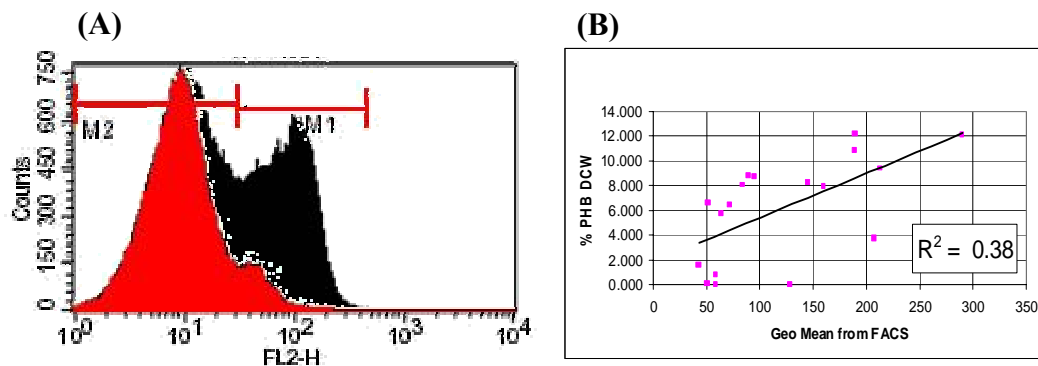**Table 1.1: Parameters for DEP-FFF**

Parameter	Value
Solution Conductivity	0.1 S/m
Voltage	10 V
Frequency	100 kHz
Electrode Width	50 um
Chamber Height	400 um

Figure 1.2: Levitation Height of Synechocystis with Different Amounts of PHB

different ratios of phosphate-buffered saline and DI water. Conductivities between 0.1 uS/m and 1 S/m were tested. The best differentiation between 9% and 0% PHB dry cell weight (DCW) *Synechocystis* was observed with 0.1 S/m conductivity. Voltage >3 V was necessary to levitate the cells. Electrical frequencies between 10 kHz and 80 MHz were evaluated. At 0.1 S/m the best discrimination was at 100 kHz. Table 1.1 shows the values chosen for each parameter. Figure 1.2 shows the equilibrium heights for a 0% and 9% PHB DCW sample under these conditions. This translates to a volume fraction separation of 3.4 mL.

This device provides an interesting, novel way to screen libraries based on their electrical properties. Further work will include: (a) sorting *Synechocystis* libraries on PHB content, (b) applying the device to *E. coli* for screening for PHB and membrane-associated caratenoids, and (c) trying new geometries which will improve the sensitivity and convenience of the separation.

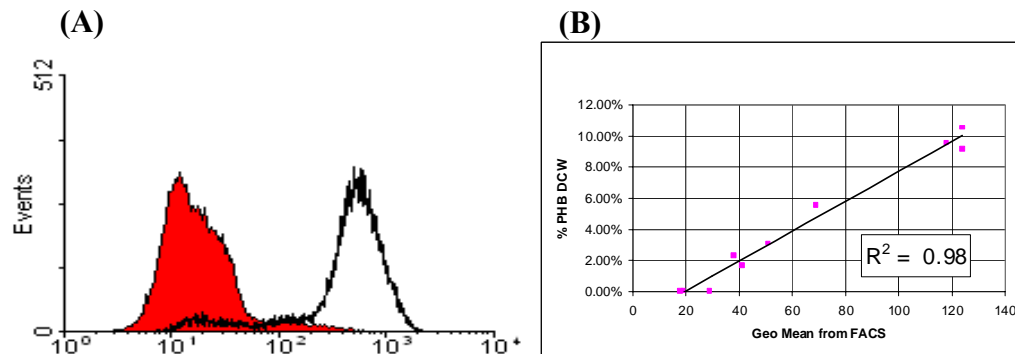## 1.3 Fluorescence Activated Cell Sorting (FACS)

FACS is an established technology for sorting cells based on optical properties and fluorescence. Cell sorting on the scale of millions is routinely achieved by FACS. Nile red, a neutral lipid stain, has been shown to fluorescently label PHB in many bacteria (references). Adapting these protocols for staining *Synechocystis* initially met with failure. Non-specific staining, presumably due to the large amount of photosynthetic membrane, made it difficult to differentiate between high and low PHB cells. Figure 1.3A shows the discrimination between a high PHB (10% DCW) and low PHB(1% DCW) population of *Synechocystis* by nile red fluorescence using an adapted protocol from the literature. Many of the cells did not stain properly, and fluoresced the same as the negative sample. Figure 1.3B shows the correlation between the mean of the fluorescence distribution and the PHB content. The R^2 showed that there was no reliable correlation between nile red fluorescence and PHB content of a culture.

(A)   **(B)**

Figure 1.3: Literature-based Staining Protocol (A) FACS-measured distribution of
PHB positive (black) and PHB negative (red) populations. Most of the positive control does not stain correctly.
(B) Correlation between geometric mean of FACS distribution and % DCW PHB. The protocol did not yield
any statistical correlation.

Adding a de-staining step to the protocol made it possible to use nile red fluorescence as
a good marker for sorting *Synechocystis.* After altering the protocol, quantitative
accuracy of PHB became possible. Figure 1.4(A) shows the fluorescence distribution for
high PHB and low PHB samples. Here there is almost two-log difference between the
positive and negative distributions. Figure 1.4(B) shows the correlation between the
mean of the fluorescence distribution and the PHB content. This shows that there is good
correlation between the fluorescence mean of the distribution and the PHB content using
the de-staining step.

Sorting experiments were used to validate the protocols ability to identify and isolate
high PHB cells. To test the ability to pick high PHB cells, 10 % PHB DCW

(A)   **(B)**

Figure 1.4: Modified Staining Protocol (A) FACS-measured distribution of PHB positive
(black outline) and PHB negative (red) populations. There is a good separation between the 2 populations. (B)
Correlation between geometric mean of FACS distribution and % DCW PHB. The protocol correlates well
PHB from 0-10% PHB DCW.

Synechocystis was mixed with PHB negative *Synechocystis* at a ratio of 1:100. High
fluorescence cells were sorted out by the protocol and recultured. The PHB content of
the sorted cells was the same as that of the high PHB cells (data not shown).

This protocol was deployed to sort a library of *Synechocystis* to identify high PHB
mutants. Transposons were stably inserted randomly into *Synechocystis* genomic DNA,
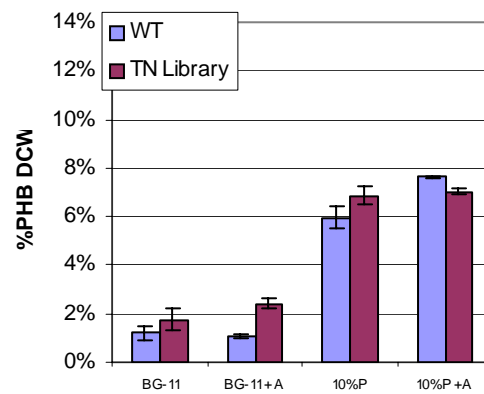disrupting genes and adding a selectable marker. The disrupted genes would provide

perturbation to metabolism and other cell processes that would affect the rate of PHB accumulation in the cell.

The library was grown under four media conditions (Table 1.2) known to affect the PHB accumulation properties of the cells. The libraries were grown to stationary phase.

Table 1.2: Media Conditions used for PHB Enrichment Experiment

Name	Media Condition	Approx. PHB Accumulation
BG-11	Regular BG-11	~1% PHB DCW
BG-11 + A	BG-11 supplemented with 10 mM sodium acetate	~3% PHB DCW
10% P	BG-11 with 10% the potassium phosphate	~10% PHB DCW
10% P + A	BG-11 with 10% the potasssium phosphate, supplemented with 10 mM sodium acetate	~11% PHB DCW

The cells were then sorted using the protocol. Cells with the highest 0.1% of the fluorescence distribution were collected and re-cultured in the same growth conditions. Figure 1.5 shows the PHB accumulation of the enriched cells (TN Library) against wild type (WT) cells. Only the library grown with BG-11 + A was found to have a significantly enriched population after sorting. BG-11 and 10% P enriched cultures showed incremental improvements and 10% P +A actually produced less in the enriched population than WT.

Figure 1.5: PHB Accumulation of Library Enriched for High PHB Mutants

Further work will focus on developing an analogous protocol for separating *E. coli* PHB libraries, and investigating the reason that PHB enrichment only occurred in one of the conditions.

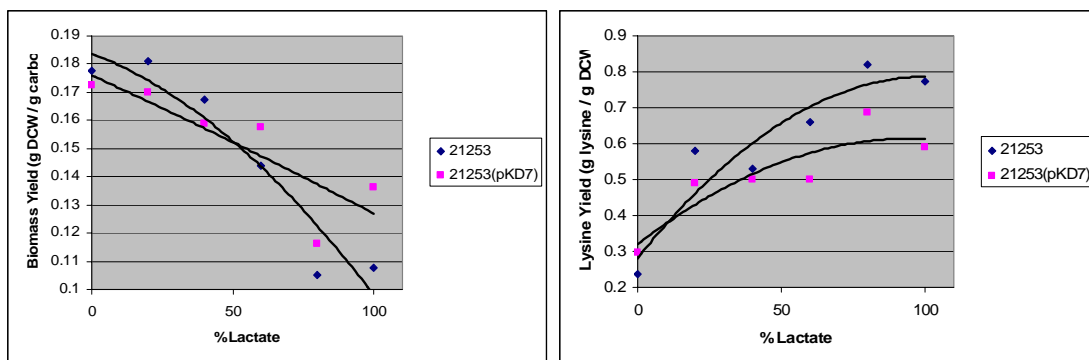## 2.0 *Corynebacterium glutamicum* Partial Genome DNA Microarrays

A series of studies have been carried out in our lab using our internally designed and manufactured partial genome DNA microarrays to explore the gene expression phenotypes of mutants of the amino acid-producing microorganism *Corynebacterium glutamicum*. *C. glutamicum* is an industrially significant bacteria noted for its biosynthesis of lysine, glutamate, and threonine, as well as the metabolic simplification caused by its relative lack of isozymes. The two mutants used in these investigations were the lysine over-producer ATCC 21253, and the pyruvate-carboxylase-over-expressing transformant 21253(pKD7). The differing abilities of these two related strains to generate biomass and lysine have been previously shown to be influenced by the type

of carbon source available in the growth medium. This work examines the transcriptomes of cells harvested under several different culture conditions to gain an understanding of the relationship between mRNA concentrations, metabolic fluxes, and changes in macroscopic culture properties brought about by, for example, such perturbations in media composition and gene over-expression.

2.1 Transcriptional Phenotypes of *C. glutamicum* Mutants Cultivated on Different Carbon Sources
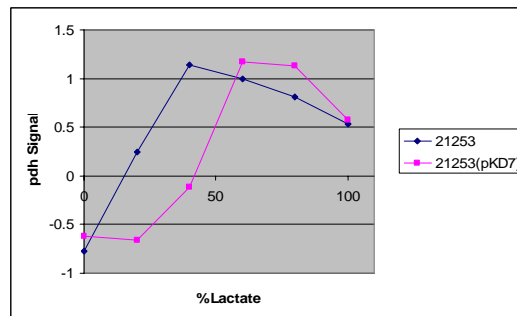
To gauge the effects that the carbon source has in different ratios of glucose and lactate, a series of twelve cultures were grown in minimal media; six for each of the two strains. The flasks had 0%, 20%, 40%, 60%, 80%, and 100% of their 20 g/l carbon source in the form of lactate, with the remainder consisting of glucose. In this way, the total concentration of carbon fed to the bacteria was kept constant at 680 mM. The growth and lysine production results of the twelve cultures are presented in Figure 2.1. Agreeing with expectations, in the case of 100% lactate, 21253(pKD7) showed a 30% higher yield of biomass on carbon than that of the control strain. Additionally, it was found that the two strains showed very similar growth behavior in the case of 100% glucose, with the disparity between the biomass yields increasing along with the amount of lactate fed to the cultures. Both strains showed a decrease in growth as the media contained more lactate, but the decline in yield for the pyruvate-carboxylase-over-expressing strain was approximately 20% while that for the parental strain was 40%. The yields of lysine on biomass for the two strains also showed a significant 30% difference for growth on lactate. In trends that mirrored those described above for growth, the lysine yields of each strain were nearly identical when the medium carbon source was solely glucose. As the percentage of lactate increased, both 21253 and 21253(pKD7) exhibited elevations in specific lysine yields, again with the more dramatic changes occurring in the control strain. The improvement in yield for 21253 was greater than 3-fold, while that of 21253(pKD7) was nearly 2-fold.

The mRNA abundances present in the twelve studied cultures were assayed by sampling total RNA pools from the cultures after 22-hr growth and applying our *C. glutamicum* DNA microarrays. The trends of relative transcript concentrations over the range of carbon substrate compositions tested was then determined for each of the genes

Figure 2.1. Yields of biomass on carbon source and of lysine on biomass. 21253 and 21253(pKD7) were grown on minimal medium containing 20 g/l glucose and/or lactate in
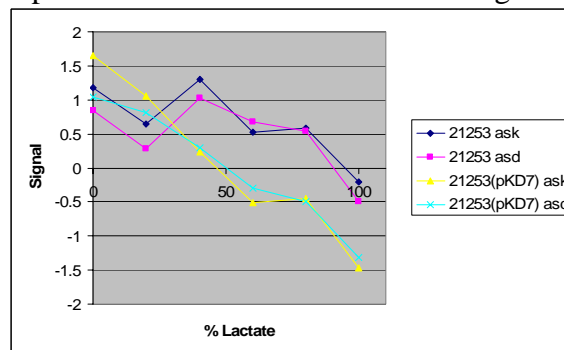
the percentages indicated. Yields were based on final biomass and lysine and initial carbon source concentrations.

Figure 2.2. Transcriptional profiles of *pdh*. Signal ratios are shown in $\log_2$-scale. 21253 and 21253(pKD7) were grown on minimal medium containing 20 g/l glucose and/or lactate in the percentages indicated. Identical control cDNA labeled with Cy5-dUTP was used in each hybridization containing sample cDNA derived from RNA harvested from one culture of 21253 or 21253(pKD7) and labeled with Cy3-dUTP.

represented on the arrays. Some of the more significant findings of the experiment include:

- Genes for five of the enzymes—*tpi*, *gap*, *pgk*, *pgm*, and *pyk*—responsible for glycolysis and gluconeogenesis, and one enzyme—*gnd*—responsible for the pentose phosphate pathway saw fairly steady declines in transcript concentration over the entire range of increasing lactate percentages. This suggests that the demands by the cells' regulatory system for transcription of these genes drops as the availability of substrates forces the network to adapt from a glycolytic mode to a gluconeogenic one.

- As lactate and glucose approached equal medium concentrations, *pdh* mRNA increased before steadying or dropping as lactate increased further from 60% to 100%. The trends for the two strains, shown in Figure 2.2, differed somewhat though, with that of the transformed strain exhibiting a lag before both the increase in transcript concentration and the eventual slight decline. This may be
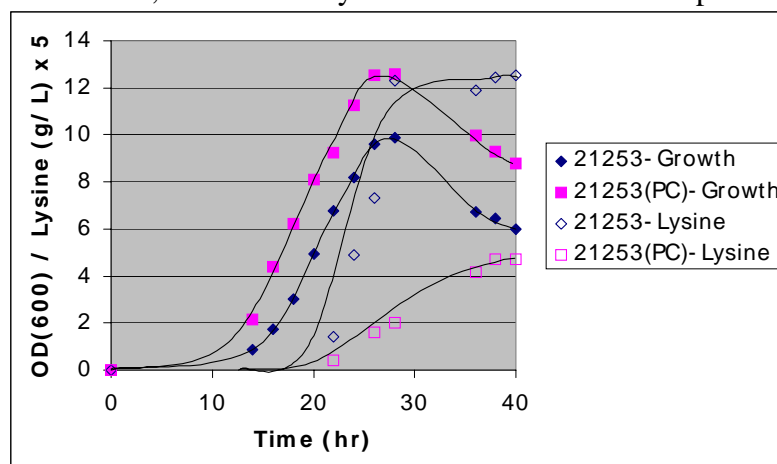

Figure 2.3. Transcriptional profiles of *ask* and *asd*. Signal ratios are shown in $\log_2$-scale. 21253 and 21253(pKD7) were grown on minimal medium containing 20 g/l glucose and/or lactate in the percentages indicated. Identical control cDNA labeled with Cy5-

because the transformant has an enhanced alternate means, in the form of *pca*, of depleting intracellular pyruvate pools without using *pdh*.

- Two genes from the tricarboxylic acid (TCA) cycle, *aco* and *icd*, had mRNA concentrations that generally increased along with the lactate content of the growth medium with the increases being much larger for 21253(pKD7) (approximately 400% and 300%, respectively) than for the control strain (approximately 70% and 40%). Increased transcription of these genes in the transformant may provide increased flux in the TCA cycle, along with an increased generation of energy and reducing equivalents allowing it to achieve its higher growth yield on lactate.

- Transcript amounts for the two glyoxylate bypass genes, *aceA* and *aceB* remain fairly steady as the lactate increases from 20% to 80%, and then rise significantly between 80% and 100% lactate. This trend directly mirrors that of *icd*, with a complementary drop seen between 80% and 100% lactate. Because *icd* and *aceA* catalyze reactions that share isocitrate as a substrate, the simultaneous increase in mRNA of one and decrease of another may signal a resulting shift in flux from one reaction path to another. In addition, the closely matching trends for *aceA* and *aceB* are very suggestive of a shared regulatory strategy.

- For all glucose/lactate mixtures tested, the mRNA abundance of *pca* was higher in the strain that had been genetically engineered for its over-expression.

- The profiles of *pca* and *malE* for 21253(pKD7) show a high degree of similarity, which may be evidence of co-regulation. Such a regulatory strategy may exist in *C. glutamicum* as an additional means of generating reducing equivalents by linking the two reactions through a modified TCA cycle or a smaller transhydrogenation cycle.

- Both *ask* and *asd* had decreases in mRNA levels for both strains as the carbon source shifted to lactate (Fig. 2.3). The decreases were much more pronounced in the transformed strain, for which they were 8-fold and 6-fold respectively, than in the parental strain, for which they were 3-fold and 2-fold respectively.
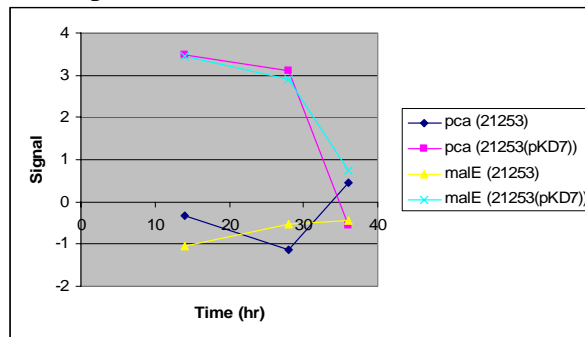
Figure 2.4. Growth and lysine production of 21253 and 21253(pKD7) grown on minimal medium containing 20 g/l lactate.

As with lysine yields, the transcript concentrations in both strains were very similar during growth on 100% glucose. The larger decreases seen in the transformant for growth on lactate may be influences on its lower lysine productivities in those conditions.

2.2 Transcriptional Phenotypes of *C. glutamicum* Mutants in Different Culture Phases

To further explore the transcriptional differences that exist between these two *C. glutamicum* strains, another experiment was carried out in which their cultures were sampled at three different time points representing different phases of growth and lysine production. In this way, we could also see how concentrations of specific mRNA sequences within the cells change over time to adapt to shifting demands on the metabolic networks of the cells. The growth and lysine generation trends of the two strains are shown in Figure 2.4. Again, the strain 21253(pKD7) was capable of supporting a larger cell concentration than 21253 when grown on an equal amount of lactate as the carbon source. Also again, the specific productivity of lysine in these transformant cells was lower than that seen in those of the parental strain. In both cultures, lysine production begins in late-exponential growth and coincides with the depletion of threonine, an essential nutrient for the cells and an inhibitor of lysine synthesis. Samples were taken at 14-, 28-, and 36-hours post-inoculation, and therefore represent the cultures in early-exponential growth, in late-exponential growth as lysine is beginning to be generated, and when the growth and lysine biosynthesis are both essentially stationary. The DNA microarrays were used to probe mRNA that was isolated from each sample, and comparisons were made between data from the different strains as well as from different timepoints.

Figure 2.5. Transcriptional profiles of *pca* and *malE*. Signal ratios are shown in $\log_2$-scale. 21253 and 21253(pKD7) were grown on minimal medium containing 20 g/l lactate. Identical control cDNA labeled with Cy5-dUTP was used in each hybridization containing sample cDNA derived from RNA harvested from one culture of 21253 or 21253(pKD7) and labeled with Cy3-dUTP.

As was expected, for all time points sampled, the transcript concentration of *pca* was higher in the strain that had been genetically engineered to over-express its gene. Also, transcripts for *malE* were significantly higher in this strain as well. For the first two time points, the two genes had on average 17-fold more copies of their representative mRNA. In addition, within each strain, the trends for the two genes showed markedly similar profiles, adding further evidence to support the hypothesis suggested by the previously described experiment that *pca* and *malE* are regulated by the cells in a coordinated manner to serve closely related roles within the metabolic network.

Three genes of the lysine biosynthesis pathway—*dapA*, *dapD*, and *lysE*—all had higher transcript levels in samples taken during lysine synthesis than in those taken during early growth. This trend is equally true of samples taken from both strains, and is indicative of a regulatory response made once the inhibitory effects of threonine, which occur upstream to these three genes in the reaction network, are removed. In particular, mRNA for the lysine exporter gene shows a dramatic 14-fold increase in 21253, coincidentally the strain with the higher yield of lysine per cells, and this is the largest time-based increase of any of the genes represented on the microarray.

In the pyruvate-carboxylase-over-expressing strain, there was more mRNA for each of the three genes of the biotin biosynthetic pathway that exist in *C. glutamicum*—*bioA*, *bioB*, and *bioD*—relative to that in the untransformed strain during both growth and lysine production. For pyruvate carboxylase to be active, the enzyme requires biotin as a cofactor, however *C. glutamicum* requires supplementation of its growth medium with biotin because it lacks the enzyme 7,8-diaminopelargonic acid synthetase and cannot synthesize biotin. It therefore seems likely that there is some type of vestigial regulatory scheme at work, providing a link between the higher expression levels of pyruvate carboxylase and the transcriptional up-regulation of those genes that in an ancestral strain might have allowed for increased biotin generation.
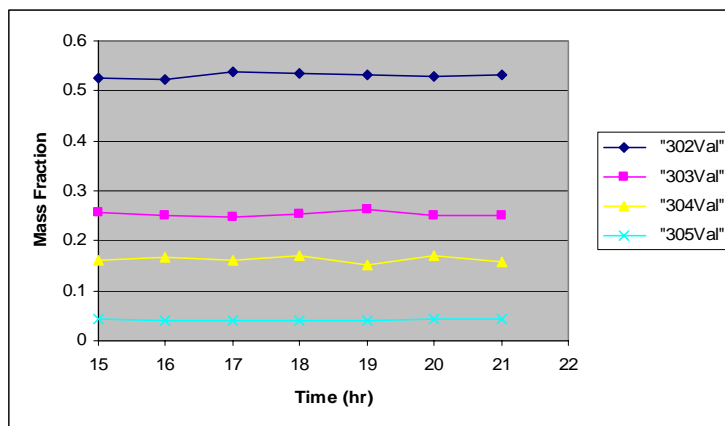
Four genes—*dtsR1*, *dtsR2*, *accC*, and *accD*—with an unconfirmed relationship to amino acid synthesis also exhibited interesting behavior. The genes are believed to be subunits of acyl-coA carboxylases, and their over-expression has been shown to coincide with diminishing production of glutamate. With less TCA cycle intermediates being diverted into the pathways of glutamate synthesis, more is then available to enter those of lysine synthesis. It is noteworthy, then, that all four of these genes are present in higher transcript levels in samples taken during lysine production than in those taken during growth. Also, the transcript concentrations are higher in the 21253 strain, with its higher lysine productivity, than in the 21253(pKD7) strain at all timepoints that were tested.

2.3   Comparisons Between Transcriptional Data and Metabolic Flux Data

The results of the two experiments described above fit in well with prior knowledge and proposed hypotheses describing the functioning of the metabolic network of *C. glutamicum*. The biological significance of these results is not an absolute, though, because the data do not make an attempt to describe to what degree the enzymes being

investigated are regulated through mRNA concentrations. As a result, any correlations made between transcriptional measurements and the activities of enzymes and flux of metabolites is based on assumptions that need to be individually proven. To better link the information gathered about the transcriptome with the actual flow of metabolites through the reaction network, cultures of 21253 and 21253(pKD7) were grown on both glucose and lactate, and samples of these cultures were used to both collect transcriptional data and construct a metabolic flux map.

Unlike samples that are used with DNA microarrays, samples to be used in current techniques of metabolic network analyses must be taken from cultures of cells that are operating under steady-state conditions. This is because the analyses seek to construct a flux map through a combination of isotopic labeling and metabolite balancing, both of which only give a depiction of current *in vivo* conditions when physiological shifts are not occurring. It has been well established that cells that are growing exponentially exist in a metabolic pseudo-steady state. In this experiment we sought to find sampling times for which exponential growth and lysine production were simultaneously taking place, so that we could describe the behavior of the metabolic network as it functions for both of these purposes. Once the period of time overlapping both growth and product formation was identified, the existence of a metabolic pseudo-steady state was confirmed by examining the labeling patterns of amino acids found in hydrolysates of sampled biomass. Representative trends of these labeling patterns, shown in Figure 2.6, demonstrate that the amino acids remained essentially isotopically unchanged during this time interval.
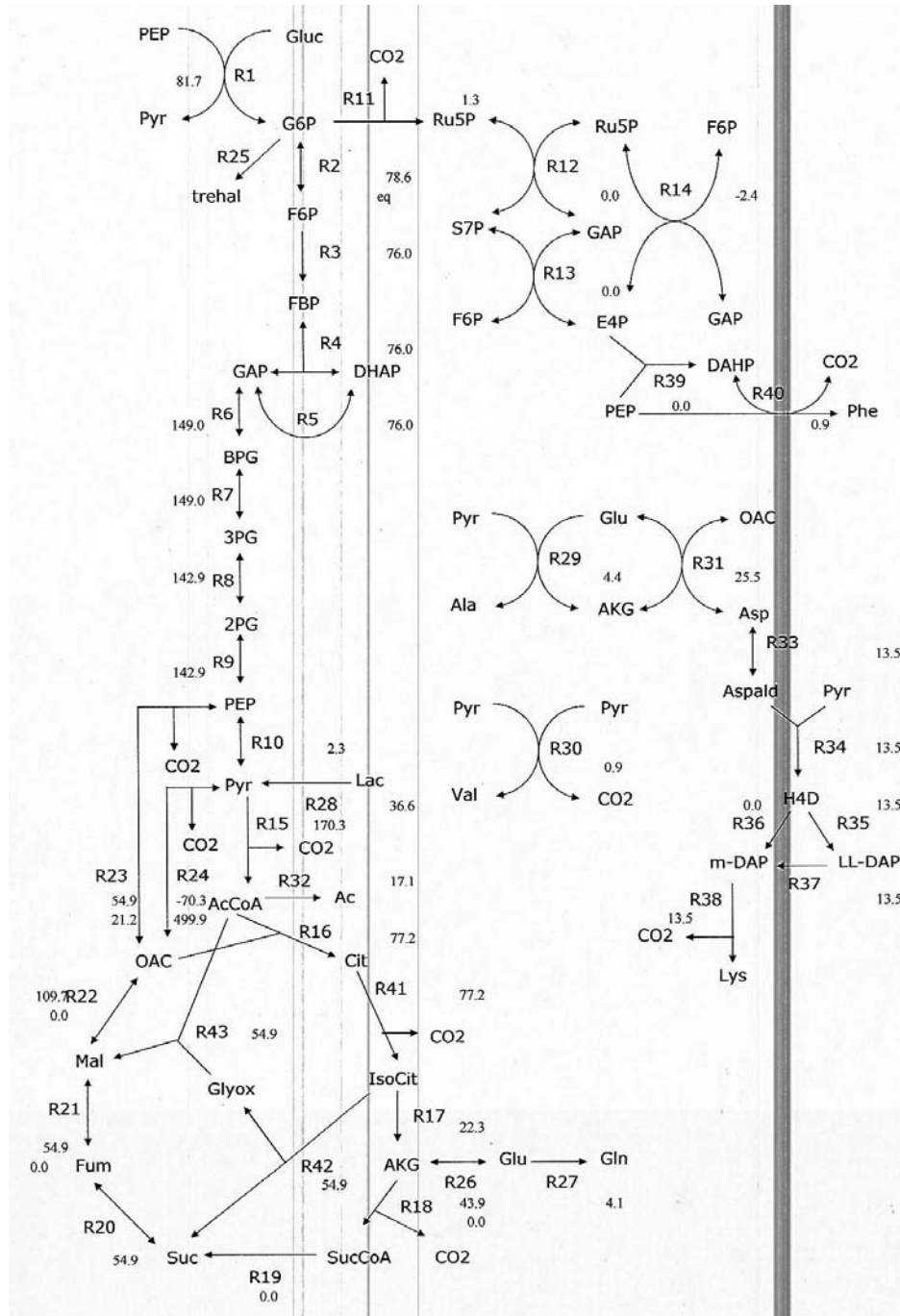

Figure 2.6. Isotopic labeling of valine from *C. glutamicum* undergoing growth and lysine production. Biomass samples taken at 1-hr intervals were hydrolyzed into constituent amino acids which were then derivitized for assaying by GC-MS.

To complete the flux map construction, data from these labeling patterns were combined with measurements of the consumption or production of nutrients, extracellular products, and biomass. The software package, MNA, was then used with the combined data to calculate the fluxes operating in the reactions of a metabolic network that had been predefined. That network, and the associated flux values for one of the four cultures, is shown in Figure 2.7. Together, the four flux maps created represent the behaviors of each of the two strains grown on each of the two carbon sources.
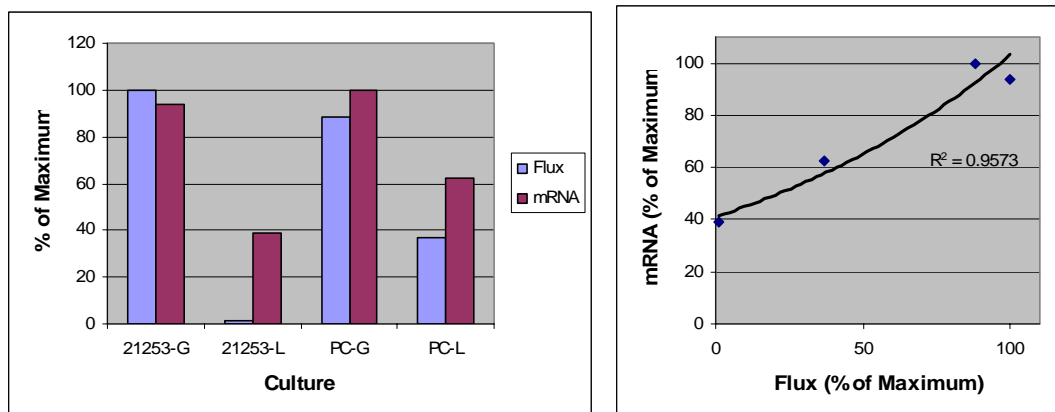
DNA microarrays were used to measure the mRNA concentrations within cell samples taken during the metabolic pseudo-steady states. These concentrations were then compared to the fluxes that were calculated for the reactions catalyzed by the enzymes encoded by the corresponding mRNA sequences. For each mRNA/flux pair, the results were normalized to a percentage of the maximum value determined for the four cultures studied. In this way, we could see the extent to which the upregulation and downregulation of transcript concentrations was related to the metabolite flux. In total, 28 reactions in the network were associated with genes represented on the microarray. Of these, 9 mRNA/flux pairs, or approximately 1/3 of the total, were found to have a poor agreement with one another. Three of these poorly fitting cases—*gpi*, *gpm*, and *fba*—are genes involved in glycolysis and gluconeogenesis. For these three genes, as expected the fluxes decreased for both strains as the glucose carbon source was replaced with lactate. The mRNA concentrations for these genes, though, stayed relatively constant among all four cases tested. Another three genes—*sucA*, *sucC*, and *sdhA*—are for TCA cycle enzymes that catalyze sequentially related reactions. It is possible that these related reactions, as well as the three of glycolysis, are regulated at a level other than transcript concentration, such as that of mRNA translation, protein stability, or enzyme activity.

Although three of the glycolytic reactions did not show a close similarity between mRNA concentration and metabolic flux, the remaining five—*pfk*, *tpi*, *gap*, *pgk*, and *eno*—did very much. For these reactions, again the flux magnitudes showed very significant drops for both strains when growth was based on lactate. Here, however, the

Figure 2.7. Flux map of 21253 grown on lactate. Fluxes are based on metabolite balances and isotopic labeling in cells of a culture in late-exponential growth that is generating lysine and is in a metabolic pseudo-steady state.

drops were matched by the values of the transcript amounts. An example is shown in Figure 2.8. Here, the fluxes, as well as the mRNA levels, were very similar between the strains when grown on glucose. On lactate, the flux in 21253 decreases to an insignificant level and the mRNA abundance decreases as well. In 21253(pKD7) a decrease on lactate is also seen, but the degree is less than that seen in the parental strain, for both the flux

14

Figure 2.8. Relative transcript concentrations and metabolite fluxes for the reaction catalyzed by *pfk*. 21253 and 21253(pKD7) were grown on minimal medium containing 20 g/l carbon source that was either 90% glucose and 10% lactate or 10% glucose and 90% lactate. Identical control cDNA labeled with Cy5-dUTP was used in each hybridization containing sample cDNA derived from RNA harvested from one culture of 21253 or 21253(pKD7) and labeled with Cy3-dUTP.
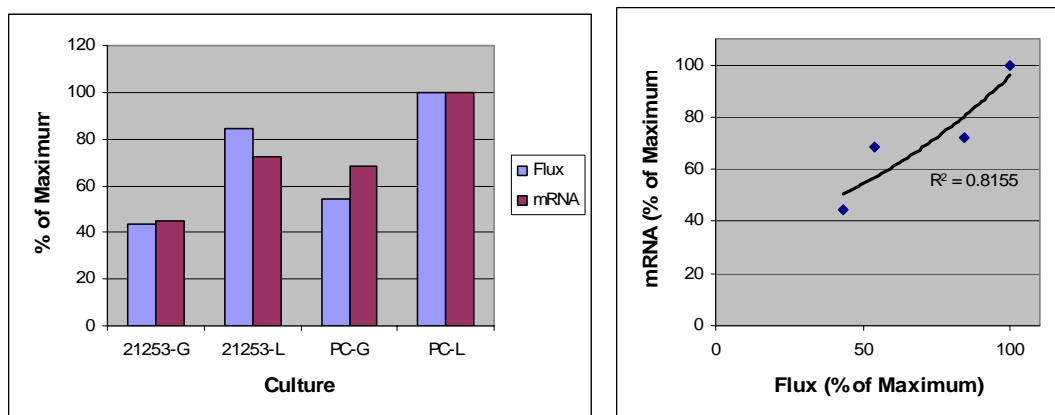
and the mRNA levels. When the relative mRNA amounts and the relative flux values are compared on a scatter plot, also shown in Figure 2.7, the resulting fit is very good, and shows that a baseline amount of gene transcription exists, even when the resulting flux does not take place.

The pentose phosphate pathway, which branches off of the glycolytic pathway, would be expected to also show a drop in carbon flow as glucose is replaced with lactate, which enters the metabolic network downstream of either pathway. Instead, the flow entering the pentose phosphate pathway remains relatively unchanged among the four cultures. This constant flow is matched by an equally steady level of the transcript of the *zwf* gene.

In contrast to the declining fluxes of glycolysis, for four of the genes involved in the TCA cycle—*aco*, *citA*, *icd*, and *fum*—increases were observed as lactate was used as the available carbon source. One example of this pattern is presented in Figure 2.9. In this case, a shift from glucose to lactate coincides with an increased need for the cells to rely on the TCA cycle to produce energy that could otherwise come from glycolysis. Also, as pyruvate carboxylase is over-expressed, there is an increased flow of material into the TCA cycle. Both of these effects can be seen in the closely matched flux and mRNA data for *fum* that is shown in the two graphs.

Other genes for which closely agreeing increases in flux and transcript levels could be seen for both strains upon growth on lactate were:
- *pdh* and *pca*, both of which direct the flow of carbon away from the pyruvate entry point of lactate into central carbon metabolism
- *pck*, which allows carbon to flow from the TCA cycle into gluconeogenesis for the purpose of forming sugar phosphates that are used as cellular building blocks

Figure 2.9. Relative transcript concentrations and metabolite fluxes for the reaction catalyzed by *fum*. 21253 and 21253(pKD7) were grown on minimal medium containing 20 g/l carbon source that was either 90% glucose and 10% lactate or 10% glucose and 90% lactate. Identical control cDNA labeled with Cy5-dUTP was used in each hybridization containing sample cDNA derived from RNA harvested from one culture of 21253 or 21253(pKD7) and labeled with Cy3-dUTP.

- *aceA* and *aceB* of the glyoxylate bypass, which is used to replenish intracellular pools of malate and oxaloacetate that are consumed in the increased TCA cycle flux

Each of the last five steps in the reaction network that lead from aspartate semialdehyde to lysine show trends in which the metabolite fluxes and gene transcript concentrations agree with one another very well. For the first of these reactions, neither the flux nor the mRNA levels of *dapB* change significantly among the four tested cultures. For the last of these reactions, growth on lactate causes the flux and transcription of *lysA* to increase for both strains. This coincides with increased lysine productivities that are seen under this condition. The three reactions that take place between these two constitute a branched pathway. In one path, catalysis is carried out by the product of the gene *dapE*. In the other, the relevant gene is *ddh*. For three of the four cultures, the fluxes and mRNA abundances of these two pairs of reactions and genes remain unchanged. For the culture of 21253(pKD7), though, an increase in both flux and mRNA titer is seen in the *ddh*-pathway, and corresponding decreases are seen in the alternative *dapE*-pathway. This suggests both some conditions for which a switch in the flow of metabolites within the network occurs, and that the switch takes place as a result of regulation at the level of gene transcription. This type of regulatory action is a good example of how DNA microarray analyses provide metabolically meaningful insights into the operation of intracellular networks.