**Supplementary Material**

**COSMIC-dFBA: A novel multi-scale hybrid framework for bioprocess modeling**

Saratram Gopalakrishnan1, William Johnson2, Miguel A. Valderrama-Gomez2, Elcin Icten2, Jasmine Tat2, Michael Ingram2, Coral Fung Shek2, Pik K. Chan2, Fabrice Schlegel2, Pablo Rolandi2, Cleo Kontoravdi3, Nathan Lewis1,4

1 Department of Pediatrics, University of California San Diego

2 Process Development, Amgen

3 Department of Chemical Engineering, Imperial College London

4 Department of Bioengineering, University of California San Diego

1. **Supplementary Methods**
   1. **Flux and Phase Calculation**

Fluxes (growth rate, uptake, and secretion rates) associated with all quantities (cell density, cell size, antibody, metabolites, and amino acids) were computed using nonlinear regression so as to minimize the variance-weighted sum of squared deviation of predicted quantities from their experimentally measured values. Metabolite concentrations are predicted using a two-state phenotypic model. Cells are assumed to be either in the growth phase or in the stationary phase, each represented by a unique set of fluxes. We define as the fluxes associated with quantity in phase (either or ). The objective is to identify the set of that recapitulates the measured concentrations of all quantities at all sampling times (). We first divide the process into M intervals based on sampling times such that interval represents the time interval . For each time interval, we define a parameter that represents the fraction of the cell population in the stationary phase. Thus, the net flux in interval associated with quantity is computed using Equation (1).

|  |  |
| --- | --- |
|  | (1) |

The time-course profile of all quantities in the interval are then computed by solving the system of ODEs described in Equation (2).

|  |  |
| --- | --- |
|  | (2) |

In the above Equation, refers to the cell density in the reactor, denotes the concentration of quantity in the perfusion medium, denotes the perfusion rate (equal to 1 ), and is the removal fraction of the quantities from the bioreactor. equals 0 for cell density and cell size at all time points, and for antibody before day 8. In all other cases, equals 1.

Thus, the parameters and are computed by solving the following nonlinear optimization problem:

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| Subject to: |  |  | (2) |
|  |  |  | (1) |
|  |  |  | (3) |
|  |  |  | (4) |
|  |  |  | (5) |
|  |  |  | (6) |
|  |  |  | (7) |
|  |  |  | (8) |
|  |  |  | (9) |

Kinetic parameters need only be computed for fluxes corresponding to taken up quantities. The kinetic rate law is defined using a Michaelis-Menten-type equation. We fix the parameters and all , replace all with the kinetic rate law and solve the above optimization problem again to compute the and terms to modulate the uptake of quantities based on the reactor concentration. The NLP problem was solved using the fmincon function within the Optimization Toolbox in MATLABTM.

* 1. **Mining metabolic tasks and task efficiencies**

We first classify the growth rate, antibody productivity, and the secreted metabolites in each phase as metabolic tasks. In order to accurately simulate the intracellular flux distribution for a given set of nutrient uptake rates, we must rank the metabolic tasks based on resource allocation determined using the *i*CHO1766 metabolic model. We also define “task efficiency” as the ratio of measured flux through a metabolic task to the maximum flux predicted using the metabolic model. The metabolic task priority and efficiencies for each phase are computed using the following algorithm.

|  |  |
| --- | --- |
| Step 1: | Impose the uptake rates of measured metabolites as bounds in the metabolic model. |
| Step 2: | Set , and . |
| Step 3: | Set |
| Step 4: | Using FBA, compute the maximum flux through each product in . |
| Step 5: | Compute for each product in . |
| Step 6: | Assign the product with the highest task efficiency to . Assign the corresponding task efficiency to . |
| Step 7: | Set the lower bound of flux through to the measured flux. Remove Priority i from L. |
| Step 8: | If is empty, then STOP. Otherwise, repeat Steps 3 – 7. |

1. **Supplementary Results**

We performed a Principal Component Analysis (PCA) on the state-specific fluxes computed as described in section 2.1. The data was normalized by computing the Z-score as described in equation (10)

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | (10) |

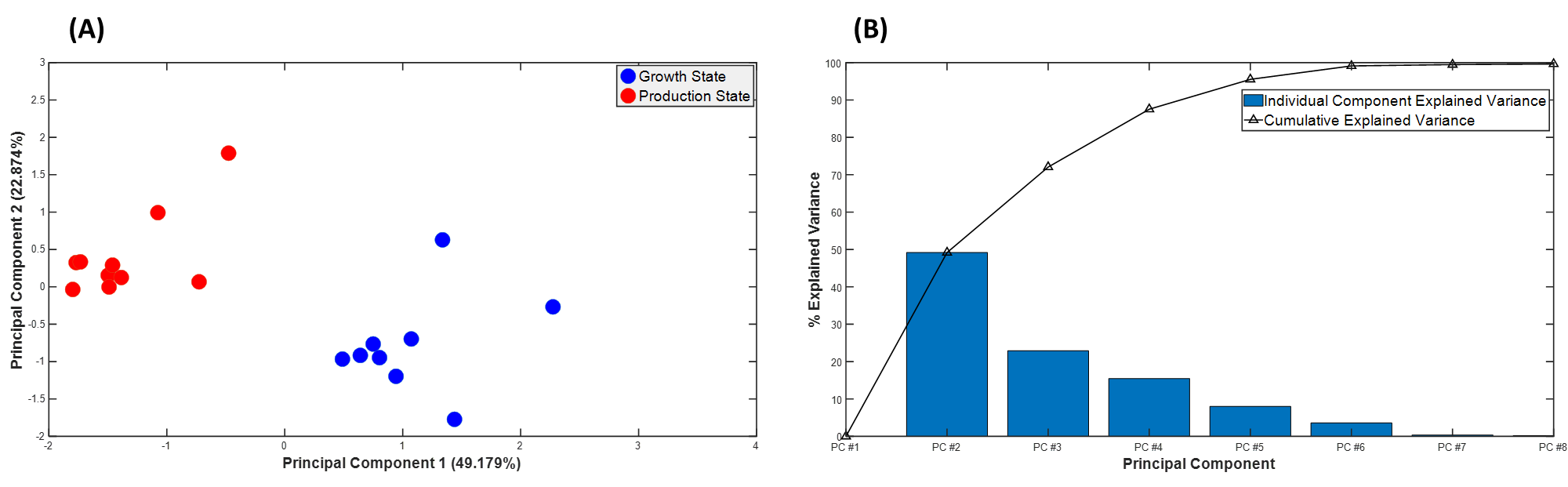
In the above equation,

is the normalized uptake/secretion flux of metabolite in condition .

is the computed uptake/secretion flux of metabolite in condition .

is the mean uptake/secretion flux of metabolite across all conditions.

is the standard deviation of the flux across all conditions.



**Figure S1:** PCA on computed uptake/secretion fluxes. (A) Fluxes projected on to the first two principal components explaining 72% of the variance in the data. Fluxes corresponding to the growth and production states are indicated using blue and red circles, respectively. Separation of data suggests that the states are metabolically distinct. (B) % variance explained by individual principal components. >99% of the data is explained by the first 7 principal components, with the first four principal components explaining over 95% of the variance in the data.

An initial analysis using PCA indicates that the metabolism in the growth and producing states is different (Figure S1A), primarily driven by changes in the growth rate and the specific production rate of the antibody. Additional shifts were seen in glucose uptake, lactate secretion, switch from glycine production to glycine consumption, asparagine, and glutamine.

**Supplementary Figure S2**: Metabolite concentration variances in the growth state and the production state .



**Supplementary Figure S3**: Concentration profiles for 25 quantities predicted by COSMIC-dFBA (blue line), Traditional dFBA (red line), and the Assumed Objectives case (orange line) compared with the experimentally measured data (magenta dots) in Reactor 1 (A), Reactor 2 (B), Reactor 3 (C), Reactor 4 (D), Reactor 5 (E), Reactor 6 (F), Reactor 7 (G), Reactor 8 (H), Reactor 9 (I), and Reactor 10 (J). The feed media changes corresponding to each reactor is specified in Supplementary Table ST 1.

|  |
| --- |
| **(A)** |
| **(B)** |
| **(C)** |
| **(D)** |
| **(E)** |
| **(F)** |
| **(G)** |
| **(H)** |
| **(I)** |
| **(J)** |