COMMUNICATION TO THE EDITOR

In Silico Design and Adaptive Evolution of *Escherichia coli* for Production of Lactic Acid

Stephen S. Fong,[1] Anthony P. Burgard,[2] Christopher D. Herring,[1] Eric M. Knight,[1] Frederick R. Blattner,[3] Costas D. Maranas,[2] Bernhard O. Palsson[1]

[1]Department of Bioengineering, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0412; telephone: 858-534-5668; fax: 858-822-3120; e-mail: palsson@ucsd.edu
[2]Department of Chemical Engineering, The Pennsylvania State University, University Park, Pennsylvania 16802
[3]Department of Genetics, University of Wisconsin, Madison, Wisconsin 53706

Received 25 January 2005; accepted 9 March 2005

Published online 16 June 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/bit.20542

Abstract: The development and validation of new methods to help direct rational strain design for metabolite overproduction remains an important problem in metabolic engineering. Here we show that computationally predicted *E. coli* strain designs, calculated from a genome-scale metabolic model, can lead to successful production strains and that adaptive evolution of the engineered strains can lead to improved production capabilities. Three strain designs for lactate production were implemented yielding a total of 11 evolved production strains that were used to demonstrate the utility of this integrated approach. Strains grown on 2 g/L glucose at 37°C showed lactate titers ranging from 0.87 to 1.75 g/L and secretion rates that were directly coupled to growth rates.

Keywords: computational model; metabolic engineering; evolution; *Escherichia coli*

INTRODUCTION

Genome-scale metabolic reconstructions are emerging as productive tools for their ability to connect genomic information to phenotypes and to simulate whole-cell physiology as an interconnected system. Using a genome-scale model of *Escherichia coli*, this systems-level approach has been shown to be reasonably accurate in its ability to describe the endpoint growth phenotypes after adaptive evolution for both wild-type (Fong et al., 2003; Ibarra et al., 2002) and single gene deletion strains (Fong and Palsson, 2004) of *E. coli*.

Correspondence to: Bernhard O. Palsson

Contract grant sponsors: The National Institutes of Health; The National Science Foundation

This article includes supplementary material available via the Internet at http://www.interscience.wiley.com/jpages/0006-3592/suppmat.

The advent of genome-scale metabolic models has concurrently spurred the development of numerous analytical tools to study biological systems in silico. In particular, a bi-level optimization algorithm, OptKnock, has been developed and implemented with a genome-scale metabolic model of *E. coli* to computationally predict gene deletion strategies for overproduction of succinate, lactate, 1,3-propanediol (Burgard et al., 2003), and amino acids (Pharkya et al., 2003). The OptKnock algorithm calculates solutions that simultaneously optimize two objective functions, biomass formation (growth rate) and secretion of a target metabolite. The premise underlying this bi-level optimization algorithm is that overproduction of target metabolites can be achieved by altering the structure of the metabolic network through gene deletions such that the stoichiometry of the perturbed network forces production of the target metabolite as normal biomass precursors are generated. With this direct stoichiometric coupling of target metabolite production to biomass, it is hypothesized that an increase in growth rate should concurrently result in an increase in the target metabolite production rate.

In this study, direct experimental testing of these computational methods and the OptKnock hypothesis of increasing production rate with increased growth rate were conducted for production of lactate in *E. coli*. Beginning with a genome-scale metabolic reconstruction of *E. coli* metabolism, the OptKnock algorithm (Burgard et al., 2003) was used to search the phenotypic solutions generated by the metabolic reconstruction to determine all single, double, triple, and quadruple gene deletion designs that would induce *E. coli* to secrete lactic acid as a by-product during optimal cellular growth. Predicted strain designs were constructed, subjected to adaptive evolution to increase growth rates of constructed

strains, and tested for by-product secretion and changes in cellular phenotypes.

MATERIALS AND METHODS

Computer Simulation

The strain optimization framework, OptKnock (Burgard et al., 2003), was used to identify multiple gene deletion combinations that force the coupling of cellular growth objectives with imposed chemical production targets. This coupling is accomplished by ensuring, through gene deletions, that the desired product becomes an obligatory by-product of growth by "shaping" the connectivity of the metabolic network (Fig. 1a). This task yields a bi-level optimization problem which is solved using an efficient solution approach inspired from concepts from LP duality theory as described by Burgard et al. (2003). Using this approach, solutions are found that optimize two objective functions, biomass formation and by-product secretion. In all cases, the OptKnock algorithm indicated that anaerobic conditions were necessary for over-production of lactic acid in *E. coli*. Initial simulations and strain designs were calculated using the iJE660 (Edwards and Palsson, 2000) reconstruction of *E. coli* metabolism. Strain designs and cellular phenotypes were later recalculated using an updated reconstruction of *E. coli*, iJR904 (Reed et al., 2003), after the updated reconstruction had been released.

Strain Construction

The wild-type *E. coli* strain K12 MG1655 (ATCC# 700926) was used as the starting strain for all experiments. The pta-adhE double deletion strain was constructed using the "gene gorging" method (Herring et al., 2003) where mutations were introduced using a donor plasmid containing a I-Sce I restriction site and recombination was facilitated using the lambda Red recombinase. All other gene deletions were conducted using homologous recombination of PCR-amplified linear fragments also using the lambda Red recombinase as described by Datsenko and Wanner (2000). Both gene deletion methods generated in-frame gene deletions with minimal residual scar regions with gene deletions being verified using PCR and physiological characteristics.

Adaptive Evolution

Adaptive evolution was conducted in 100 mL of M9 minimal medium supplemented with 2 g/L glucose at 37°C for 60 days. Cultures were sparged aseptically using nitrogen gas and serially passed into fresh, pre-heated medium daily. Average dilution at the time of passage was 1:100,000 (approximately 1 μL into 100 mL) every 24 h such that approximately 10 doublings occurred each day (exact number of doublings varied throughout evolution as growth rates changed). The inoculum was changed throughout the course of evolution to maintain cultures in exponential growth for each 24 h culture period prior to being passed into

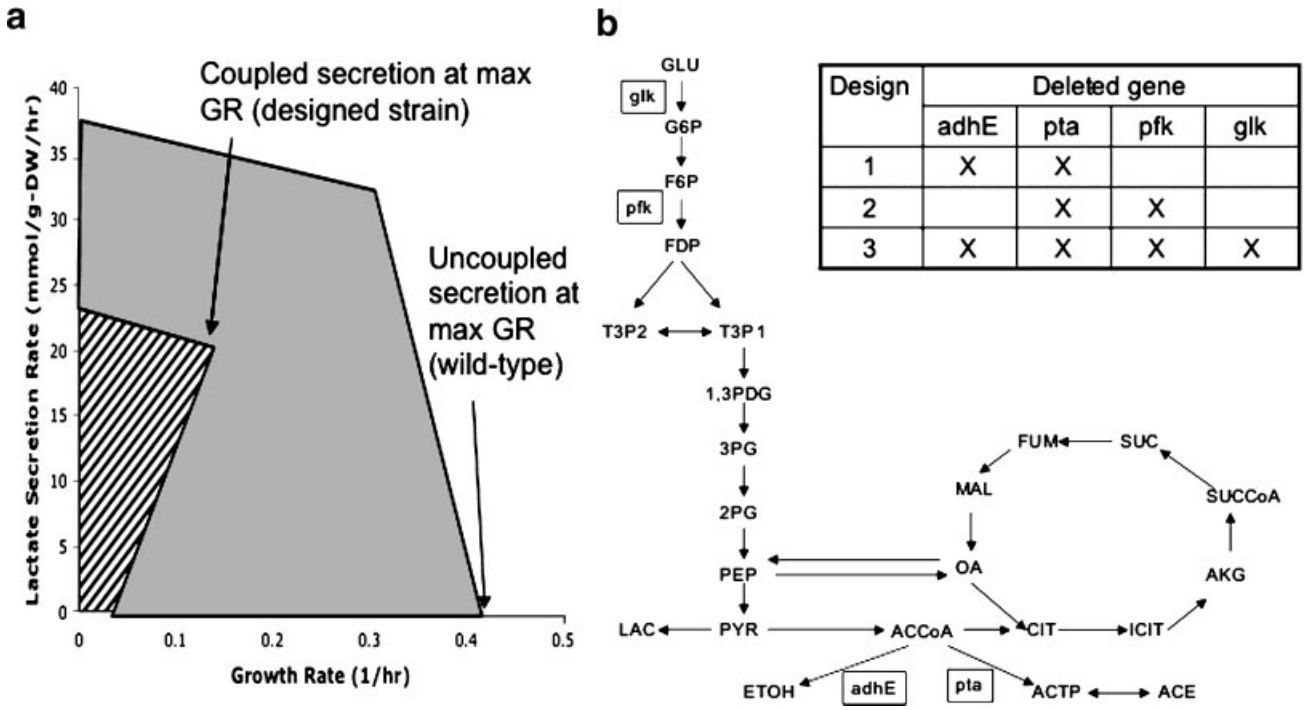

Figure 1. Design strategy for lactic acid secretion. **a**: Computational solution spaces illustrating by-product secretion coupling to biomass for designed strain compared to uncoupled secretion for wild-type strain. **b**: Schematic representation of *E. coli* central metabolism. Genes deleted for different strains in this study are shown in boxes. Inset table indicates genes deleted for each design.

fresh medium. Cultures were frozen and stored at regular intervals throughout adaptive evolution.

Phenotype Assessment

Each strain was tested in a batch culture in M9 minimal medium supplemented with 2 g/L glucose at 37°C to measure growth rate, glucose uptake rate, and by-product secretion rate. Batch cultures were sampled at regular intervals to monitor optical density (A600) for determination of growth rate and to monitor chemical composition of the media for determination of glucose uptake rate and by-product secretion rate. The glucose concentration in the media was measured by enzymatic assay (Sigma). By-product concentrations were determined using UV detection by HPLC using an Aminex87-H ion exchange column at 65°C with 5 mM H_2SO_4 as the mobile phase. Batch cultures were sealed and re-sparged with nitrogen gas at the time of sampling.

RESULTS AND DISCUSSION

Three different designs for production of lactate were selected based upon the OptKnock calculations: (1) *pta-adhE* double deletion strain, (2) *pta-pfk* double deletion strain, and (3) *pta-adhE-pfk-glk* quadruple deletion strain (Fig. 1b). While the *pta-adhE* design is an intuitively logical design that removes pathways for secretion of alternate by-products, deletion of *pfk* and *glk* was non-intuitive. Deletion of *pfk* was computationally suggested to promote secretion of lactate by forcing a rerouting of fluxes through the Entner-Doudoroff pathway and pentose phosphate pathway thus increasing the supply of of NADH and pyruvate. The deletion of *glk* was computationally selected to ensure that the phosphotransferase system (PTS), which converts phosphoenolpyruvate to pyruvate, was maximally utilized for glucose uptake as deletion of *pfk* causes a switch from PTS to glucokinase as the main means of glucose phosphorylation (Roehl and Vinopal, 1976). Both NADH and pyruvate are necessary reactants for production of lactate through lactate dehydrogenase.

Although the OptKnock framework was applied to the iJE660 model (Edwards and Palsson, 2000), an updated *E. coli* stoichiometric model, iJR904 (Reed et al., 2003), has since been published based on more recent annotation information as well as biochemical literature. The newer model includes 904 genes that encode proteins which catalyze 931 unique reactions. The model iJR904 has been tested against iJE660 and exhibited improved predictive power in a number of growth conditions that have been experimentally verified. Accordingly, the iJR904 model has replaced the older version iJE660 for all calculations pertaining to this work. It should be noted, however, that the predicted phenotype of the *pta-pfk* strain showed large changes when calculated using the iJR904 model as compared to initial simulations with the iJE660 model. In this instance, the newer model contradicts the original strain design predictions by suggesting that the growth rate and lactic acid production rate for this strain are not directly coupled to one another. Both models anticipate coupled cell growth and lactic acid production for the *pta-adhE* and *pta-adhE-pfk-glk* designs.

Prior to experimental implementation, it was decided that only the major isozyme, *pfkA*, that accounts for approximately 90% of the phosphofructokinase activity in *E. coli* (Blangy et al., 1968) would be deleted from the latter two designs as deletion of both *pfkA* and *pfkB* is lethal to *E. coli* (Daldal and Fraenkel, 1981; Vinopal and Fraenkel, 1975). These three designs were then constructed from the wild-type K12 MG1655 *E. coli* strain and subjected to adaptive evolution under anaerobic conditions for 60 days at 37°C in glucose minimal medium. Subjection of each strain to adaptive evolution was intended to increase the growth rate of each strain which should concurrently increase the lactate secretion rate (as the two functions are designed to be directly coupled) and also to allow each strain to refine its own metabolic functionality in response to the introduced gene deletions.

Approximately 1,000 generations of growth were achieved in the 60 day period of adaptive evolution (∼900 generations for the *pta-adhE* strains and ∼1100 generations for the *pta-pfk* and *pta-adhE-pfk-glk* strains). Independent parallel evolution cultures of each design were conducted (5 *pta-adhE* cultures, 3 *pta-pfk* cultures, and 3 *pta-adhE-pfk-glk* cultures). Each strain was tested for growth rate, glucose uptake rate, and by-product secretion at 10-day intervals during the course of adaptive. A compilation of these measurements for each individual evolution strain is shown in Table I. Average values for the phenotype measurements of growth rate (1/h), glucose uptake rate (mmol/g-DW.h), and lactate secretion rate (mmol/g-DW.h) for the *pta-adhE* double deletion strains (Fig. 2), the *pta-pfk* double deletion strains (Fig. 3), and the *pta-adhE-pfk-glk* quadruple deletion strains (Fig. 4) are shown with different computational solution spaces corresponding to the different glucose uptake rates found throughout evolution. Direct coupling of lactic acid secretion to growth rate is computationally predicted in the *pta-adhE* and *pta-pfk-glk-adhE* computational solution spaces where growth at the higher growth rates cannot be achieved without at least a minimum amount of lactic acid secretion. Calculation of the *pta-pfk* design with the most recent *E. coli* model was not predicted to strictly couple lactic acid secretion to growth with only strict coupling occurring upon additional deletion of *glk* and *adhE*.

By measuring the growth rates, lactic acid secretion rates, and glucose uptake rates, the experimental phenotypes could be directly compared to the computationally predicted solution spaces for each design. Both the *pta-adhE* strains (Fig. 2a) and the *pta-pfk* strains (Fig. 3a) showed good agreement with the computationally determined solution spaces. The *pta-adhE-pfk-glk* strains (Fig. 4a) consistently exhibited faster growth rates experimentally than were predicted computationally though the lactate secretion rates did fall within the computationally predicted range.

Table I. Phenotype measurements for individual strains on 2 g/L glucose minimal medium with designed evolution strains designated by deleted genes.

Strain	Min GR (1/h)	Max GR (1/h)	Min lactate titer (g/L)	Max lactate titer (g/L)	Min lactate secretion rate (mmol/g-DW.h)	Max lactate secretion rate (mmol/g-DW.h)
Unevolved wild-type		0.48		0.0		0.0
pta-adhE 1	0.19	0.25	1.25	1.66	22.9	30.6
pta-adhE 2	0.19	0.25	1.25	1.75	20.5	30.0
pta-adhE 3	0.19	0.28	1.25	1.67	22.8	34.9
pta-adhE 4	0.18	0.26	1.25	1.69	22.1	30.6
pta-adhE 5	0.16	0.24	1.25	1.66	17.0	31.6
pta-pfk 1	0.09	0.26	0.91	1.32	8.2	16.6
pta-pfk 2	0.09	0.26	0.90	1.32	8.2	20.2
pta-pfk 3	0.09	0.26	0.85	1.32	8.2	17.7
pta-adhE-pfk-glk 1	0.11	0.24	0.88	1.32	9.6	17.4
pta-adhE-pfk-glk 2	0.12	0.25	0.87	1.32	10.5	18.6
pta-adhE-pfk-glk 3	0.12	0.26	0.87	1.32	10.6	19.7

Minimum (Min) and maximum (Max) values for each parameter are the minimum and maximum values occurring for a given strain throughout the course of evolution. Note: main by-products of wild-type strain under anaerobic growth are acetate, formate, and ethanol. Error of less than 5% in all parameters was determined through replicate testing.

In all strains, increases in growth rate occurred over the course of adaptive evolution. The *pta-adhE* strains showed an average increase in growth rate of 21% (Fig. 2a), the *pta-pfk* strains had an average growth rate increase of 133% (Fig. 3a), and the *pta-adhE-pfk-glk* strains showed an average growth rate increase of 75% (Fig. 4a). One of the most important criteria used for selecting the deletion designs was that the design should directly couple lactic acid secretion with growth rate such that an increase in growth rate would result in an increase in the lactic acid secretion rate meaning that the growth and secretion objectives were directly stoichiometrically coupled. This coupling was shown to be true for all three constructed designs as increases in lactic acid secretion were observed concurrently with the increases in growth rate. The *pta-adhE* strains showed an average increase in secretion rate of 25% (Fig. 2a), the *pta-pfk* strains averaged an increase in lactic acid secretion of 73% (Fig. 3a), and the *pta-adhE-pfk-glk* strains showed an average increase of 55% (Fig. 4a).

Although all evolved strains showed an increase in lactic acid secretion rate over the course of adaptive evolution, the

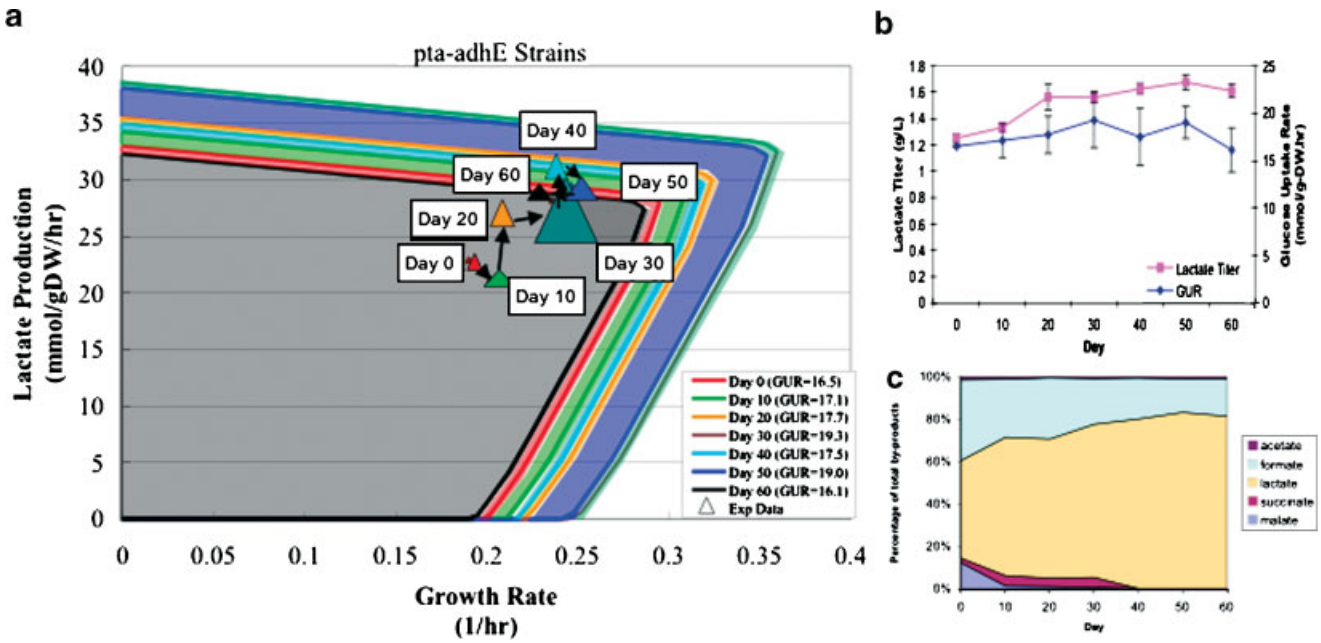

Figure 2. Phenotype of *pta-adhE* deletion strains during adaptive evolution. **a**: Average values of growth rate (1/h), lactate secretion rate (mmol/g-DW.h), and glucose uptake rate (GUR, mmol/g-DW.h) are shown every 10 days of evolution on computationally predicted solution spaces that are calculated based upon the experimental glucose uptake rate. Colors of experimental points match the correspondingly colored solution space calculated using measured glucose uptake rates. Variance between the independently evolved strains is indicated by the base and height of each point. **b**: Maximum lactate titer (g/L) and glucose uptake rate over the course of adaptive evolution. Error bars indicate variance between the independently evolved strains. **c**: Proportions of all measured by-products (% of total by-products) over the course of adaptive evolution.

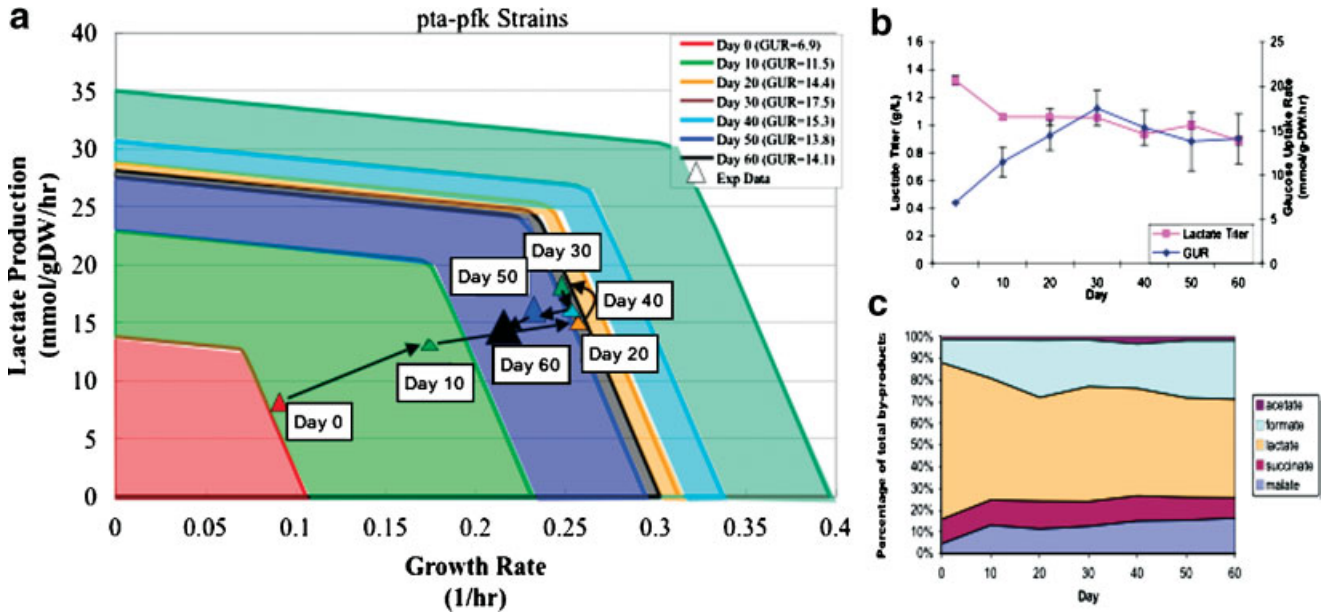

Figure 3. Phenotype of *pta-pfk* deletion strains during adaptive evolution. **a**: Average values of growth rate (1/h), lactate secretion rate (mmol/g-DW.h), and glucose uptake rate (GUR, mmol/g-DW.h) are shown every 10 days of evolution on computationally predicted solution spaces that are calculated based upon the experimental glucose uptake rate. Colors of experimental points match the correspondingly colored solution space calculated using measured glucose uptake rates. Variance between the independently evolved strains is indicated by the base and height of each point. **b**: Maximum lactate titer (g/L) and glucose uptake rate over the course of adaptive evolution. Error bars indicate variance between the independently evolved strains. **c**: Proportions of all measured by-products (% of total by-products) over the course of adaptive evolution.

overall lactic acid titer did not increase for all strains. The *pta-adhE* strains did exhibit an increase (~35%) in lactic acid titer (Fig. 2b) with a final maximum concentration of 1.69 g/L. The increase in lactic acid titer for this strain was largely attributable to the decrease in secretion of other secondary by-products such as formate, malate, and succinate (Fig. 2c). Both the *pta-pfk* strains and the *pta-adhE-pfk-glk* strains showed decreases in lactic acid titer at the end of adaptive evolution to final maximum concentrations of 0.89 g/L (Fig. 3b) and 0.87 g/L (Fig. 4b), respectively. In both cases,

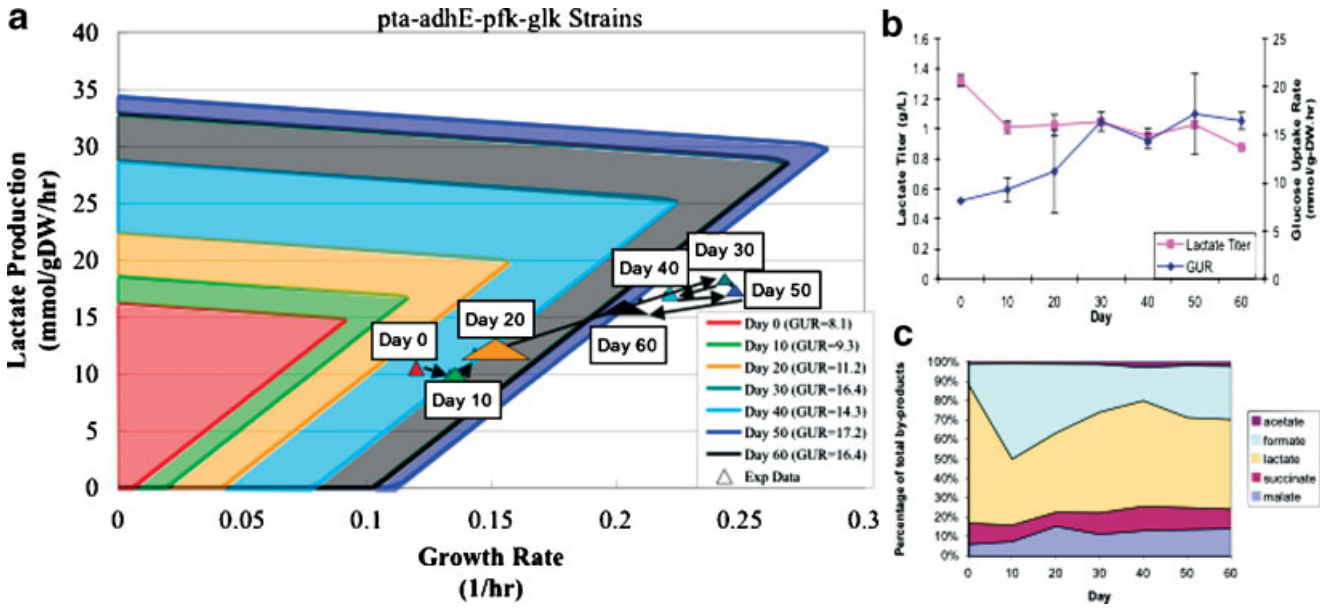

Figure 4. Phenotype of *pta-adhE-pfk-glk* deletion strains during adaptive evolution. **a**: Average values of growth rate (1/h), lactate secretion rate (mmol/g-DW.h), and glucose uptake rate (GUR, mmol/g-DW.h) are shown every 10 days of evolution on computationally predicted solution spaces that are calculated based upon the experimental glucose uptake rate. Colors of experimental points match the correspondingly colored solution space calculated using measured glucose uptake rates. Variance between the independently evolved strains is indicated by the base and height of each point. **b**: Maximum lactate titer (g/L) and glucose uptake rate over the course of adaptive evolution. Error bars indicate variance between the independently evolved strains. **c**: Proportions of all measured by-products (% of total by-products) over the course of adaptive evolution.

secondary by-product secretion increased (Figs. 3c and 4c) with formate being secreted as the most abundant secondary by-product. In all cases, the by-product secretion profiles stabilized after approximately 20 days of adaptive evolution with all strains showing sustained elevated lactic acid titers throughout the course of adaptive evolution over the wild-type strain (secretes trace amounts of lactic acid).

Overall, significant secretion of lactic acid for the duration of adaptive evolution and increases in growth rate and lactic acid secretion rate throughout evolution were observed for all evolved strains. Through comparison of the different designs, the *pfkA* gene deletion appeared to have a dominant effect on the cellular phenotype, such that the phenotype of the *pta-pfk-glk-adhE* deletion strain more closely resembled the *pta-pfk* deletion strain rather than the *pta-adhE* deletion strain. It has been shown that deletion of the *pfk* gene causes intracellular accumulation of fructose 6-phosphate which causes the degradation of *ptsG* mRNA, thus inhibiting glucose transport (Morita et al., 2003). Given this consideration and the apparent dominating effect of the gene deletion, it is suggested that future strain designs should avoid deletion of the *pfk* gene if alternate deletion strategies can be found.

While engineered strains of *E. coli* already exist that produce lactic acid (Chang et al., 1999; Dien et al., 2001; Zhou et al., 2003) and one of the designs here (*pta-adhE* mutant) has been previously described (Gupta and Clark, 1989), the goal of this study was to experimentally test computationally predicted strain designs calculated from a genome-scale metabolic model using the OptKnock algorithm. For the designs generated, it was shown that this combination of computational approaches can prospectively and effectively calculate strain designs for over-production of lactic acid. Adaptive evolution of the different designed strains showed that: (1) the computationally predicted phenotypes are experimentally reproducible and consistent, (2) the process of adaptive evolution leads to increased secretion rates of a target metabolite and can lead to improved product titers, and (3) the generation of stable production strains can be achieved through this method. Overall, all 11 evolved strains exhibited secretion profiles that supported the OptKnock hypothesis that metabolite over-production can be stoichiometrically coupled to biomass generation. This was demonstrated through the increased lactic acid secretion rates that were found concurrent with increased growth rates during evolution. While these examples successfully coupled these two objectives, evolution did not always lead to increased product titer. This may be attributed to specific evolutionary changes and transcriptional regulation that were not accounted for in this study. Thus, the described method represents a systems-level approach to metabolic engineering that, in the future and through further refinement, has the potential to be extended to applications involving over-production of different metabolites, gene additions, and also different organisms.

References

Blangy D, Buc H, Monod J. 1968. Kinetics of the allosteric interactions of phosphofructokinase from *Escherichia coli*. J Mol Biol 31:13–35.

Burgard AP, Pharkya P, Maranas CD. 2003. Optknock: A bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. Biotechnol Bioeng 84:647–657.

Chang DE, Jung HC, Rhee JS, Pan JG. 1999. Homofermentative production of D- or L-lactate in metabolically engineered *Escherichia coli* RR1. Appl Environ Microbiol 65:1384–1389.

Daldal F, Fraenkel DG. 1981. Tn10 insertions in the pfkB region of *Escherichia coli*. J Bacteriol 147:935–943.

Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci USA 97:6640–6645.

Dien BS, Nichols NN, Bothast RJ. 2001. Recombinant *Escherichia coli* engineered for production of L-lactic acid from hexose and pentose sugars. J Ind Microbiol Biotechnol 27:259–264.

Edwards JS, Palsson BO. 2000. The *Escherichia coli* MG1655 in silico metabolic genotype: Its definition, characteristics, and capabilities. Proc Natl Acad Sci 97:5528–5533.

Fong SS, Marciniak JY, Palsson BO. 2003. Description and interpretation of adaptive evolution of *Escherichia coli* K-12 MG1655 by using a genome-scale in silico metabolic model. J Bacteriol 185:6400–6408.

Fong SS, Palsson BO. 2004. Metabolic gene-deletion strains of *Escherichia coli* evolve to computationally predicted growth phenotypes. Nat Genet 36:1056–1058.

Gupta S, Clark DP. 1989. *Escherichia coli* derivatives lacking both alcohol dehydrogenase and phosphotransacetylase grow anaerobically by lactate fermentation. J Bacteriol 171:3650–3655.

Herring CD, Glasner JD, Blattner FR. 2003. Gene replacement without selection: Regulated suppression of amber mutations in *Escherichia coli*. Gene 311:153–163.

Ibarra RU, Edwards JS, Palsson BO. 2002. *Escherichia coli* K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. Nature 420:186–189.

Morita T, El-Kazzaz W, Tanaka Y, Inada T, Aiba H. 2003. Accumulation of glucose 6-phosphate or fructose 6-phosphate is responsible for destabilization of glucose transporter mRNA in *Escherichia coli*. J Biol Chem 278:15608–15614.

Pharkya P, Burgard AP, Maranas CD. 2003. Exploring the overproduction of amino acids using the bilevel optimization framework OptKnock. Biotechnol Bioeng 84:887–899.

Reed JL, Vo TD, Schilling CH, Palsson BO. 2003. An expanded genome-scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR). Genome Biol 4:R54.

Roehl RA, Vinopal RT. 1976. Lack of glucose phosphotransferase function in phosphofructokinase mutants of *Escherichia coli*. J Bacteriol 126: 852–860.

Vinopal RT, Fraenkel DG. 1975. PfkB and pfkC loci of *Escherichia coli*. J Bacteriol 122:1153–1161.

Zhou S, Causey TB, Hasona A, Shanmugam KT, Ingram LO. 2003. Production of optically pure D-lactic acid in mineral salts medium by metabolically engineered *Escherichia coli* W3110. Appl Environ Microbiol 69:399–407.