

Genome-Scale Model for *Clostridium acetobutylicum*: Part II. Development of Specific Proton Flux States and Numerically Determined Sub-Systems

Ryan S. Senger,^{1,2,3} Eleftherios T. Papoutsakis^{1,2,3}

¹Delaware Biotechnology Institute, University of Delaware, 15 Innovation Way, Newark, Delaware 19711; telephone: 302-831-8376; fax: 302-831-4841; e-mail: senger@dbi.udel.edu

²Department of Chemical Engineering, Colburn Laboratory, University of Delaware, Newark, Delaware

³Department of Chemical and Biological Engineering; Northwestern University, Evanston, Illinois

Received 14 December 2007; revision received 23 May 2008; accepted 28 May 2008

Published online 5 June 2008 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bit.22009

ABSTRACT: A regulated genome-scale model for *Clostridium acetobutylicum* ATCC 824 was developed based on its metabolic network reconstruction. To aid model convergence and limit the number of flux-vector possible solutions (the size of the phenotypic solution space), modeling strategies were developed to impose a new type of constraint at the endo–exo-metabolome interface. This constraint is termed the *specific proton flux state*, and its use enabled accurate prediction of the extracellular medium pH during vegetative growth of batch cultures. The *specific proton flux* refers to the influx or efflux of free protons (per unit biomass) across the cell membrane. A specific proton flux state encompasses a defined range of specific proton fluxes and includes all metabolic flux distributions resulting in a specific proton flux within this range. Effective simulation of time-course batch fermentation required the use of independent flux balance solutions from an optimum set of specific proton flux states. Using a real-coded genetic algorithm to optimize temporal bounds of specific proton flux states, we show that six separate specific proton flux states are required to model vegetative-growth metabolism and accurately predict the extracellular medium pH. Further, we define the *apparent proton flux stoichiometry* per weak acids efflux and show that this value decreases from ~ 3.5 mol of protons secreted per mole of weak acids at the start of the culture to ~ 0 at the end of vegetative growth. Calculations revealed that when specific weak acids production is maximized in vegetative growth, the net proton exchange between the cell and environment occurs primarily through weak acids efflux (apparent proton flux stoichiometry is 1). However, proton efflux through cation channels during the

early stages of acidogenesis was found to be significant. We have also developed the concept of *numerically determined sub-systems* of genome-scale metabolic networks here as a sub-network with a one-dimensional null space basis set. A numerically determined sub-system was constructed in the genome-scale metabolic network to study the flux magnitudes and directions of acetylornithine transaminase, alanine racemase, and D-alanine transaminase. These results were then used to establish additional constraints for the genome-scale model.

Biotechnol. Bioeng. 2008;101: 1053–1071.

© 2008 Wiley Periodicals, Inc.

KEYWORDS: *Clostridium acetobutylicum*; metabolic flux analysis; proton flux state; numerically determined sub-system; genome-scale model

Introduction

Clostridium acetobutylicum ATCC 824 is a strict anaerobe that undergoes an acidogenic phase of vegetative growth followed by acid re-uptake, solventogenesis, and sporulation in the later stages of the culture (Husemann and Papoutsakis, 1988; Jones and Woods, 1986; Monot et al., 1982; Papoutsakis and Meyer, 1985a; Roos et al., 1985). To generate a regulated genome-scale model of an organism in which differentiation involves a cascading expression of sigma-factors (Paredes et al., 2005), a model describing the metabolic events (including vegetative growth) leading up to the expression of the first sigma-factor of the cascade (Spo0A in *C. acetobutylicum* (Alsaker et al., 2004; Harris et al., 2002; Wilkinson et al., 1995)) is essential. The primary metabolism of *C. acetobutylicum* has been extensively

Correspondence to: R.S. Senger

Contract grant sponsor: NIH

Contract grant number: F32GM078947

Additional Supporting Information may be found in the online version of this article.

studied and has been further characterized by the first flux balance analysis (Papoutsakis, 1984; Papoutsakis and Meyer, 1985a,b). Further developments addressed a key singularity of the metabolic network and model through the use of a non-linear constraint (Desai et al., 1999a,b). In this article, the model of primary metabolism was combined with other resolved metabolic pathways of macromolecular biosynthesis and cell proliferation to develop a genome-scale model for the vegetative growth phase of *C. acetobutylicum*. The metabolic reaction network reconstruction was presented in Part 1 of this work (Senger and Papoutsakis, 2008).

Using our metabolic network of *C. acetobutylicum*, we present a genome-scale model and a constraint that originates at the endo–exo-metabolome interface and becomes further identified at the fluxome-level of regulation (Nielsen and Oliver, 2005). To do this, we introduce the concept of *specific proton flux states* into constraints-based optimization in metabolic flux analysis. The specific proton flux state is defined by a range of allowable specific proton flux values across the cell membrane. Flux solutions located in the phenotypic solution space that conform to a specific proton flux state contain intracellular and membrane transport specific flux values that are uniquely defined. Also, using the genome-scale reconstructed metabolic network of *C. acetobutylicum*, we present a study of metabolic flux capacity using a sub-network we call a *numerically determined sub-system*. We defined a sub-network in the underdetermined genome-scale stoichiometric matrix that contains a one-dimensional null space basis set. Through the assumption of various growth rates, we (i) demonstrate how numerically determined sub-systems can further identify the phenotypic solution space around matrix singularities and (ii) determine the metabolic impact of experimentally uncharacterized physiological processes.

Methods

Genome-Scale Model of *C. acetobutylicum*

The metabolic network reconstruction for *C. acetobutylicum* ATCC 824 was described in Part 1 of this work (Senger and Papoutsakis, 2008). The reconciled metabolic network for *C. acetobutylicum* includes 422 intracellular metabolites involved in 552 reactions, including 80 membrane transport reactions. The model was composed based on the “biochemical reaction” formalism (e.g., $\text{ATP} + \text{H}_2\text{O} \leftrightarrow \text{ADP} + \text{Orthophosphate}$) as opposed to a charge-balanced “chemical reaction” formalism ($\text{ATP}^{4-} + \text{H}_2\text{O} \leftrightarrow \text{ADP}^{3-} + \text{HPO}_4^{2-} + \text{H}^+$) since intracellular pH, temperature, pressure, and ionic strength were held constant in the model (Alberty, 1991, 1993). The use of biochemical equations in the model was later verified by a hydrogen atom balance around the cell boundary with membrane transport and biomass constituting equations written with the chemical reaction formalism. The full list of reactions of the metabolic network reconstruction was given in Part 1 of this series (Senger and Papoutsakis, 2008). The

genome-scale model and an updated set of reaction constraints are listed in Supplementary Appendix 1 of this manuscript. The model contains many more constrained and irreversible reactions than did the previous version. Constraints were applied based on the irreversibility (or known direction) of metabolic reactions as well as applying order-of-magnitude approximations to constraints. For example dGTP is produced by the metabolic network because of its inclusion in the DNA biosynthesis equation, which is a component of the biomass constituting equation (see Supplementary Appendix 1). If a maximum growth rate of 0.5 h^{-1} is assumed, the maximum production rate of dGTP is equal to the growth rate multiplied by the stoichiometric coefficient of DNA in the biomass constituting equation (0.03) multiplied by the stoichiometric coefficient of dGTP in the DNA equation (0.33) to yield a required dGTP biosynthesis flux of $5 \times 10^{-3} \text{ mmol h}^{-1} \text{ g biomass}^{-1}$. The corresponding stoichiometric matrix was constructed using MATLAB[®] (The Mathworks, Inc.; Natick, MA) and the flux balance equation, $S \cdot v = 0$, was solved using LINDO API (Lindo Systems, Inc., Chicago, IL). We explored the objective function of maximizing the specific growth rate, but we found that maximizing the rate of reduced ferredoxins production yielded a ratio of H_2/CO_2 formation closer to experimentally observed values (McLaughlin et al., 1985; Vasconcelos et al., 1994). In *C. acetobutylicum*, NADPH regeneration is achieved through oxidation of the ferredoxins instead of through the pentose phosphate pathway.

Sampling and Reduction of the Phenotypic Solution Space

Optimization of the flux vector, v , of the flux balance equation is an underdetermined problem for a genome-scale metabolic network with more reactions than metabolites. Thus, multiple solutions of flux distributions to the flux balance equation exist upon constrained optimization by linear programming. The solution or family of flux vector solutions, v , corresponding to an observable phenotype resides within the multidimensional convex polytope that is the phenotypic solution space. Methods of constraining and sampling of the phenotypic solution space to yield computational phenotypes consistent with experimental observations has been of interest to recent computational research (Choi et al., 2007; Covert et al., 2003; Famili et al., 2005; Price et al., 2004; Wiback et al., 2004). In the following sections, we present novel algorithms for developing additional constraints to the phenotypic solution space to reduce the set of possible flux distributions. To obtain a representative flux distribution from the constrained phenotypic solution space, a stochastic sampling algorithm was applied (Wiback et al., 2004), and multiple results for individual fluxes were averaged. The constrained solution space was sampled 10^2 times to obtain a representative set of fluxes, v . The set of membrane transport equation with their

upper and lower constraints that were varied stochastically to probe the phenotypic solution space are provided in Supplementary Appendix 2. These particular transport reactions were chosen because they directly determine the specific proton flux across the cell membrane. Their constraints were chosen so that specific proton flux ranges from $-100 \text{ mmol H}^+ \text{ h}^{-1} \text{ g biomass}^{-1}$ (efflux) to $5 \text{ mmol H}^+ \text{ h}^{-1} \text{ g biomass}^{-1}$ (influx) could be explored. The averaged flux values for CO_2 formation given different sampling sizes and proton flux states are shown in Supplementary Appendix 3.

The development of tools to further reduce the size of the phenotypic solution space (and number of possible solutions) is the subject of much on-going research. Here, we propose two additional methods for constraining the resulting phenotypic solution space of flux balance analysis (FBA) solutions: (i) defining a specific proton flux state and examining its relationship with extracellular medium pH changes and (ii) generating sub-networks with a one-dimensional null space vector (we call these *numerically determined sub-systems*) through the addition of reaction flux relationships. We argue that the two proposed methods are particularly useful when dealing with organisms (such as the clostridia) that lack the wealth of literature data and well-developed experimental tools for constructing genetic perturbations (MacCarthy et al., 2005; Reed et al., 2006; Tegner et al., 2003) to determine flux constraints and probe metabolic capacity.

Calculation of the Specific Proton Flux

Several studies of *C. acetobutylicum* fermentation kinetics (Husemann and Papoutsakis, 1988; Roos et al., 1985), transcriptional data (Alsaker and Papoutsakis, 2005; Alsaker et al., 2004; Tummala et al., 2003) and fluxes of the primary metabolic pathways (Desai et al., 1999a,b; Papoutsakis, 1984; Papoutsakis and Meyer, 1985a,b) have shown that the rates of butyric and acetic acids production and consumption (re-uptake) vary widely over the course of a batch fermentation. Here, we propose that the rates at which free protons are excreted by the cell can be used as an identifier of dominant metabolic programs and can be used to reduce the phenotypic solution space. To do this, we consider the *specific proton flux* ($q_{\text{H}^+_{\text{ext}}}$), which is defined as the flux of free protons exchanged between the cell and the extracellular environment (a negative flux corresponds to the excretion of protons). The proton exchange reaction across the cellular membrane is represented as,

$$\frac{d\text{H}^+_{\text{ext}}}{dt} = q_{\text{H}^+_{\text{ext}}} X \quad (1)$$

where H^+_{ext} is the total (not necessarily free) extracellular hydrogen ion concentration and X is the biomass concentration. We define the extracellular hydrogen ion concentration as the sum of (i) the extracellular free proton

concentration, H^+_{free} (calculated from medium pH, $\text{pH}_{\text{extracellular}} = -\log(\text{H}^+_{\text{free}})$), and (ii) those hydrogen ions associated with weak acids given specific medium pH and $\text{p}K_a$ values. For example, the total extracellular hydrogen ion concentration (H^+_{ext}) is defined in Equation (2), for the minimal medium (Monot et al., 1982), as the summation of the extracellular molar concentrations of: (i) free protons (H^+_{free}), (ii) protonated butyrate (butyric acid) (H_{butyrate}), (iii) acetic acid (H_{acetate}), (iv) lactic acid (H_{lactate}), (v) carbonic acid ($H_{\text{carbonate}}$), (vi) ammonium ion (H_{ammonium}), and (vii) monobasic and dibasic potassium phosphates ($H_{\text{phosphates}}$).

$$\text{H}^+_{\text{ext}} = \text{H}^+_{\text{free}} + H_{\text{butyrate}} + H_{\text{acetate}} + H_{\text{lactate}} + H_{\text{carbonate}} + H_{\text{phosphates}} + H_{\text{ammonium}} \quad (2)$$

We recognize this approximation of the extracellular proton concentration excludes other minor proton sources/sinks. However, these species were chosen to calculate the specific proton flux as they are (i) abundant in the minimal medium and (ii) contain $\text{p}K_a$ values within the operating pH range of batch fermentation. Thus, consideration of these species is particularly useful in approximating the extracellular hydrogen ion concentration from minimal media fermentation data. Given a pH value and the $\text{p}K_a$ of a weak acid, the fraction of protonated and unprotonated species can be directly calculated using the Henderson-Hasselbalch equation.

The specific proton flux ($q_{\text{H}^+_{\text{ext}}}$) was calculated directly for the genome-scale model. This was done through the summation of specific fluxes over all membrane transport exchange reactions protons (M in total), where q_r is the flux of a reaction exchanging proton(s) with the extracellular environment and h_r is the stoichiometric coefficient of H^+ in the membrane transport reaction.

$$q_{\text{H}^+_{\text{ext}}} = \sum_{r=1}^M h_r q_r \quad (3)$$

The list of membrane transport equations and related stoichiometry in the genome-scale model considered in the calculation of the specific proton flux is shown as Table I. The Transport Classification (TC) numbers of each reaction were obtained from the Transport Classification Database (TCDB) (Busch and Saier, 2002) and are listed in Table I.

Calculation of Extracellular Medium pH

To effectively study the significance of specific proton flux states of the genome-scale metabolic model, calculation of the extracellular medium pH was required. A semi-mechanistic, partial buffering model (Dougherty et al., 2006) was derived from a charge balance and adapted for a minimal medium (Monot et al., 1982) supporting

Table 1. Transport reactions with contribution to the proton flux state of the culture for growth in minimal media.

TC ^a	Transport Reaction ^{b,c,d,e}	Constrained?
2.A.1	Butyric acid → butyrate (ext.) + H ⁺ (ext.)	Yes
2.A.1	Acetic acid → acetate (ext.) + H ⁺ (ext.)	Yes
2.A.14	Lactic acid → lactate (ext.) + H ⁺ (ext.)	Yes
2.A.88	Biotin (ext.) + H ⁺ (ext.) → biotin + H ⁺	Yes
2.A.46	4-aminobenzoate (ext.) + H ⁺ (ext.) → 4-aminobenzoate + H ⁺	Yes
3.A.2	3 H ⁺ (ext.) + ADP + orthophosphate ↔ 3 H ⁺ + ATP + H ₂ O	No
2.A.35	K ⁺ (ext.) + H ⁺ (ext.) ↔ K ⁺ + H ⁺	No
2.A.37	Na ⁺ (ext.) + H ⁺ ↔ Na ⁺ + H ⁺ (ext.)	No
3.A.3	3 Na ⁺ + 2 K ⁺ (ext.) + ATP + H ₂ O ↔ 3 Na ⁺ (ext.) + 2 K ⁺ + ADP + orthophosphate	Yes

^aTransporter classification (TC) families from the Transport Classification Database (Busch and Saier 2002) are given for each exchange reaction.

^bAll metabolites are located within the cell well unless specifically labeled as “extracellular” (“ext.”).

^cTransport reactions were either constrained to the forward direction (→) or were left unconstrained with respect to direction (↔).

^dProtonated species of weak acids (e.g., butyric acid, C₄H₈O₂) are secreted by *C. acetobutylicum* and dissociate into the ionic species in the extracellular medium as dictated by pK_a and medium pH values.

^eThe P-type ATPase (TC 3.A.3) was constrained to a flux value of 1 since it and the F₁F₀ ATPase (TC 3.A.2) create a singularity when both are unconstrained.

C. acetobutylicum growth. It is shown as Equation (4), and model parameters are defined in Nomenclature.

$$\begin{aligned}
 & \sum_{\text{Acids}} C_A \frac{\sum_{i=1}^D \left[(d+1-i) [\text{H}_{\text{free}}^+]^{i-1} \prod_{j=1}^{D+1-i} K_{aj} \right]}{\sum_{i=1}^{D+1} \left([\text{H}_{\text{free}}^+]^{(i-1)} \prod_{j=1}^{D+1-i} K_{aj} \right)} + \frac{K_w}{\text{H}_{\text{free}}^+} - \text{H}_{\text{free}}^+ \dots \\
 & - \sum_{\text{Bases}} C_B \frac{\sum_{i=1}^D \left[(d+1-i) [\text{H}_{\text{free}}^+]^{i-1} \prod_{j=1}^{D+1-i} K_{aj} \right]}{\sum_{i=1}^{D+1} \left([\text{H}_{\text{free}}^+]^{(i-1)} \prod_{j=1}^{D+1-i} K_{aj} \right)} - \sum_{\text{Counterions}} C_C \sum_{k=1}^{n_c} z_k = 0
 \end{aligned} \tag{4}$$

The weak acids (C_A) included in the model consisted of those of the initial media formulation and those produced/ consumed during fermentation: acetic acid (pK_a = 4.76), butyric acid (pK_a = 4.83), lactic acid (pK_a = 3.08), carbonic acid (pK_{a,1} = 6.35, pK_{a,2} = 10.33), monobasic and dibasic potassium phosphate (pK_{a,1} = 2.15, pK_{a,2} = 7.20, pK_{a,3} = 12.35). The weak base (C_B) considered by the pH model was ammonium ion (pK_a = 9.25). Monobasic and dibasic potassium phosphates were treated as electrolyte-associated buffers; thus, the charge balance pH model included a term to account for potassium counterions (C_C) as fully described by the authors in development of the pH model (Dougherty et al., 2006). The concentration of carbonic acid was held constant and was based on CO₂ solubility in fermentation broth (Gros et al., 1999). The pH model (Eq. 4) was solved for H_{free}⁺ (pH_{extracellular} = -log(H_{free}⁺)) using a numerical root-finding algorithm.

Consideration of a Specific Proton Flux State in Modeling Metabolism

From direct calculations of intracellular fluxes in the primary metabolism of *C. acetobutylicum*, it has long been known that these fluxes do not remain constant during the

vegetative stage of growth for batch cultures (Desai et al., 1999a,b; Papoutsakis, 1984). In addition, calculation of the specific proton flux from fermentation data (using Eqs. 1 and 2) yields a continuous function over the time-course of vegetative growth due to the pH dependence associated with end-product efflux of weak acids. However, the solution of the flux balance equation, S·v = 0, yields a single set of intracellular and membrane transport fluxes. To address this problem, much research has been dedicated to the development of constraint-based regulatory networks (Covert et al., 2003; Gianchandani et al., 2006). With time-dependent flux profiles, a large (if not infinite) number of genome-scale flux profiles is required to model the vegetative growth stage of *C. acetobutylicum*. We argue that the specific proton flux can be used to discretize this continuous time-dependent set of fluxes into a manageable number useful for time-course growth predictions by the genome-scale model. To do this, we define multiple specific proton flux states over the course of a batch fermentation in order to limit the available phenotypic space. This is further illustrated in Figure 1. Only a limited set of possible intracellular and membrane transport fluxes (a sub-space of the phenotypic solution space) result in a specific proton flux (q_{H_{ext}⁺}) within a specified range. Thus, the specific proton flux is an additional parameter capable of limiting the number of possible flux solutions to the flux balance equation. We argue that this systematic method for reducing the phenotypic space is effective for organisms without highly developed transcriptional regulatory networks and may lead to further insights in better studied organisms. We have not incorporated the thermodynamics of transmembrane ion transport (Henry et al., 2006, 2007) into our approach; although, we anticipate these calculations may be useful in transitioning between specific proton flux states. This requires an accurate representation of the transmembrane ΔpH parameter, and more recent genome-scale models of *Escherichia coli* K-12 (Reed et al., 2003) and *Saccharomyces cerevisiae* (Duarte et al., 2004) have been constructed paying particular attention to the protonation

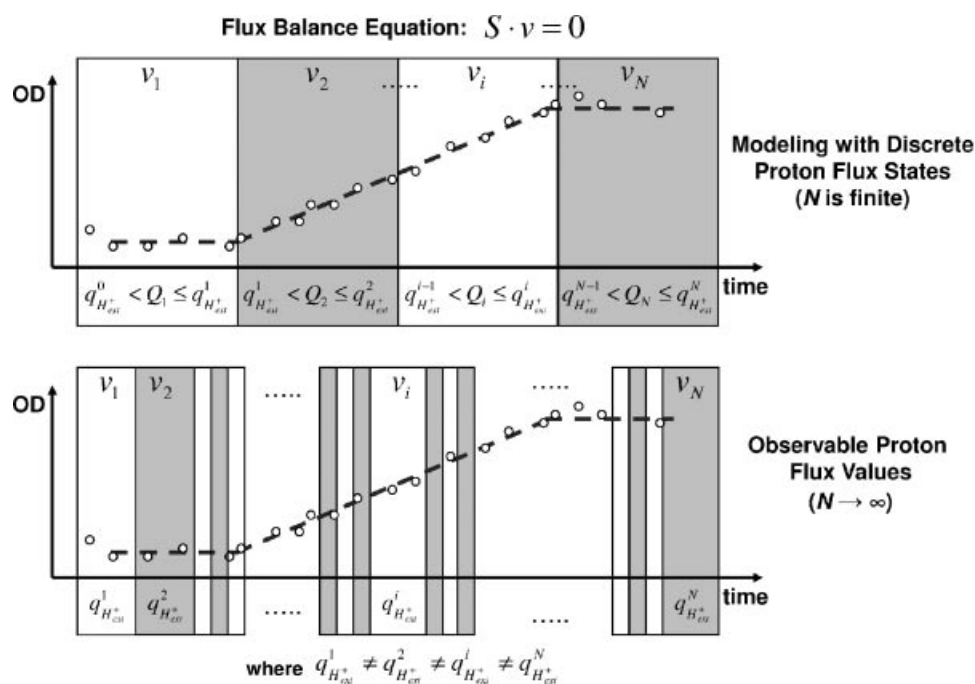


Figure 1. Illustration of discretized and continuous proton flux states using fictitious optical density (OD) data. Specific proton flux values are represented by $q_{H_{ext}}^i$ and discrete proton flux states are represented by Q_i . The number of flux solutions, v , to the flux balance equation, $S \cdot v = 0$, based on proton flux state, is represented by v^N .

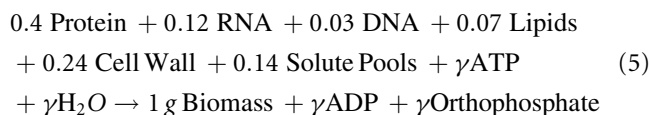
state of biological compounds at physiological pH. Effective use of the specific proton flux parameter to simulate vegetative growth and metabolism requires separating the function into discrete specific proton flux states to limit the number of flux vector solutions. Prediction of the extracellular pH was used as the benchmark for determining whether flux solutions conditioned to fit growth and metabolite data captured the cell-to-environment interactions that determine extracellular medium pH.

Batch Culture Simulations

Batch culture growth and metabolism were simulated using intracellular and membrane transport reaction flux values (the optimized flux vector, v) obtained from the flux balance equation, $S \cdot v = 0$. The simulation was performed over a time-course of 15 h of batch growth using the 4th-order Runge-Kutta numerical method with a step-size of 0.01 h. The length of the time-course simulation was chosen to correspond with typical vegetative growth of the culture. The pH of the extracellular medium was calculated at each time step using the described pH model (Dougherty et al., 2006) and $\text{pH}_{\text{extracellular}} = -\log(\text{H}_{\text{free}}^+)$. The set of intracellular and membrane transport fluxes used at each time-point were unique to the specific proton flux state of the culture. The choice of specific proton flux states and their optimized temporal bounds for *C. acetobutylicum* batch growth are discussed in the following section.

Optimization of the Biomass Constituting Equation

As also presented in Part 1 of this work (Senger and Papoutsakis, 2008), the biomass constituting equation used in the genome-scale model of *C. acetobutylicum* was adapted from one previously published for *Staphylococcus aureus* N315 (Heinemann et al., 2005). It was anticipated that the biomass composition of the relatively under-studied *C. acetobutylicum* differs (possibly significantly) from that of *S. aureus* and would be apparent when comparing calculations of the specific growth rate of *C. acetobutylicum* by the genome-scale model to experimentally measured values. In addition, it is known that *C. acetobutylicum* morphology and physiology change over the course of batch fermentation due to sporulation (Paredes et al., 2005) and by developing acid- and solvent-tolerant phenotypes (Alsaker and Papoutsakis, 2005; Alsaker et al., 2004; Borden and Papoutsakis, 2007; Harris et al., 2002; Tomas et al., 2004). We investigated the ATP requirement, γ , designated to cell maintenance in the biomass constituting equation, shown in Equation (5), for each specific proton flux state investigated.



For each specific proton flux state, the membrane transport fluxes of (i) glucose, (ii) butyric acid, (iii) acetic

acid, (iv) lactic acid, (v) acetone, (vi) butanol, and (vii) ethanol were tightly constrained to experimentally observed values (Monot et al., 1982). The ATP maintenance requirement, γ , of the biomass constituting equation was varied until the calculated specific growth rate matched experimental observations. Optimum values of γ were obtained for each specific proton flux state examined.

Optimization of Specific Proton Flux States

We argue that discrete specific proton flux states may be used to reduce the phenotypic solution space. To study the effectiveness of using specific proton flux states to model vegetative growth, a genetic algorithm was implemented to optimize the temporal bounds of different defined specific proton flux states. The different chosen discretizations (called *Sets*) of the specific proton flux state are given in Table II. The precise bounds of the specific proton flux states of Table II were chosen based on the availability of raw data points and the desire to have different specific proton flux states exist over similar time intervals when characterizing batch growth. Solutions to the flux balance equation (optimized flux vector, ν) were obtained for each specific proton flux state. This was done using the following procedure: (i) the phenotypic solution space was generated by optimizing the flux vector, ν , given the objective function of maximizing the production of reduced ferredoxins, (ii) the resulting phenotypic space was probed by varying constraints around the membrane transport fluxes of Supplementary Appendix 2, (iii) for each flux vector sampled from the phenotypic solution space, the specific proton flux was calculated using Equation (3) and the membrane transport reactions of Table I, (iv) if the specific proton flux fell within the range of a designated specific proton flux state, the flux vector was retained, and (v) for each discrete specific proton flux state, 10^2 flux vectors were identified and averaged to generate a representative flux vector. For each of the four sets of discretized specific proton flux states (Table II), a real-coded genetic algorithm was used to optimize the time point at which each flux state was implemented in vegetative growth simulations. The objective function of the optimization algorithm was the minimization of the mean square error (MSE) between

model predictions and the following minimal medium experimental observations (Monot et al., 1982): (i) biomass concentration, (ii) glucose, (iii) butyric acid, (iv) acetic acid, and (v) butanol. We refer to this procedure of optimizing temporal bounds of specific proton flux states to predict biomass and metabolite concentrations as model *training*. Thus, the *training data set* consists of 25 experimental data points (Monot et al., 1982). The trained model was then used to calculate the medium pH. These predictions were compared to the five reported experimental data points composing the extracellular pH profile (Monot et al., 1982). We refer to this procedure as *testing* the trained model. Details of the genetic algorithm parameters applied here have been published (Senger et al., 2006).

Identification of Numerically Determined Sub-Systems

The stoichiometric matrix of the metabolic network reconstruction of *C. acetobutylicum* is underdetermined (422 metabolites involved in 552 reactions) (Senger and Papoutsakis, 2008). In turn, the solution of the flux balance equation ($S \cdot \nu = 0$) is a multi-dimensional polytope due to the presence of singularities in the stoichiometric matrix. Here, we present a novel method to examine the impact of these singularities on metabolic capacity and the intracellular flux distribution. These results are then used to further constrain the stoichiometric matrix. We do this by first extracting, from the stoichiometric matrix, the reactions resulting in one of the multiple singularities of the matrix. Currently, we are focusing on singularities in reactions leading to macromolecular biosynthesis and the production of biomass. The reactions leading to biomass synthesis are extracted from the point of the singularity. These compounds and reactions are used to create a new stoichiometric matrix, S' , and flux balance equation ($S' \cdot \nu' = 0$). Additional relationships (e.g., thermodynamic, regulatory, etc.) may be applied to the reactions resulting in the singularity to study their impact on distribution of intracellular fluxes and metabolic capacity. Since the sub-network, extracted from the stoichiometric matrix, contains one singularity and relations constraining it specifically, we define this system as a numerically determined sub-system of the metabolic network reconstruction.

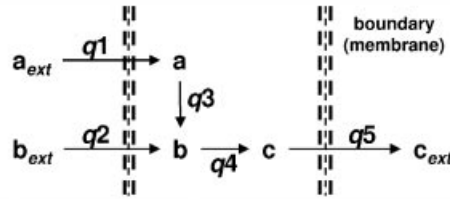
Thermodynamics (Henry et al., 2007; Kummel et al., 2006), observed metabolite selection (Desai et al., 1999b), or other investigated regulatory constraints (Choi et al., 2007; Covert et al., 2003; Price et al., 2003; Shlomi et al., 2007) may provide relationships necessary to constrain fluxes about a singularity. We have incorporated a method for including such information within the stoichiometric matrix to generate systems with a unique null space basis set. Our method is similar to the recent development of *artificial metabolites* (Choi et al., 2007). An elementary example of this concept is presented as Figure 2. This example consists of extracellular metabolites a_{ext} and b_{ext} crossing a system boundary (cell membrane) through separate transport

Table II. Investigated sets of specific proton flux states^a.

	Set 1	Set 2	Set 3	Set 4
State 1	-200 to 5	-200 to -100	-200 to -55	-200 to -100
State 2		-100 to -50	-55 to -35	-100 to -60
State 3		-50 to -20	-35 to -25	-60 to -40
State 4		-20 to 5	-25 to -15	-40 to -30
State 5			-15 to -5	-30 to -20
State 6			-5 to 5	-20 to -10
State 7				-10 to 0
State 8				0 to 5

^aSpecific proton flux states have units of $\text{mmol H}^+ \text{h}^{-1} \text{g biomass}^{-1}$.

I. Reaction scheme with singularity



1. Three boundary transport reactions with fluxes, q_1 , q_2 , q_5 .
2. Two intracellular reactions with fluxes q_3 and q_4 . Metabolites a , b and c (internal and external).

II. Corresponding Stoichiometric Matrix

$$S' = \begin{matrix} & r1 & r2 & r3 & r4 & r5 \\ \begin{matrix} a \\ b \\ c \end{matrix} & \begin{bmatrix} 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix} & \end{matrix}$$

1. Rank(S') = 3
2. Number of reactions = 5
3. Null Space Dimensions = 2
4. \therefore No unique solution
5. Singularities to be resolved = 1

III. Identification of a Flux Relationship

$$\frac{q_1}{q_2} = 0.6$$

$$q_1 - 0.6(q_2) = 0$$

1. Simple hypothetical case shown here: Ratio of fluxes is known
2. Relationships are dependent on thermodynamic parameters.

IV. Incorporation into the Stoichiometric Matrix

$$S' = \begin{matrix} & r1 & r2 & r3 & r4 & r5 \\ \begin{matrix} a \\ b \\ c \\ ratio \end{matrix} & \begin{bmatrix} 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 1 & -0.6 & 0 & 0 & 0 \end{bmatrix} & \end{matrix}$$

1. Rank(S') = 4
2. Number of reactions = 5
3. Null Space Dimensions = 1
4. \therefore Unique basis set solution exists

V. Calculation of the Null Space Basis Set (n) ($S' \cdot v' = 0$)

$$n = \begin{bmatrix} 0.23 \\ 0.38 \\ 0.23 \\ 0.61 \\ 0.61 \end{bmatrix} \quad v' = \frac{n}{n_i} \cdot \text{scalar}_i$$

1. Scalar derived from enzyme kinetics for reaction i .
2. For genome-scale models, this can be an assumed growth rate or substrate uptake rate.
3. Kinetics needed for one reaction to numerically-define the model.

Figure 2. Simple example of numerically defining a sub-system by resolving a singularity with a flux ratio relationship and applying kinetic parameters.

reactions with fluxes q_1 and q_2 . Intracellular metabolite a is converted to b with reaction flux q_3 , and b is converted to c (intracellular) with flux q_4 . Finally, c is transported across the system boundary with reaction flux q_5 . The steady-state assumption reveals the following relationship between transport fluxes: $q_5 = q_1 + q_2$, and the relationship between q_4 and q_5 is apparent ($q_4 = q_5$). However, the flux relationships between q_1 , q_2 and q_3 cannot be discerned without information regarding the transport of a_{ext} and b_{ext} into the system. This is also apparent by constructing the

stoichiometric matrix of the system, as shown in Figure 2. The resulting matrix has a rank of three for five reactions (columns), making the null space two-dimensional. Thus, for the system to be numerically determined, the null space must be one-dimensional, meaning that resolution of one singularity is required in this system. To resolve the singularity in this example, boundary (membrane) metabolite selectivity was assumed. An arbitrary reaction flux ratio relationship was assumed between q_1 and q_2 (assuming a flux relationship between q_1 or q_2 and q_5 would also

resolve the singularity). This relationship may then be built-in to the stoichiometric matrix as an additional row as shown in Figure 2. The resulting stoichiometric matrix has a rank of four and contains five columns (reactions), so a unique null space basis set solution exists. To the unique basis set, an assumed transport flux (or specific growth rate in genome-scale models) can be applied to numerically define the system. This process is further illustrated in Figure 2.

Results and Discussion

Contributors to the Specific Proton Flux State and the Environmental Response

The fluxes of membrane transport reactions used for calculating the specific proton flux in minimal medium simulations are shown in Table I. Due to the absence of specific transporter enzymes in *C. acetobutylicum* (Nolling et al., 2001), weak organic acid secretion occurs by the diffusion of the protonated weak acid species (e.g., butyric acid) across the cell membrane (Desai et al., 1999a,b). The protonated weak acid dissociates into ionic species (e.g., butyrate) in the extracellular medium as dictated by the pK_a and medium pH values. Weak acid membrane transport with full dissociation is shown in Table I, but the degree of dissociation was fully determined by the Henderson-Hasselbalch equation and the pH model. This approach allowed for calculating the contribution of weak organic acid secretion to the specific proton flux. It also enabled the incorporation of proton excretion through ion channels, such as the F-type ATPase, as well. Although multiple mechanisms of ion transport through the cell membrane are known to exist (Dills et al., 1980; Holland and Blight, 1999; Konings et al., 1995), a simplified reaction model of transmembrane ion exchange with the extracellular environment was included in the metabolic network and is listed in Table I. TC numbers and membrane transport reaction stoichiometry were obtained from the Transport Classification Database (Saier et al., 2006). In generating and sampling the phenotypic solution space, the fluxes of membrane transport equations of weak acids were constrained to the limits of observable values (Monot et al., 1982). However, the F_0F_1 proton-translocating ATP synthase (F-type ATPase) encoded by the *atp* operon (CAC2864–CAC2872) was left unconstrained. The use of the F-type ATPase with the P-type ATPase (TC 3.A.3), K^+ symporter (TC 2.A.35), and Na^+ antiporter (TC 2.A.37) creates an artificial cycle in the metabolic network. For the following simplified example, incorporation of K^+ and Na^+ into biomass as part of the solutes pool is considered to be zero. Considering the ATPases and cation transporters of Table I, if the flux value of the F-type ATPase is -1 (exporting protons at the expense of ATP) and the flux value of the P-type ATPase is 5, this yields a net proton efflux value of 28 (3 from the F-type ATPase, 15 from Na^+ antiport, and 10

from K^+ symport) and an ATP hydrolysis flux value of 6. If the flux value of the F-type ATPase is -6 and the flux value of the P-type ATPase is 0, the proton efflux is 18 with an ATP hydrolysis flux of -6. We realize the relationship between the ATPase, K^+ symporter, and Na^+ antiporter fluxes is governed by membrane transport kinetics under the influence of the proton motive force. These relationships remain unresolved in this research, so we constrained the flux of the P-type ATPase to a value of 1 for all simulations. Thus the “physiological observation” of flux through the F-type ATPase, P-type ATPase and cation transporters will exist as some multiple of the calculations presented here, with the net results of proton transport per ATP synthesis/hydrolysis being identical.

Optimization of Specific Proton Flux States

The continuous function of specific proton flux was discretized into multiple specific proton flux states, as described previously. These states are given in Table II. Representative flux vector solutions were obtained for each specific proton flux state through random sampling of the phenotypic solution space and given constraints of measured extracellular metabolites around observed values (Monot et al., 1982). Simulation results of fitted extracellular metabolite and biomass concentrations (the training data set) by real-coded genetic algorithm optimization of the specific proton flux state temporal bounds are shown in Figure 3. Results are shown for Set 1 and Set 3 of optimized specific proton flux states given in Table II. The prediction of extracellular pH (the testing data set) by these flux distributions are shown in Figure 4. Six temporal specific proton flux states (Set 3 of Table II) were required to accurately predict the experimentally observed (Monot et al., 1982) extracellular pH profile during vegetative growth. Specific fluxes for metabolites exchanged with the extracellular medium and specific growth rates are given for all six specific proton flux states in Table III. As shown in Figures 3 and 4, the model composed of six specific proton flux states is compared with a model considering only a single specific proton flux state (Set 1 of Table II). Although a single specific proton flux state model can fit vegetative growth metabolite and growth data with reasonable accuracy, as shown in Figure 3, this model fails in prediction of extracellular medium pH (see Fig. 4). Likewise, a model with four discrete specific proton flux states (Set 2 of Table II) was effectively fit to metabolite and growth data but failed to predict the extracellular medium pH. Further, the model with eight discrete specific proton flux states (Set 4 or Table II) fit experimental metabolite and growth data and accurately predicted extracellular medium pH with excellent accuracy. In this case, the model containing six specific proton flux states is preferable given the time requirement to sample the phenotypic solution space for each specific proton flux state. Results for Set 2 and Set 4 of Table II are not displayed in Figure 3 or Figure 4.

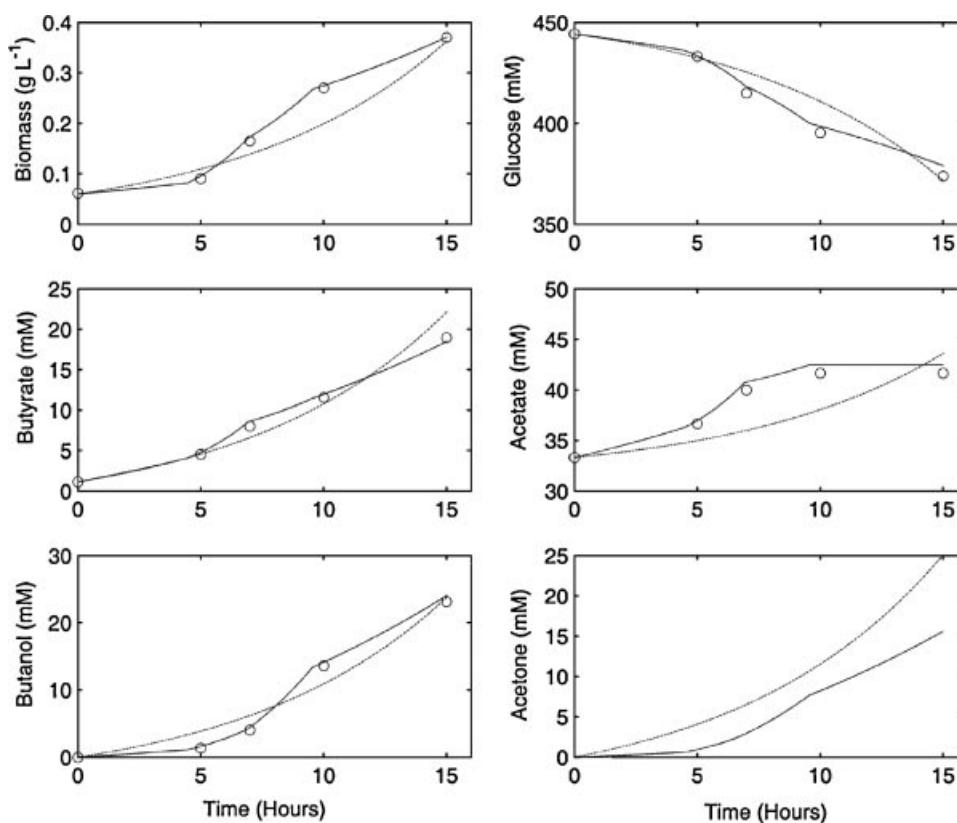


Figure 3. Raw data (circles) and optimized genome-scale model predictions (lines) for biomass production, glucose consumption, acids and solvents production during exponential growth of *C. acetobutylicum* on minimal media (Monot et al., 1982). The following model predictions are shown: (i) six discrete proton flux states (Set 3 of Table II) (solid lines) and (ii) single proton flux state model (Set 1 of Table II) (dashed lines). Note: experimental observations with minimal media were not reported for acetone; only model predictions are shown for this case.

To study the response of the pH model, in a separate set of simulations, the cation membrane transport equations (TC 2.A.37, 2.A.38, 3.A.2, 3.A.3, see Tables I and III) were inactivated for the model with six specific proton flux states (Set 2 of Table II). This modified model was used to predict extracellular medium pH, assuming the specific membrane transport fluxes and specific growth rates listed in Table III. Secretion of butyric and acetic acids allows the liberation of at most 1 mol of protons per mole of weak organic acids for the extracellular medium pH range of 4–7. Here, we define the *apparent proton flux stoichiometry* per weak acids efflux as the specific proton flux per flux of butyric and acetic acids (e.g., $q_{H^+}/(q_{\text{butyric acid}} + q_{\text{acetic acid}})$). An apparent proton flux stoichiometry of 1 reveals proton exchange with the environment is primarily through weak acids secretion by the cell. An apparent proton flux stoichiometry greater than 1 reveals significant secretion of protons through cation channels (often at the expense of ATP). Finally, an apparent proton flux stoichiometry less than 1, implies a significant influx of protons into the cell. Simulations predicting medium pH with different values of the apparent proton flux stoichiometry are shown in Figure 4. In particular, we

assumed that: (i) the proton flux through cation transport equations was zero (apparent proton flux stoichiometry = 1), (ii) the proton flux through cation transport equations was equal to the cation transport associated with weak acids secretion (apparent proton flux stoichiometry = 2), and (iii) the proton flux through cation transport equations was 2, 3, and 4 times greater than that associated with weak acids secretion (apparent proton flux stoichiometries = 3, 4, and 5, respectively). As shown in Figure 4, no single apparent proton flux stoichiometry accurately predicted the observed pH profile, and an apparent proton flux stoichiometry of 4 approximated the extracellular medium pH well for the first specific proton flux state (-100 to -55 mmol H⁺ h⁻¹ g biomass⁻¹). However, for all cases with an apparent proton flux stoichiometry greater than 1 over the duration of the batch fermentation, the extracellular proton concentration eventually exceeded the buffer capacity of the medium, and the medium pH decreased dramatically.

While it is recognized that ion transport in bacterial cells occurs through more intricate mechanisms (Das et al., 1997; Dills et al., 1980; Holland and Blight, 1999; Jones and

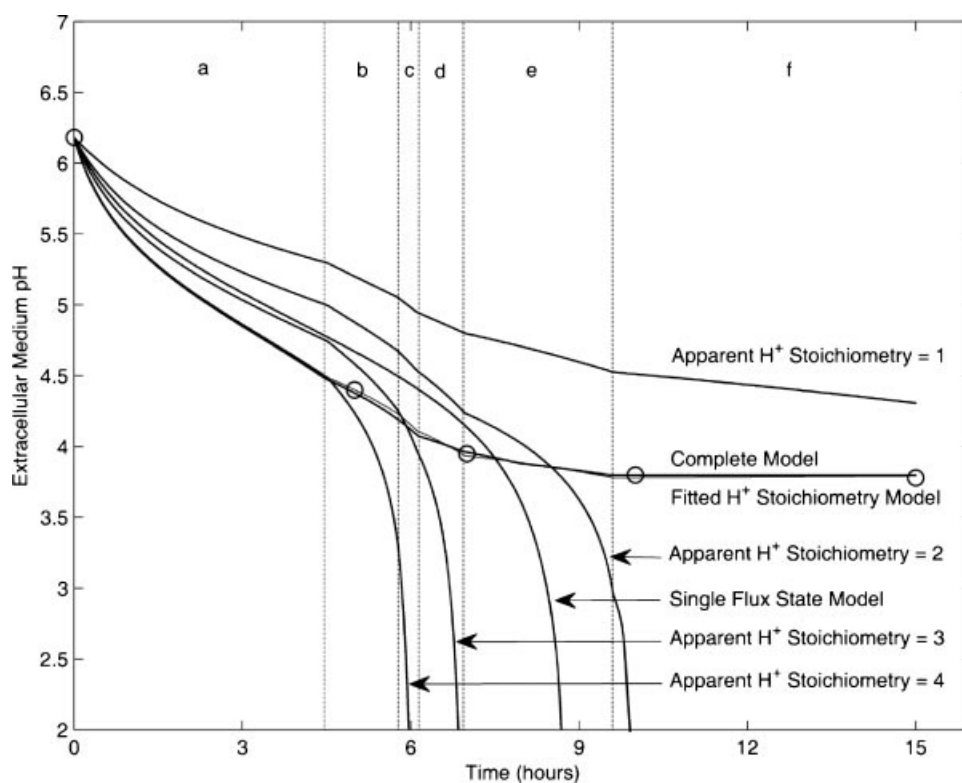


Figure 4. Model-derived values (lines) and raw data points (circles) of extracellular media pH for batch growth of *C. acetobutylicum* in minimal media (Monot et al., 1982). Proton flux states are labeled by letters: (a) -100 to -55 $\text{mmol H}^+ \text{h}^{-1} \text{g biomass}^{-1}$, (b) -55 to -35 $\text{mmol H}^+ \text{h}^{-1} \text{g biomass}^{-1}$, (c) -35 to -25 $\text{mmol H}^+ \text{h}^{-1} \text{g biomass}^{-1}$, (d) -25 to -15 $\text{mmol H}^+ \text{h}^{-1} \text{g biomass}^{-1}$, (e) -15 to -5 $\text{mmol H}^+ \text{h}^{-1} \text{g biomass}^{-1}$, and (f) -5 to 5 $\text{mmol H}^+ \text{h}^{-1} \text{g biomass}^{-1}$. The *Complete Model* is composed of six discrete proton flux states with specific fluxes and growth rates shown in Table III and growth and metabolite predictions shown in this figure. The *Single Flux State Model* consists of a single proton flux state with growth and metabolite predictions shown in this figure. The *Apparent H⁺ Stoichiometry* curves correspond to specific fluxes in Table III with proton flux from cation transport reactions ignored. Stoichiometric coefficient for protons associated with acetate and butyrate efflux was adjusted from 1 to 4 and is listed for each case. The *Fitted H⁺ Stoichiometry Model* contains adjusted stoichiometric coefficients for proton efflux with weak acids to fit the observed extracellular medium pH profile.

Woods, 1986; Konings et al., 1995; Riebeling and Jungermann, 1976) than the simplified model proposed here, the reaction model used in this study is a first attempt to study the cellular requirement to synthesize/hydrolyze ATP and import/export protons as required by the metabolic network. Fluxes through these reactions were also used to generate a better understanding of the stoichiometry of free-protons transferred to the extracellular medium in addition to butyric and acetic acids efflux for the six optimized specific proton flux states. For the case with inactivated specific proton flux through cation channels (apparent proton flux stoichiometry = 1), the predicted medium pH was greater than the observed pH profile. This case illustrates that a proton pump (proton efflux) must be operated by the cell through cation channels and ATPases, at the expense of hydrolyzing ATP, to obtain the observed phenotype of the culture. On the other hand, for apparent proton flux stoichiometries yielding an extracellular medium pH prediction below the observable values, the net flux of protons across the cell membrane through the F-type ATPase must be positive (inward), driving the synthesis of ATP. Given the optimized flux values of the F-type

ATPase in Table III, the dynamics of proton flux through cation transport mechanisms other than efflux with weak acids over the course of batch fermentation is apparent. Both of the cases described above (influx and efflux of protons through cation channels) were observed over the course of batch fermentation with proton efflux occurring early in the fermentation and proton influx apparent at the end of exponential growth.

The specific proton fluxes associated with the membrane transport equations of Table III were calculated (positive values are influx, negative is efflux) and are given in Table IV. These values were used to obtain an estimate of the upper-bound of the stoichiometry of protons associated with weak acid efflux for all six specific proton flux states. As shown in Tables I and III, the stoichiometry of protons associated with butyric and acetic acids efflux was 1 for all specific proton flux states. To calculate the apparent proton flux stoichiometry, the specific proton flux from weak acid efflux reactions was added to the specific proton flux (efflux or influx) of the cation transport reactions, as shown in Table IV. This value was then divided by the combined butyric and acetic acids efflux to calculate the moles of

Table III. Optimized specific flux values ($\text{mmol h}^{-1} \text{g biomass}^{-1}$) for selected membrane transport reactions for the six optimized proton flux states.

Classification ^a	Membrane transport reaction ^{b,c,d,e}	Specific proton flux state ($\text{mmol H}^+ \text{h}^{-1} \text{g biomass}^{-1}$)					
		-100 to -55	-55 to -35	-35 to -25	-25 to -15	-15 to -5	-5 to 5
Biomass growth equation	0.4 protein + 0.12 RNA + 0.03 DNA + 0.07 lipid + 0.24 cell wall + 0.14 solute pools + γ ATP	0.070 ($\gamma=680$)	0.30 ($\gamma=410$)	0.30 ($\gamma=400$)	0.30 ($\gamma=430$)	0.17 ($\gamma=700$)	0.060 ($\gamma=1050$)
Glucose uptake	+ γ H ₂ O → 1 g biomass + γ ADP + γ orthophosphate	29	80	89	90	41	13
Weak acid production	Phosphoenolpyruvate + D-glucose (ext.) → pyruvate + D-glucose 6-phosphate						
	Butyric acid ↔ butyrate (ext.) + H ⁺ (ext.)	9.5	13	14	13	5.3	2.4
Solvent Production	Acetic acid ↔ acetate (ext.) + H ⁺ (ext.)	9.5	14	14	14	3.5	0
	Lactic acid ↔ lactate (ext.) + H ⁺ (ext.)	3.5	2.3	0.4	0	0	0
Freely-exchanged small molecules	Butanol (ext.) ↔ 1-butanol	-4.5	-9.5	-9.6	-9.4	-14	-5.4
	Acetone (ext.) ↔ acetone	-2.5	-7.5	-7.5	-7.8	-8.8	-2.0
Anion transport	Ethanol (ext.) ↔ ethanol	-0.41	-0.40	-0.41	-0.39	-0.40	-0.40
	CO ₂ (ext.) ↔ CO ₂	-57 (±0.6)	-190 (±3)	-220 (±10)	-260 (±5)	-180 (±10)	-66 (±2)
F ₀ F ₁ ATPase	H ₂ O (ext.) ↔ H ₂ O	12 (±0.7)	7.5 (±1)	5.2 (±3)	7.0 (±0.4)	3.2 (±0.3)	6.6 (±2)
	NH ₃ (ext.) ↔ NH ₃	0.39 (±0.08)	0.26 (±0.3)	0.48 (±0.2)	0.40 (±0.1)	0.21 (±0.1)	0.03 (±0.03)
Cation transport	H ₂ (ext.) ↔ H ₂	-72 (±2)	-300 (±3)	-370 (±3)	-370 (±10)	-140 (±9)	-57 (±4)
	Sulfate (ext.) + ATP + H ₂ O ↔ sulfate + ADP + orthophosphate	0.11 (±0.01)	0.026 (±0.05)	0.29 (±0.03)	0.31 (±0.01)	0.29 (±0.9)	0.078 (±0.05)
Cation transport	Nitrate (ext.) + ATP + H ₂ O ↔ nitrate + ADP + orthophosphate	0 (±0)	2.3 (±0.1)	2.3 (±0.2)	2.4 (±0)	1.9 (±3.7)	0.46 (±0.1)
	Orthophosphate (ext.) + ATP + H ₂ O ↔ 2 orthophosphate	0.081 (±0.01)	0.23 (±0.2)	0.36 (±0.03)	0.38 (±0)	0.9 (±2)	0 (±0)
Cation transport	ADP + orthophosphate + 3 H ⁺ (ext.) ↔ ATP + H ₂ O + 3 H ⁺	-11 (±1)	-6.1 (±0.4)	-0.27 (±0.2)	4.2 (±0.8)	2.0 (±0.5)	2.3 (±0.7)
	K ⁺ (ext.) + H ⁺ (ext.) ↔ K ⁺ + H ⁺	-2	-2	-2	-2	-2	-2
Cation transport	Na ⁺ (ext.) + H ⁺ ↔ Na ⁺ + H ⁺ (ext.)	3	3	3	3	3	3
	3 Na ⁺ + 2 K ⁺ (ext.) + ATP + H ₂ O ↔ 3 Na ⁺ (ext.) + 2 K ⁺ + ADP + Orthophosphate	1	1	1	1	1	1

^aTight constraints were used for the biomass growth equation, glucose uptake, weak acid production, and solvent production fluxes. These are further detailed in Supplementary Appendix 2.

^bPositive flux values correspond to the forward (left to right) direction of the transport reaction as written. Negative values denote the reverse direction of the reaction as written.

^cAll compounds labeled “extracellular” (“ext.”) are located outside the cell (in the medium).

^dError values are given in parentheses and correspond to 1 standard deviation.

^eThe flux values of the P-type ATPase were constrained to a value of 1 in all cases.

Table IV. Net flux of metabolites in cation transport equations of Table III with calculated and fitted apparent proton flux stoichiometry with weak acids efflux.

	Specific proton flux state (mmol H ⁺ h ⁻¹ g biomass ⁻¹)					
	<-55	-55 to -35	-35 to -25	-25 to -15	-15 to -5	-5 to 5
H ⁺ ^a	-38	-23	-5.8	7.6	1.0	1.9
Na ⁺ ^a	0	0	0	0	0	0
K ⁺ ^a	0	0	0	0	0	0
ATP ^b	-11	-7.1	-1.3	3.2	1.0	1.3
Calculated specific proton flux ^c	-61	-52	-34	-20	-7.8	-0.50
Calculated apparent proton flux stoichiometry with weak acid efflux ^d	3.2	1.9	1.2	0.72	0.89	0.21
Curve-fitted apparent proton flux stoichiometry with weak acid efflux ^e	4.0	1.5	1.0	1.0	0.40	0.20

^aNet flux calculated from *cation transport* flux values of Table III.

^bNet flux of ATP synthesized (positive)/hydrolyzed (negative) from F₀F₁ ATPase and Na⁺/K⁺ ATPase in Table III.

^cCalculated from the fluxes of reactions listed in Table I. The value must be within the proton flux state ranges listed above.

^dThe calculated Specific Proton Flux value divided by the sum of acetate and butyrate fluxes of Table III.

^eOptimized value from curve-fitting the modified model to the observed pH profile.

protons secreted per mole of weak acid. This number was recalculated by curve-fitting the extracellular medium pH profile using the abbreviated flux model with inactivated cation transfer equations (see above). The apparent proton flux stoichiometry associated with butyric and acetic acids efflux was adjusted for each specific proton flux state to fit the experimental pH data (Monot et al., 1982). The calculated results and curve-fit results of apparent proton flux stoichiometry associated with weak acid efflux are in good agreement, as shown in Table IV. The apparent proton flux stoichiometry associated with weak acid efflux is near 4 (or slightly lower) at the start of the culture and is reduced to roughly 1 when weak acid effluxes are maximized. Toward the end of vegetative growth of *C. acetobutylicum*, model predictions conclude that the net flux of protons across the cell membrane is roughly zero. It has been shown (McLaughlin et al., 1985) that ATP is in high demand (low availability) during early acidogenic stages of batch fermentation. Our model results suggest this is likely because the proton efflux through cation channels occurs at the expense of hydrolyzing ATP and is quite significant early in batch fermentation. Conversely, toward the end of exponential growth, the extracellular pH is much lower than the intracellular pH (yielding a large Δ pH) (Husemann and Papoutsakis, 1988). Here, the import (or reduced export) of protons drives ATP synthesis (or reduces ATP hydrolysis) through, particularly, the F-type ATPase. This is reflected in experimental findings of excess intracellular ATP during the later stages of exponential growth (McLaughlin et al., 1985).

Regarding Optimization of the Biomass Equation and Balance of Hydrogen Atoms

The results of optimizing the ATP maintenance coefficient, γ , of Equation (5) for the six optimized specific proton flux

states discussed above are shown in Figure 5. The calculated specific growth rates are shown for six separate values of γ , ranging between 100 and 1,000 for each specific proton flux state. Experimentally observed specific growth rates (Monot et al., 1982) are indicated for each specific proton flux state by a dashed horizontal line marker in Figure 5. Optimized values of γ are printed above the data for each specific proton flux state in Figure 5. Here, we show that the ATP maintenance coefficient, γ , is correlated with the specific proton flux state. Further investigation of optimized values revealed that relatively low values of γ occur during the acidogenic growth phase coupled with maximum growth rates. Larger values of γ were calculated for the culture lag phase and for specific proton flux states corresponding to solventogenesis (e.g., -5 to 5 mmol H⁺ h⁻¹ g biomass⁻¹). We argue that the high ATP maintenance requirements for the latter solventogenic specific proton flux states are coupled with the onset of sporulation. Sporulation has been described as an *energy-intensive* process (Dworkin and Losick, 2005; Parker et al., 1996), and the coupling of sporulation to solventogenesis in *C. acetobutylicum* has been well-characterized in recent literature (Alsaker and Papoutsakis, 2005; Harris et al., 2002; Paredes et al., 2005; Tomas et al., 2003). However, we also anticipate metabolic programs responsible for solvent-tolerant phenotypes are responsible for the increased ATP requirement of biomass growth during solventogenesis (Borden and Papoutsakis, 2007; Lepage et al., 1987; Tomas et al., 2004). Thus, we recognize that other stoichiometric coefficients (e.g., lipids) of the biomass constituting equation are also functions of the specific proton flux state. Here, we have approximated a *dynamic* biomass constituting equation by investigating the ATP maintenance coefficient only. A maintenance energy coefficient of 14 (mmol ATP h⁻¹ g biomass⁻¹) was calculated for *C. acetobutylicum* cultures grown in a complex medium containing yeast extract (Meyer and Papoutsakis, 1989). Using the values of γ calculated in this study,

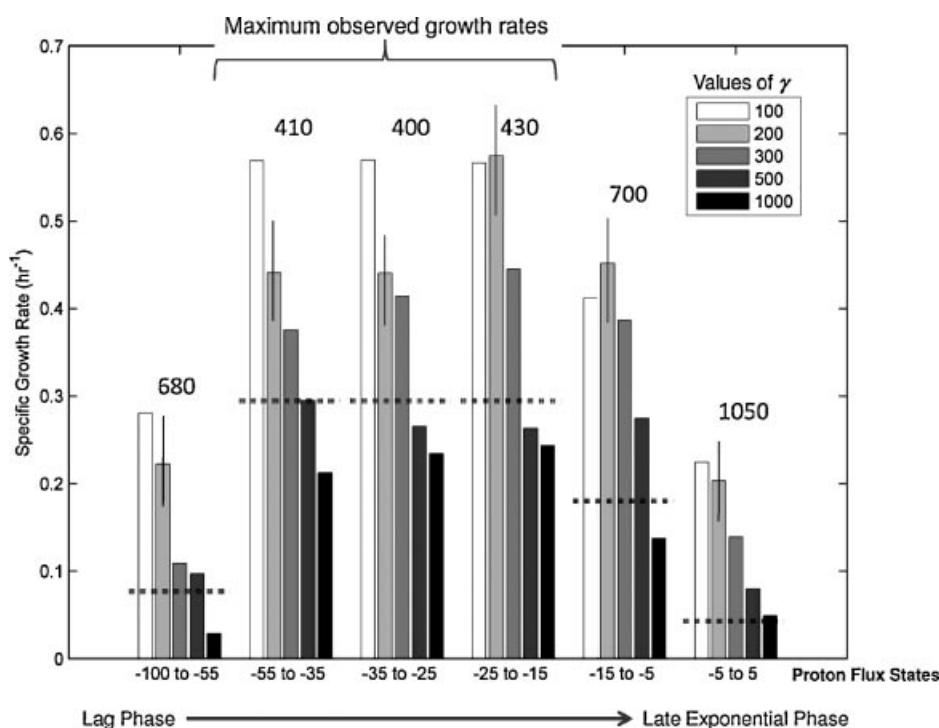


Figure 5. Calculated specific growth rate for specified proton flux states given multiple values of the stoichiometry of ATP in the biomass constituting equation (Eq. 5). The dashed horizontal lines correspond to the experimentally observed specific growth rate (Monot et al., 1982). Numerical values printed above the data correspond to the optimized value of the stoichiometric coefficient of ATP (also referred to as γ in Eq. 5) for each proton flux state. Error bars are given for the case in which the stoichiometric coefficient of ATP is equal to 200 and represent one standard deviation.

maintenance energy coefficients between 40 and 150 ($\text{mmol ATP h}^{-1} \text{g biomass}^{-1}$) result. It is expected that these maintenance energy coefficients should be larger than previously published values since this study involves a minimal medium. However, we acknowledge that these large values of γ may also indicate inadequacies of the biomass constituting equations, which were adapted for clostridia from the genome-scale model of *S. aureus* N315 (Heinemann et al., 2005). We believe this may also help explain the relatively large flux values of H_2 and CO_2 formation.

To verify the use of the biochemical reaction formalism in construction of the genome-scale model, a hydrogen atom balance was performed over the entire cell using membrane transport reactions and biomass constituting equations written using the chemical reaction formalism. The full balance for all proton flux states is given as Supplementary Appendix 4. Results showed an accumulation of hydrogen atoms of 0.63–2.13% of the total number of hydrogen atoms (not just protons) transferred across the cell membrane. Given the stochastic nature of sampling the phenotypic solution space, leading to averaged flux values with measured standard deviations, arriving at a hydrogen atom balance yielding 2% or less accumulation is sufficient to assume that hydrogen atoms are not being created or

destroyed within the metabolic network when written using the biochemical reaction formalism.

“Unconstrained” Simulations of the Genome-Scale Model

Simulations of the genome-scale model were performed with unconstrained production of acids (butyric, acetic, and lactic) and solvents (acetone, butanol, and ethanol). The selectivity coefficient of acids to solvents for each specific proton flux state is shown in Figure 6. The genome-scale model with applied constraints is given in Supplementary Appendix 1. Stochastically varied constraints, including tight constraints, for every simulation are given in Supplementary Appendix 2. Results show that for all three cases examined (i) tightly constrained specific glucose uptake and growth rates, (ii) constrained glucose uptake rate only, and (iii) unconstrained specific glucose uptake and growth rates, the specific proton flux state is capable of characterizing the start of the exponential growth phase of *C. acetobutylicum* as acidogenic (selectivity of acids to solvents production greater than 1) and the end of exponential growth as solventogenic (selectivity less than 1). The success of the “unconstrained” simulations suggest that

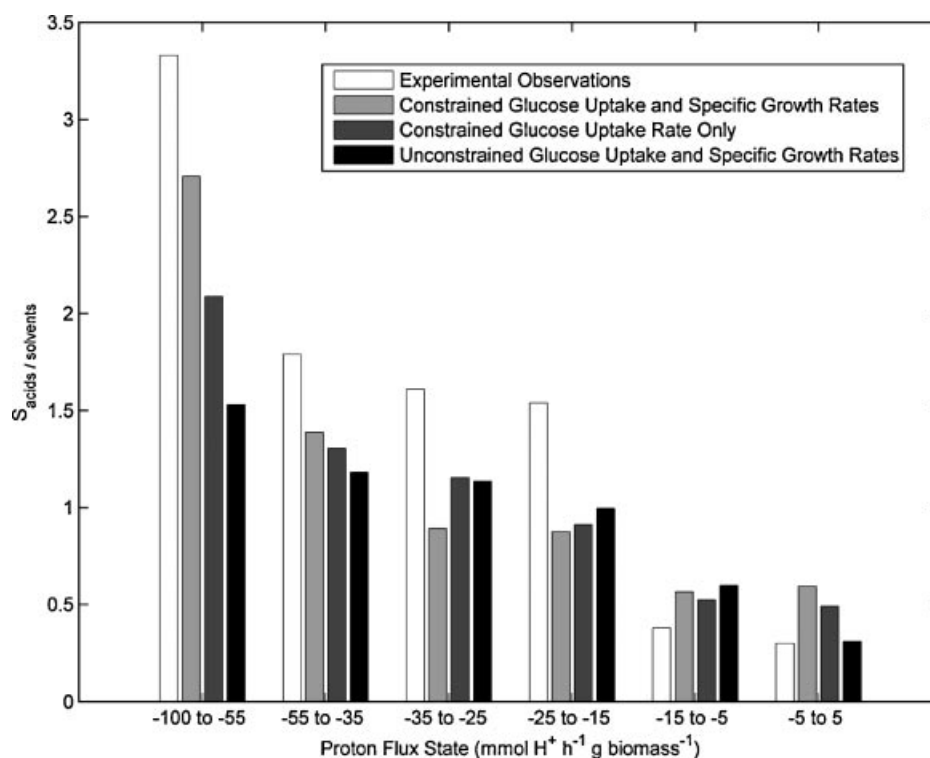


Figure 6. The selectivity of acids to solvents for: (i) experimental observations, (ii) simulation of the genome-scale model in which glucose uptake and specific growth rates were constrained to experimentally observed values (Monot et al., 1982), (iii) simulations in which only the glucose uptake rate was constrained, and (iv) model simulations with no constraints on glucose uptake or specific growth rates. In all cases, reaction fluxes determining the proton flux state (including butyric, acetic, and lactic acids efflux) and solvent (acetone, butanol, and ethanol) effluxes were left unconstrained. The selectivity of acids to solvents is defined as the sum of butyric, acetic, and lactic acids efflux divided by the sum of acetone, butanol, and ethanol effluxes.

the use of the genome-scale model with proton flux states will be useful to metabolic engineering ventures that seek to maximize solvents production. However, simulation results of Figure 6 do show some apparent differences with experimental observations. For example, the selectivity of acids to solvents at the start of the culture (specific proton flux state -100 to -55 $\text{mmol H}^+ \text{h}^{-1} \text{g biomass}^{-1}$) was observed to be higher than model predictions, and the incorporation of the tightly constrained specific glucose uptake and growth rates significantly improved model predictions. Additionally, the exponential growth phase was observed to be slightly more acidogenic than model predictions, and the latter stages of fermentation were slightly more solventogenic than calculations predicted. It has been well-established in recent literature that solventogenesis in *C. acetobutylicum* is under control of the Spo0A transcriptional regulator (Alsaker et al., 2004; Harris et al., 2002). However, these regulatory models have not yet been built into the *C. acetobutylicum* genome-scale model. Regardless, these data illustrate the effectiveness of specific proton flux states for predicting acidogenic and solventogenic phenotypes while showing where additional regulatory models could improve genome-scale model predictions.

Analysis of a Singularity Using a Numerically Determined Sub-System

In Part 1 (Senger and Papoutsakis, 2008), we described two pathways in the metabolic network for the biosynthesis of L-glutamate. These were: (i) conversion of pyruvate and L-ornithine to L-glutamate and L-alanine by the acetylornithine transaminase (ArgD, EC 2.6.1.11, CAC2388) and (ii) biosynthesis of L-glutamate through the L-arginine biosynthesis pathway in the presence of a large intracellular L-glutamate solute pool. Thus, L-alanine is a product of L-glutamate biosynthesis by ArgD. However, in *C. acetobutylicum*, L-alanine may also be biosynthesized from pyruvate through D-alanine by D-alanine transaminase (EC2.6.2.21, CAC0792) and alanine racemase (EC 5.1.1.1, CAC0492). This creates a singularity in the metabolic network. To examine fluxes around this singularity, the numerically determined sub-system, shown in Figure 7, was extracted from the genome-scale metabolic network and analyzed in detail by varying the ratio of fluxes responsible for L-glutamate biosynthesis. These reactions are shown in Figure 7 as Reaction 1 (L-glutamate biosynthesis by ArgD) and Reaction 2 (L-glutamate biosynthesis through the

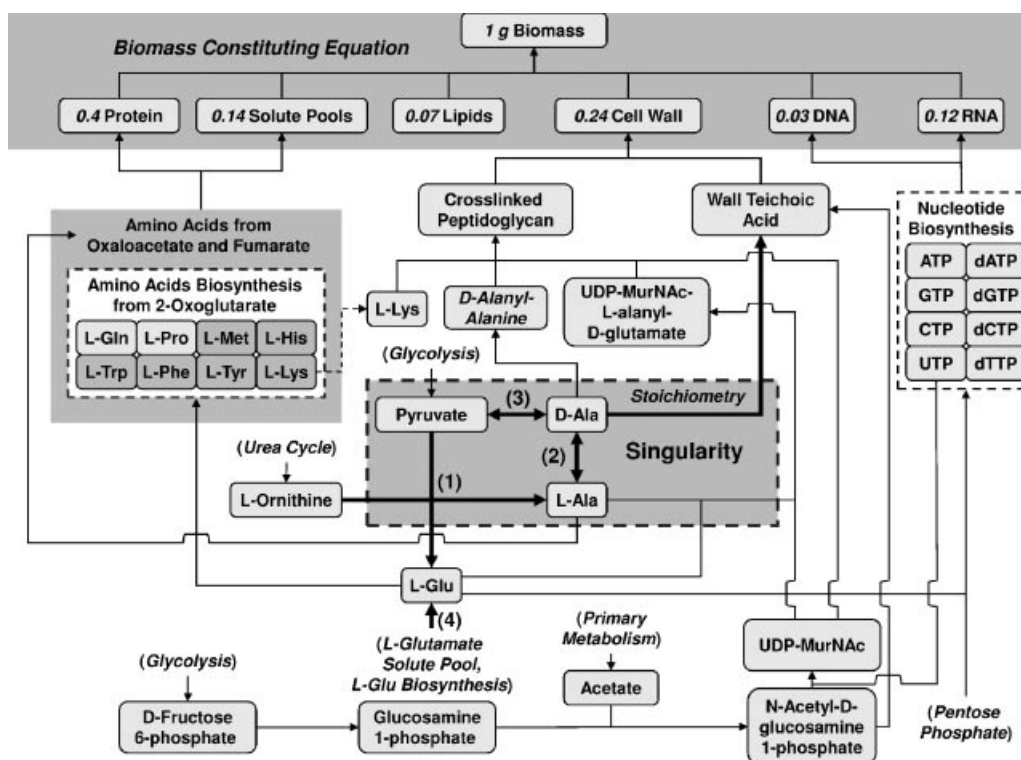


Figure 7. Sub-system of the genome-scale model to investigate flux constraint bounds around D-alanine and probe metabolic capacity based on its incorporation into D-alanylation of wall teichoic acids. The location of the singularity of the sub-system is identified. It was resolved by varying the ratio of Reaction 1 (through acetylornithine transaminase (ArgD, EC 2.6.1.11, CAC2388)) to Reaction 4 (L-glutamate biosynthesis through L-arginine biosynthesis pathway). Reaction 2 is catalyzed by the alanine racemase (EC 5.1.1.1, CAC0492) and D-alanine transaminase (EC 2.6.2.21, CAC0792) drives Reaction 3. Not present in the diagram above, but assumed to be available in excess were: (i) all L-amino acids not derived from L-glutamate, (ii) all required lipids for biomass synthesis, (iii) phosphorylated carbohydrate required by nucleotide biosynthesis, (iv) all intracellular solute pools, (v) sources of all additional molecules required by synthesis reactions, (vi) sinks for all byproducts of synthesis reactions and (vii) all energy requirements.

L-arginine biosynthesis pathway). The stoichiometric matrix for this sub-system is given in Supplementary Appendix 5. All precursors and energetic requirements are assumed readily available for this sub-system. For example, the sub-system requires the stoichiometric amount of D-fructose-6-phosphate, available from glycolysis (see Fig. 7), to support the specified growth rate.

The sub-system of Figure 7 was used to directly study the flux size and direction of alanine racemase (EC 5.1.1.1, CAC0492) (shown as Reaction 2 in Fig. 7) as well as D-alanine transaminase (EC 2.6.1.21, CAC0792) (shown as Reaction 3 in Fig. 7). The sub-system was solved by fixing a value of the specific growth rate and by assuming flux ratios for L-glutamate biosynthesis by ArgD (shown as Reaction 1 in Fig. 7) and through the L-arginine biosynthesis pathway (shown as Reaction 4 in Fig. 7). First, the reaction fluxes through alanine racemase (Reaction 2) and are shown in Figure 8a as a function of the specific flux of L-alanine through ArgD (Reaction 1) for a range of specific growth rates from 0.05 h^{-1} through 0.35 h^{-1} . The catalyzing direction of the alanine racemase enzyme was found to vary (forward or reverse) under normal growth conditions and quantitatively illustrates how reaction direction changes

with growth rate and available precursors. Every specific growth rate has a maximum calculated flux of L-alanine through ArgD (Reaction 1) and alanine racemase (Reaction 2), as shown in Figure 8a. Above these maximum flux values, products of these reactions exceed the demand required of cell growth. Thus, these maximum flux values are labeled as a *theoretical boundary* in Figure 8a, and the shaded *infeasible region* of Figure 8a represents flux values exceeding this maximum *metabolic capacity*. The theoretical boundary also represents the line at which the flux ratio of Reaction 1 to Reaction 4 goes to infinity. The boundary where the flux ratio of Reaction 1 to Reaction 4 approaches zeros is the ordinate axis of Figure 8a. Thus, the feasible solution space of L-alanine flux through ArgD and alanine racemase is *triangular-shaped* and is bounded by: (i) the maximum theoretical boundary discussed above, (ii) the ordinate axis, and (iii) the operating line of the maximum growth rate. This solution space provides ranges for constraints of these reactions in optimization of the flux vector of the flux balance equation using linear programming. Of course, through this approach, we assume adequate ATP biosynthesis and cofactor regeneration by other parts of the metabolic network (e.g., glycolysis and weak acid/solvent

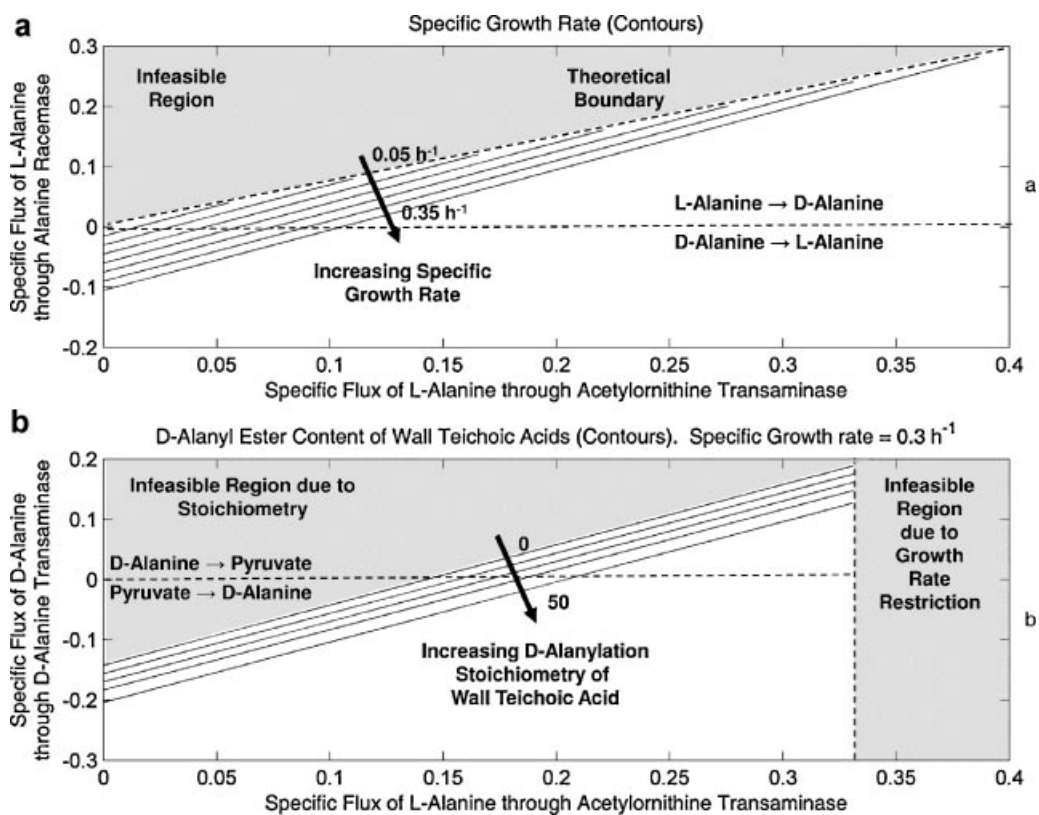


Figure 8. Results of simulations for the numerically defined sub-system shown in Figure 6. **a:** The flux ratio for L-glutamate production from ArgD to nitrogen assimilation (Reactions 1 and 4 in Fig. 6), was varied to produce ratios of the specific flux of L-alanine through alanine racemase (Reaction 2 in Fig. 6) against the flux of L-alanine through acetylornithine transaminase (Reaction 1 in Fig. 6). The simulation was performed for multiple assumed specific growth rates between 0.05 h⁻¹ and 0.35 h⁻¹. **b:** The relationship between the specific flux of D-alanine through D-alanine transaminase (Reaction 3 in Fig. 6) and the specific flux of L-alanine through acetylornithine transaminase (Reaction 1 in Fig. 6) was produced by varying the flux ratio for L-glutamate production. This is shown for an increasing number (0–50) of D-alanine residues involved in D-alanylation of wall teichoic acids. For these calculations, a specific growth rate of 0.3 h⁻¹ was assumed.

production) to support specific growth rates that make-up the bounds of Figure 8a. With substrate limitations, specific growth rates will approach zero.

With this sub-system, we were also able to probe metabolic capacity in the presence of modified teichoic acids in the cell wall, as has been observed for many pathogenic clostridia and *B. subtilis* (Neuhaus and Baddiley, 2003; Pollack and Neuhaus, 1994). Thus, intracellular fluxes of the sub-system were also found to be dependent upon the possible presence of D-alanylation of teichoic acids in *C. acetobutylicum*. Although, the presence of D-alanylation in *C. acetobutylicum* is unlikely, due to the absence of a complete *dlt* operon, its consideration provides a perturbation useful for discovering limits of intracellular flux values given altered metabolic demands. D-alanyl content of wall teichoic acids was varied from 0 to 50 residues. The flux ratio of Reactions 1 and 4 in Figure 5 were varied to produce the flux relationship between acetylornithine transaminase (ArgD) (Reaction 1) and D-alanine transaminase (Reaction 3), as shown in Figure 8b. With a specific growth rate of 0.3 h⁻¹, results show that at a fixed metabolic flux of ArgD,

increasing the D-alanyl content of wall teichoic acids from 0 to 50 increases the metabolic flux of pyruvate through D-alanine transaminase to D-alanine by 0.07 mmol h⁻¹ g biomass⁻¹. This accounts for approximately 0.08% of the specific influx of glucose observed for this growth rate (Monot et al., 1982), making the inclusion of D-alanylation relatively insignificant in the metabolic network of *C. acetobutylicum*. In summary, this example demonstrates that numerically defined sub-systems of the metabolic network can be probed to reveal the metabolic impact of including physiological processes that have not yet been identified through experimental means.

Conclusions

Using the reconstructed metabolic network developed previously (Senger and Papoutsakis, 2008), a constrained genome-scale model for *C. acetobutylicum* was developed for published minimal medium data (Monot et al., 1982). Novel techniques were developed in this research that may be used

to further reduce the potential phenotypic space of genome-scale models while generating further understanding of metabolic capacities and cell-to-environment interactions. The concept of specific proton flux states was coupled to the pH of the extracellular environment in the first such development for genome-scale models. The specific proton flux state is a type of constraint derived from the endo-exo-metabolome interface and may prove useful in supplementing regulatory models derived around the transcriptome or fluxome levels of metabolic regulation. The specific proton flux was found to become less-negative (fewer protons leaving the cell) over the course of exponential growth in minimal media and achieved a plateau value close to zero toward the end of vegetative growth of the culture. Optimized discretization of the continuous function of specific proton flux resulted in six discrete temporal states that allowed accurate prediction of the extracellular medium pH throughout the vegetative growth stage. In addition, further calculations revealed the apparent stoichiometry of protons secreted with weak acids (predominantly butyric and acetic acids) during vegetative growth. At the start of the culture, given specific proton fluxes less than $-55 \text{ mmol h}^{-1} \text{ g biomass}^{-1}$ (proton efflux), approximately 3.5 mol of protons are secreted per mole of butyric and acetic acids. Following an initial drop in extracellular medium pH, this value reduces to 1 and coincides with the maximum observed specific production of butyric and acetic acids. Further, optimization of the biomass constituting equation revealed maintenance ATP demands of the lag and near-stationary phases exceeded those of the exponential growth phase by a factor of ~ 2.5 . It is speculated this is due to changing morphology due to endospore germination in the lag phase, development of solvent tolerant phenotypes, and sporulation in the latter stages of batch growth. It is likely these calculations will lead to a dynamic biomass constituting equation to describe batch growth. Short-comings of this model are recognized in that calculated energy maintenance coefficients were significantly larger than previously published values and may be tied to inadequacies of the biomass constituting equations for describing clostridial cultures. In addition, elevated flux values of H_2 and CO_2 formation may have resulted from this issue or from the presence of secreted byproducts (e.g., acetoin) not measured in the original minimal medium study (Monot et al., 1982).

The concept of numerically defined sub-spaces of genome-scale networks for limiting the phenotypic solution space was also introduced in this study. A particular sub-system surrounding L-alanine and D-alanine biosyntheses/degradations was presented and was analyzed in detail given a singularity created by including multiple pathways of L-glutamate biosynthesis in the metabolic network. The sub-system described, quantitatively, the ranges and direction of specific fluxes surrounding L- and D-alanine biosynthesis for various physiological demands for D-alanine in the biosynthesis of peptidoglycan and D-alanylation of teichoic acids.

From the sub-system presented here, we propose the following requirements and strategy for extracting and analyzing a numerically defined sub-system: (i) the genome-scale metabolic network is complete, (ii) the singularity to be studied has been identified, (iii) no other singularities exist in the sub-system, (iv) a specific growth rate or metabolite membrane exchange flux is known, (v) the biosynthetic precursor entering the sub-system leads to cell growth and is not degraded for energy production. In addition, the following steps are given for the general identification of numerically defined sub-systems in genome-scale metabolic network reconstructions: (i) identify a singularity in the stoichiometric matrix involving one or more reactions; (ii) identify relationships (e.g., thermodynamics, etc.) to describe the relationship of fluxes that cause the singularity; (iii) reconstruct the metabolic network between the reactions of the singularity and the biomass constituting equation; (iv) allow precursors to reactions of the singularity and energetic requirements to be imported into the sub-network as required; (v) construct the sub-system stoichiometric matrix, S' , and flux vector, v' , based on the compounds and reactions present in the sub-network; (vi) add flux relationships of the singularity to the stoichiometric matrix as shown in Figure 2; (vii) obtain the unique null space basis set solution and flux vector values as also shown for the simplified system in Figure 2; (viii) vary the flux relationships about the singularity (if necessary) to obtain new unique solutions; and (ix) correlate intracellular flux distributions to those flux ratios about the singularity. Numerically defined sub-systems can also be used to probe changes in intracellular fluxes and metabolic capacity in response to (among others): (i) protein glycosylation, (ii) changes in membrane lipids composition, (iii) possible solvent stress responses, and (iv) physiological processes associated with sporulation.

Nomenclature

Acids	all acids in minimal medium: acetic, butyric, lactic, carbonic, phosphate ion, mono- and di-basic potassium phosphate
Bases	all bases in minimal medium: ammonia
counterions	all species with counterions in minimal medium: mono- and di-basic potassium phosphate
C_A	total acids concentration of the extracellular medium (mM)
C_B	total base concentration of the extracellular medium (mM)
C_C	total concentration of species with counterions in the extracellular medium (mM)
d	number of acidic dissociation sites
D	total number of acidic and basic dissociation sites
γ	stoichiometric coefficient of ATP in the biomass constituting equation
H_{ext}^+	total extracellular hydrogen ion concentration (contains those of protonated weak acids) (mM)
H_{free}^+	extracellular free proton concentration ($\text{pH} = -\log(\text{H}_{\text{free}}^+)$) (mM)
$\text{H}_{\text{butyrate}}^+$	extracellular protonated butyrate (butyric acid) concentration (mM)

H^+_{acetate}	extracellular acetic acid concentration (mM)
H^+_{lactate}	extracellular lactic acid concentration (mM)
$H^+_{\text{carbonate}}$	extracellular carbonic acid concentration (mM)
$H^+_{\text{phosphates}}$	extracellular mono- and dibasic potassium phosphates concentration (mM)
H^+_{ammonium}	extracellular ammonium concentration (mM)
h_r	stoichiometric coefficient of H^+ in membrane transport equation r
K_{aj}	acid dissociation constant of species j
K_W	water ionization constant ($=10^{-14}$)
M	number of membrane transport equations
N	number of specific proton flux states needed to model fermentation data
n	null space basis set vector
n_c	number of counterions
$q_{H^+_{\text{ext}}}$	specific proton flux ($\text{mmol h}^{-1} \text{g biomass}^{-1}$)
q_r	specific flux of reaction r ($\text{mmol h}^{-1} \text{g biomass}^{-1}$)
Q_i	Discrete specific proton flux state i with bounds $q_{H^+_{\text{ext}}}^{i-1}$ and $q_{H^+_{\text{ext}}}^i$
S	stoichiometric matrix of the genome scale model
S'	stoichiometric matrix of the sub-system
t	time (h)
v	vector of specific flux values for the genome scale model
v'	vector of specific flux values for the sub-system
X	biomass concentration (g L^{-1})
z_k	signed charge of the k th counterion

This work was supported by NSF grant BES-0418157. R.S.S. was supported by NIH NRSA post-doctoral training grant F32GM078947.

References

- Alberty RA. 1991. Equilibrium compositions of solutions of biochemical species and heats of biochemical reactions. *Proc Natl Acad Sci USA* 88(8):3268–3271.
- Alberty RA. 1993. Levels of thermodynamic treatment of biochemical reaction systems. *Biophys J* 65(3):1243–1254.
- Alsaker KV, Papoutsakis ET. 2005. Transcriptional program of early sporulation and stationary-phase events in *Clostridium acetobutylicum*. *J Bacteriol* 187(20):7103–7118.
- Alsaker KV, Spitzer TR, Papoutsakis ET. 2004. Transcriptional analysis of *spo0A* overexpression in *Clostridium acetobutylicum* and its effect on the cell's response to butanol stress. *J Bacteriol* 186(7):1959–1971.
- Borden JR, Papoutsakis ET. 2007. Dynamics of genomic-library enrichment and identification of solvent tolerance genes for *Clostridium acetobutylicum*. *Appl Environ Microbiol* 73(9):3061–3068.
- Busch W, Saier MH Jr. 2002. The transporter classification (TC) system, 2002. *Crit Rev Biochem Mol Biol* 37(5):287–337.
- Choi HS, Kim TY, Lee DY, Lee SY. 2007. Incorporating metabolic flux ratios into constraint-based flux analysis by using artificial metabolites and converging ratio determinants. *J Biotechnol* 129(4):696–705.
- Covert MW, Famili I, Palsson BO. 2003. Identifying constraints that govern cell behavior: A key to converting conceptual to computational models in biology? *Biotechnol Bioeng* 84(7):763–772.
- Das A, Ivey DM, Ljungdahl LG. 1997. Purification and reconstitution into proteoliposomes of the F1F0 ATP synthase from the obligately anaerobic gram-positive bacterium *Clostridium thermoautotrophicum*. *J Bacteriol* 179(5):1714–1720.
- Desai RP, Harris LM, Welker NE, Papoutsakis ET. 1999a. Metabolic flux analysis elucidates the importance of the acid-formation pathways in regulating solvent production by *Clostridium acetobutylicum*. *Metab Eng* 1(3):206–213.
- Desai RP, Nielsen LK, Papoutsakis ET. 1999b. Stoichiometric modeling of *Clostridium acetobutylicum* fermentations with non-linear constraints. *J Biotechnol* 71(1–3):191–205.
- Dills SS, Apperson A, Schmidt MR, Saier MH. 1980. Carbohydrate transport in bacteria. *Microbiol Rev* 44(3):385–418.
- Dougherty DP, Da Conceicao Neta ER, McFeeters RF, Lubkin SR, Breidt F Jr. 2006. Semi-mechanistic partial buffer approach to modeling pH, the buffer properties, and the distribution of ionic species in complex solutions. *J Agric Food Chem* 54(16):6021–6029.
- Duarte NC, Herrgard MJ, Palsson BO. 2004. Reconstruction and validation of *Saccharomyces cerevisiae* iND750, a fully compartmentalized genome-scale metabolic model. *Genome Res* 14(7):1298–1309.
- Dworkin J, Losick R. 2005. Developmental commitment in a bacterium. *Cell* 121(3):401–409.
- Famili I, Mahadevan R, Palsson BO. 2005. k-Cone analysis: Determining all candidate values for kinetic parameters on a network scale. *Biophys J* 88(3):1616–1625.
- Gianchandani EP, Papin JA, Price ND, Joyce AR, Palsson BO. 2006. Matrix formalism to describe functional states of transcriptional regulatory systems. *PLoS Comput Biol* 2(8):e101.
- Gros JB, Dussap CG, Catte M. 1999. Estimation of O_2 and CO_2 solubility in microbial culture media. *Biotechnol Prog* 15(5):923–927.
- Harris LM, Welker NE, Papoutsakis ET. 2002. Northern, morphological, and fermentation analysis of *spo0A* inactivation and overexpression in *Clostridium acetobutylicum* ATCC 824. *J Bacteriol* 184(13):3586–3597.
- Heinemann M, Kummel A, Ruinatscha R, Panke S. 2005. *In silico* genome-scale reconstruction and validation of the *Staphylococcus aureus* metabolic network. *Biotechnol Bioeng* 92(7):850–864.
- Henry CS, Jankowski MD, Broadbelt LJ, Hatzimanikatis V. 2006. Genome-scale thermodynamic analysis of *Escherichia coli* metabolism. *Biophys J* 90(4):1453–1461.
- Henry CS, Broadbelt LJ, Hatzimanikatis V. 2007. Thermodynamics-based metabolic flux analysis. *Biophys J* 92(5):1792–1805.
- Holland IB, Blight MA. 1999. ABC-ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans. *J Mol Biol* 293(2):381–399.
- Husemann MHW, Papoutsakis ET. 1988. Solventogenesis in *Clostridium acetobutylicum* fermentations related to carboxylic acid and proton concentrations. *Biotechnol Bioeng* 32(7):843–852.
- Jones DT, Woods DR. 1986. Acetone-butanol fermentation revisited. *Microbiol Rev* 50(4):484–524.
- Konings WN, Lolkema JS, Poolman B. 1995. The generation of metabolic energy by solute transport. *Arch Microbiol* 164(4):235–242.
- Kummel A, Panke S, Heinemann M. 2006. Putative regulatory sites unraveled by network-embedded thermodynamic analysis of metabolome data. *Mol Syst Biol* 2: 2006. 0034.
- Lepage C, Fayolle F, Hermann M, Vandecasteele J-P. 1987. Changes in membrane lipid composition of *Clostridium acetobutylicum* during acetone-butanol fermentation: Effects of solvents, growth temperature and pH. *J Gen Microbiol* 133(1):103–110.
- MacCarthy T, Pomiankowski A, Seymour R. 2005. Using large-scale perturbations in gene network reconstruction. *BMC Bioinformatics* 6:11.
- Mclaughlin JK, Meyer CL, Papoutsakis ET. 1985. Gas chromatography and gateway sensors for online state estimation of complex fermentations (butanol acetone fermentation). *Biotechnol Bioeng* 27(8):1246–1257.
- Meyer CL, Papoutsakis ET. 1989. Continuous and biomass recycle fermentations of *Clostridium acetobutylicum*. 2. Novel patterns in energetics and product-formation kinetics. *Bioprocess Eng* 4(2):49–55.
- Monot F, Martin JR, Petitdemange H, Gay R. 1982. Acetone and butanol production by *Clostridium acetobutylicum* in a synthetic medium. *Appl Environ Microbiol* 44(6):1318–1324.
- Neuhaus FC, Baddiley J. 2003. A continuum of anionic charge: Structures and functions of D-alanyl-teichoic acids in gram-positive bacteria. *Microbiol Mol Biol Rev* 67(4):686–723.

- Nielsen J, Oliver S. 2005. The next wave in metabolome analysis. *Trends Biotechnol* 23(11):544–546.
- Nolling J, Breton G, Omelchenko MV, Makarova KS, Zeng Q, Gibson R, Lee HM, Dubois J, Qiu D, Hitti J, Wolf YI, Tatusov RL, Sabathe F, Doucette-Stamm L, Soucaille P, Daly MJ, Bennett GN, Koonin EV, Smith DR. 2001. Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. *J Bacteriol* 183(16):4823–4838.
- Papoutsakis ET. 1984. Equations and calculations for fermentations of butyric-acid bacteria. *Biotechnol Bioeng* 26(2):174–187.
- Papoutsakis ET, Meyer CL. 1985a. Equations and calculations of product yields and preferred pathways for butanediol and mixed-acid fermentations. *Biotechnol Bioeng* 27(1):50–66.
- Papoutsakis ET, Meyer CL. 1985b. Fermentation equations for propionic acid bacteria and production of assorted oxychemicals from various sugars. *Biotechnol Bioeng* 27(1):67–80.
- Paredes CJ, Alsaker KV, Papoutsakis ET. 2005. A comparative genomic view of clostridial sporulation and physiology. *Nat Rev Microbiol* 3(12):969–978.
- Parker GF, Daniel RA, Errington J. 1996. Timing and genetic regulation of commitment to sporulation in *Bacillus subtilis*. *Microbiology* 142(Pt 12):3445–3452.
- Pollack JH, Neuhaus FC. 1994. Changes in wall teichoic acid during the rod-sphere transition of *Bacillus subtilis* 168. *J Bacteriol* 176(23):7252–7259.
- Price ND, Papin JA, Schilling CH, Palsson BO. 2003. Genome-scale microbial *in silico* models: The constraints-based approach. *Trends Biotechnol* 21(4):162–169.
- Price ND, Reed JL, Palsson BO. 2004. Genome-scale models of microbial cells: Evaluating the consequences of constraints. *Nat Rev Microbiol* 2(11):886–897.
- 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(9):R54.
- Reed JL, Patel TR, Chen KH, Joyce AR, Applebee MK, Herring CD, Bui OT, Knight EM, Fong SS, Palsson BO. 2006. Systems approach to refining genome annotation. *Proc Natl Acad Sci USA* 103(46):17480–17484.
- Riebeling V, Jungermann K. 1976. Properties and function of clostridial membrane ATPase. *Biochim Biophys Acta* 430(3):434–444.
- Roos JW, McLaughlin JK, Papoutsakis ET. 1985. The effect of pH on nitrogen supply, cell-lysis, and solvent production in fermentations of *Clostridium acetobutylicum*. *Biotechnol Bioeng* 27(5):681–694.
- Saier MH Jr, Tran CV, Barabote RD. 2006. TCDB: The Transporter Classification Database for membrane transport protein analyses and information. *Nucleic Acids Res* 34(Database issue):D181–D186.
- Senger RS, Papoutsakis ET. 2008. Genome-scale model for *Clostridium acetobutylicum*: Part 1. Metabolic network resolution and analysis. *Biotechnol Bioeng*, DOI: 10.1002/bit.22010
- Senger RS, Phisalaphong M, Karim MN, Linden JC. 2006. Development of a culture sub-population induction model: Signaling pathways synergy and taxanes production by *Taxus canadensis*. *Biotechnol Prog* 22(6):1671–1682.
- Shlomi T, Eisenberg Y, Sharan R, Ruppin E. 2007. A genome-scale computational study of the interplay between transcriptional regulation and metabolism. *Mol Syst Biol* 3:101.
- Tegner J, Yeung MK, Hasty J, Collins JJ. 2003. Reverse engineering gene networks: Integrating genetic perturbations with dynamical modeling. *Proc Natl Acad Sci USA* 100(10):5944–5949.
- Tomas CA, Alsaker KV, Bonarius HP, Hendriksen WT, Yang H, Beamish JA, Paredes CJ, Papoutsakis ET. 2003. DNA array-based transcriptional analysis of asporogenous, nonsolventogenic *Clostridium acetobutylicum* strains SKO1 and M5. *J Bacteriol* 185(15):4539–4547.
- Tomas CA, Beamish J, Papoutsakis ET. 2004. Transcriptional analysis of butanol stress and tolerance in *Clostridium acetobutylicum*. *J Bacteriol* 186(7):2006–2018.
- Tummala SB, Junne SG, Paredes CJ, Papoutsakis ET. 2003. Transcriptional analysis of product-concentration driven changes in cellular programs of recombinant *Clostridium acetobutylicum* strains. *Biotechnol Bioeng* 84(7):842–854.
- Vasconcelos I, Girbal L, Soucaille P. 1994. Regulation of carbon and electron flow in *Clostridium acetobutylicum* grown in chemostat culture at neutral pH on mixtures of glucose and glycerol. *J Bacteriol* 176(5):1443–1450.
- Wiback SJ, Famili I, Greenberg HJ, Palsson BO. 2004. Monte Carlo sampling can be used to determine the size and shape of the steady-state flux space. *J Theor Biol* 228(4):437–447.
- Wilkinson SR, Young DI, Morris JG, Young M. 1995. Molecular genetics and the initiation of solventogenesis in *Clostridium beijerinckii* (formerly *Clostridium acetobutylicum*) NCIMB 8052. *FEMS Microbiol Rev* 17(3):275–285.