Maximum entropy and population heterogeneity in continuous cell cultures, validation with experimental data

Jose A. Pereiro-Morejón, Jorge Fernandez-de-Cossio-Diaz, Roberto Mulet

September 16, 2020

1 Introduction

The study of cellular metabolism is a research area with direct potential impact in biological sciences, medicine and industry. An example is that cell culture-derived products are a major part of the multi-billion biotechnological industry portfolio. This products are obtained by exploding the capability of cellular metabolism to produce molecules with a wide range of chemical complexity. For this propose, cells are cultivated in three common modes: batch, fed-batch and continuous. In batch, cultivation starts with a medium rich in nutrients that will be consumed completely by the cells, often till starvation. Similarly, fed-batch cultures starts will a nutrient pool, but it is resupplied in discrete time intervals. On the other hand, in continuous mode, fresh medium constantly replaces culture fluid at a given rate.

The chemostat is an example of a continuous cultivation devise developed in the 50's (Monod, 1949; Novick & Szilard, 1950). Chemostat cultures are often run at constant working volume and in steady state condition, which is reach when the macroscopic variables of the culture stays constant with time (mainly cell and external metabolites concentrations). Although advantages of continuous cultivation have been commonly mentioned in the literature (Werner, Walz, Noé, & Konrad, 1992; Griffiths, 1992; Kadouri & Spier, 1997; Werner & Noe, 1998; Croughan, Konstantinov, & Cooney, 2015), the preferred use of these technics over batch or fed-batch struggle with the complexity displayed by continuous systems, i.e., hysteresis, multi-stability or sharp transitions between metabolic states (Mulukutla, Yongky, Grimm, Daoutidis, & Hu, 2015; Europa, Gambhir, Fu, & Hu, 2000; C et al., 2001; Hayter et al., 1992; Gambhir et al., 2003; Follstad, Balcarcel, Stephanopoulos, & Wang, 1999; Fernandez-de-Cossio-Diaz, Leon, & Mulet, 2017). Mathematical modeling can guide our efforts to understand the cellular metabolism and then suggest strategies to improve production efficiency using continuous cultivation methods. This could lead to a decrease in the production costs and subsequently, to a reduction on medicine

prices and a broader accessibility to treatments. Therefor, for the industry, it is fundamental the development of theoretical frameworks that allows, for example, predict how to reach an optimal stable cultivation state given a cell of interest and the medium to be used.

A major driver of biological discovery are currently found in increasingly accurate experimental techniques that generate unprecedented amount of thoughtful data. Information about cellular metabolism, at individual reaction (gene) level, had lead to the development of genome-scale metabolic networks ($GEMs$) (Kanehisa et al., 2014; Caspi et al., 2016). These networks have been used to build mathematical models of the cellular metabolism (Palsson, 2006). Constraint-based technics as Flux balance Analysis (FBA) had have good results predicting, for example, cellular metabolism in growth phase of batch cultures (Palsson, 2006). Meanwhile, a methodology for applying (FBA) in the context of a chemostat continuous cultivation have been proposed by (Fernandez-de-Cossio-Diaz et al., 2017). It introduces a detailed characterization of the steady state of the chemostat, which is achieved by coupling cell metabolism with the dynamics of extra-cellular concentrations. Although this approach can explain different phenomena in the context of a detailed metabolic model, it have the drawback that it assume cell population to be homogeneous.

Cells in a culture are not identical, many different processes can contribute to heterogeneity (Kiviet et al., 2014; Elowitz, Levine, Siggia, & Swain, 2002; Fernandez-de-Cossio-Diaz, Mulet, & Vazquez, 2019; Huh & Paulsson, 2011; Wang et al., 2016). In this context, with the goal of understanding which is the impact of considering cell population in a continuous culture to be homogeneous, a model that include cell heterogeneity is develop by (Fernandez-de-Cossio-Diaz et al., 2019). It is proposed to apply the maximum entropy principle ($MaxEnt$) (Jaynes, 1957, 2003; C et al., 2001; Harte & Newman, 2014; Schneidman, Berry, Segev, & Bialek, 2006; De Martino, MC Andersson, Bergmiller, Guet, & Tkačik, 2018; De Martino, Capuani, & De Martino, 2016, 2017) to model the heterogeneity in a continuous bioreactor. The authors also shows that the methodology is applicable to GEMs of realistic sizes with commonly available computational resources.

The objective of this work is to search thorough literature-available data of chemostat cultures and test the impact of including heterogeneity. For that propose we will face the two model proposed in (Fernandez-de-Cossio-Diaz et al., 2017) and (Fernandez-de-Cossio-Diaz et al., 2019) to the same data and using the same metabolic networks for different cell lines.

2 Materials and Method

2.1 Model framework

The model framework used in this work is explained in detail in (Fernandez-de-Cossio-Diaz et al., 2017) and (Fernandez-de-Cossio-Diaz et al., 2019)

2.2 Chemostat experimental data

2.2.1 Escherichia coli KJ134

Continuous cultivation data for Escherichia coli ($E.coli$) were taken from (van Heerden & Nicol, 2013). The used Escherichia coli KJ134 strain was genetically modified for succinic acid fermentation. A dozen of steady states were recorded at different dilution rates and GLC availability in the feed medium. Data for the effluent concentration of GLC, succinic acid (SA), acetic acid (AcA), formic acid (FA) and malic acid (MA) for the different steady states were given. Also, the cell concentration was reported.

2.2.2 Human cell line AGE1.HN.AA1

Experimental data were taken from (Rath, 2017). In this work, the author performed 6 continuous cultures of the cell line AGE1.HN.AA1. Its parental line AGE1.HN was established by the company ProBioGen (ProBioGen AG, Berlin, Germany) from a tissue sample of a human brain. The experiments were run under various conditions, differing mainly in the dilution rate (D) and the feed medium composition of glucose (GLC), glutamine (GLN) and galactose (GAL) $Table$ 2.

For each experiment, a steady-state condition was reached, (steady states were labeled A, B, C, D, E, $F01$), and several observables were reported. Particularly relevant for this work was the growth rate (μ), D, the viable cell concentration (Xv) and the medium concentration (s) and derived uptake rate (q) for a set of metabolites (GLC, lactose (LAC), GLN, ammonium ($NH4$), GAL, pyruvate (PYR), glutamate (GLU), alanine (ALA), asparagine (ASP)). A unit conversion was required to make experimental data and our model compatible. For this propose the only external data needed was the cell mass density, 0.25 pgDW/ μm^3 (Niklas, Schräder, Sandig, Noll, & Heinzle, 2011).

2.3 Preparing GEMs

2.3.1 Escherichia coli KJ134

For modeling the cellular metabolism, the $E.coli$ genome-scale metabolic model (GEM) iJR904 (download link: bigg.ucsd.edu) (Reed, Vo, Schilling, & Palsson, 2003) was modified to match with (van Heerden & Nicol, 2013) data. The reported genetic modifications for the Escherichia coli KJ13 strain were included. It was also added enzymatic cost constraints in concordance with (Beg et al., 2007).

2.3.2 Human cell line AGE1.HN.AA1

With the same propose of modeling the metabolism of the human-derived cell AGE1.HN.AA1, it was used cell-line specific human GEMs derived from the generic GEM (download link: zenodo) described at (Robinson et al., 2020). In this work, many cell specific GEMs were derived from the general model using

omic data from several sources. We use models produced using data from brain samples, including healthy and cancer, as substitute of an AGE1.HN.AA1 specific network (Robinson et al., 2020). The GEMs biomass equations were modified in agreement with the biomass composition for AGE1.HN.AA1 reported in (Niklas et al., 2011). The production of α1-antitrypsin ($A1AT$) was also considered, it was used the protein sequence at drugbank, and the reported production rate for $A1AT$ (Rath, 2017) to set maintenance for the required amino acids. Additionally, an atp demand ($ATPM$) was included, taking the maintenance energy demand for mammalian cells mentioned by (Fernandez-de-Cossio-Diaz & Vazquez, 2018). Furthermore, enzymatic constraint were added using the methodology reported at (Sánchez et al., 2017) and data from (Robinson et al., 2020).

References

Beg, Q. K., Vazquez, A., Ernst, J., de Menezes, M. A., Bar-Joseph, Z., Barabási, A.-L., & Oltvai, Z. N. (2007, July). Intracellular crowding defines the mode and sequence of substrate uptake by Escherichia coli and constrains its metabolic activity. *Proceedings of the National Academy of Sciences*, *104*(31), 12663–12668. doi: 10.1073/pnas.0609845104

C, A., A, I., A, C., X, G., Jj, C., & C, G. (2001, November). *Analysis of CHO cells metabolic redistribution in a glutamate-based defined medium in continuous culture* (Vol. 17) (No. 6). https://pubmed.ncbi.nlm.nih.gov/11735437/. Biotechnol Prog. doi: 10 .1021/bp0100981

Caspi, R., Billington, R., Ferrer, L., Foerster, H., Fulcher, C. A., Keseler, I. M., ... Karp, P. D. (2016, January). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Research*, *44*(D1), D471-D480. doi: 10.1093/nar/gkv1164

Croughan, M. S., Konstantinov, K. B., & Cooney, C. (2015). The future of industrial bioprocessing: Batch or continuous? *Biotechnology and Bioengineering*, *112*(4), 648–651. doi: 10.1002/bit.25529

De Martino, D., Capuani, F., & De Martino, A. (2016, May). Growth against entropy in bacterial metabolism: The phenotypic trade-off behind empirical growth rate distributions in E. coli. *Physical Biology*, *13*(3), 036005. doi: 10.1088/1478-3975/13/3/036005

De Martino, D., Capuani, F., & De Martino, A. (2017, July). Quantifying the entropic cost of cellular growth control. *Physical Review E*, *96*(1), 010401. doi: 10.1103/PhysRevE.96.010401

De Martino, D., MC Andersson, A., Bergmiller, T., Guet, C. C., & Tkačik, G. (2018, July). Statistical mechanics for metabolic networks during steady state growth. *Nature Communications*, *9*(1), 2988. doi: 10.1038/s41467-018-05417-9

Elowitz, M. B., Levine, A. J., Siggia, E. D., & Swain, P. S. (2002). Stochastic gene expression in a single cell. *Science*, *297*(5584), 1183–1186. doi: 10.1126/science.1070919

Europa, A. F., Gambhir, A., Fu, P.-C., & Hu, W.-S. (2000). Multiple steady states with distinct cellular metabolism in continuous culture of mammalian cells. *Biotechnology and Bioengineering*, *67*(1), 25–34. doi: 10.1002/(SICI)1097-0290(20000105)67:1⟨25::AID-BIT4⟩3.0.CO;2-K

Fernandez-de-Cossio-Diaz, J., Leon, K., & Mulet, R. (2017). Characterizing steady states of genome-scale metabolic networks in continuous cell cultures. *PLoS Computational Biology*, *13*(11), 1–22. doi: 10.1371/journal.pcbi.1005835

Fernandez-de-Cossio-Diaz, J., Mulet, R., & Vazquez, A. (2019, June). Cell population heterogeneity driven by stochastic partition and growth optimality. *Scientific Reports*, *9*(1), 1–7. doi: 10.1038/s41598-019-45882-w

Fernandez-de-Cossio-Diaz, J., & Vazquez, A. (2018, May). A physical model of cell metabolism. *Scientific Reports*, *8*(1), 1–13. doi: 10.1038/s41598-018-26724-7

Follstad, B. D., Balcarcel, R. R., Stephanopoulos, G., & Wang, D. I. C. (1999). Metabolic flux analysis of hybridoma continuous culture steady state multiplicity. *Biotechnology and Bioengineering*, *63*(6), 675–683. doi: 10.1002/(SICI)1097-0290(19990620)63:6⟨675::AID-BIT5⟩3.0.CO;2-R

Gambhir, A., Korke, R., Lee, J., Fu, P.-C., Europa, A., & Hu, W.-S. (2003, January). Analysis of cellular metabolism of hybridoma cells at distinct physiological states. *Journal of Bioscience and Bioengineering*, *95*(4), 317–327. doi: 10.1016/S1389-1723(03)80062-2

Griffiths, J. B. (1992). Animal cell culture processes - batch or continuous? *Journal of Biotechnology*, *22*(1-2), 21–30. doi: 10.1016/0168-1656(92)90129-W

Harte, J., & Newman, E. A. (2014, July). Maximum information entropy: A foundation for ecological theory. *Trends in Ecology & Evolution*, *29*(7), 384–389. doi: 10.1016/j.tree.2014.04.009

Hayter, P. M., Curling, E. M., Baines, A. J., Jenkins, N., Salmon, I., Strange, P. G., ... Bull, A. T. (1992). Glucose-limited chemostat culture of chinese hamster ovary cells producing recombinant human interferon-γ. *Biotechnology and Bioengineering*, *39*(3), 327–335. doi: 10.1002/bit.260390311

Huh, D., & Paulsson, J. (2011). Random partitioning of molecules at cell division. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(36), 15004–15009. doi: 10.1073/pnas.1013171108

Jaynes, E. T. (1957, May). Information Theory and Statistical Mechanics. *Physical Review*, *106*(4), 620–630. doi: 10.1103/PhysRev.106.620

Jaynes, E. T. (2003, April). *Probability Theory: The Logic of Science.* /core/books/probability-theory/9CA08E224FF30123304E6D8935CF1A99. Cambridge University Press. doi: 10.1017/CBO9780511790423

Kadouri, A., & Spier, R. E. (1997, July). Some myths and messages concerning the batch and continuous culture of animal cells. *Cytotechnology*, *24*(2),

89. doi: 10.1023/A:1007932614011

Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., & Tanabe, M. (2014). Data, information, knowledge and principle: Back to metabolism in KEGG. *Nucleic Acids Research*, *42*(D1), 199–205. doi: 10.1093/nar/gkt1076

Kiviet, D. J., Nghe, P., Walker, N., Boulineau, S., Sunderlikova, V., & Tans, S. J. (2014, October). Stochasticity of metabolism and growth at the single-cell level. *Nature*, *514*(7522), 376–379. doi: 10.1038/nature13582

Monod, J. (1949). The Growth of Bacterial Cultures. *Annual Review of Microbiology*, *3*(1), 371–394. doi: 10.1146/annurev.mi.03.100149.002103

Mulukutla, B. C., Yongky, A., Grimm, S., Daoutidis, P., & Hu, W.-S. (2015, March). Multiplicity of Steady States in Glycolysis and Shift of Metabolic State in Cultured Mammalian Cells. *PLOS ONE*, *10*(3), e0121561. doi: 10.1371/journal.pone.0121561

Niklas, J., Schräder, E., Sandig, V., Noll, T., & Heinzle, E. (2011). Quantitative characterization of metabolism and metabolic shifts during growth of the new human cell line AGE1.HN using time resolved metabolic flux analysis. *Bioprocess and Biosystems Engineering*, *34*(5), 533–545. doi: 10.1007/s00449-010-0502-y

Novick, A., & Szilard, L. (1950, December). Description of the Chemostat. *Science*, *112*(2920), 715–716. doi: 10.1126/science.112.2920.715

Palsson, B. Ø. (2006). *System Biology - Properties of Reconstructed Networks.*

Rath, A. (2017). *Characterisation of cell growth, metabolism and recombinant protein production during transient and steady state conditions for the human cell line AGE1.HN-AAT* (PhD Thesis).

Reed, J. L., Vo, T. D., Schilling, C. H., & Palsson, B. O. (2003, August). An expanded genome-scale model of Escherichia coli K-12 (iJR904 GSM/GPR). *Genome Biology*, *4*(9), R54. doi: 10.1186/gb-2003-4-9-r54

Robinson, J. L., Kocabaş, P., Wang, H., Cholley, P.-E., Cook, D., Nilsson, A., ... Nielsen, J. (2020, March). An atlas of human metabolism. *Science Signaling*, *13*(624). doi: 10.1126/scisignal.aaz1482

Sánchez, B. J., Zhang, C., Nilsson, A., Lahtvee, P.-J., Kerkhoven, E. J., & Nielsen, J. (2017). Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints. *Molecular Systems Biology*, *13*(8), 935. doi: 10.15252/msb.20167411

Schneidman, E., Berry, M. J., Segev, R., & Bialek, W. (2006, April). Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature*, *440*(7087), 1007–1012. doi: 10.1038/nature04701

van Heerden, C. D., & Nicol, W. (2013, September). Continuous and batch cultures of Escherichia coli KJ134 for succinic acid fermentation: Metabolic flux distributions and production characteristics. *Microbial Cell Factories*, *12*(1), 80. doi: 10.1186/1475-2859-12-80

Wang, J., Cazzato, E., Ladewig, E., Frattini, V., Rosenbloom, D. I., Zairis, S., ... Rabadan, R. (2016). Clonal evolution of glioblastoma under therapy. *Nature Genetics*, *48*(7), 768–776. doi: 10.1038/ng.3590

Werner, R. G., & Noe, W. (1998, March). Letter to the Editor. *Cytotechnology*, *26*(2), 81–82. doi: 10.1023/A:1007985828899

Werner, R. G., Walz, F., Noé, W., & Konrad, A. (1992, January). Safety and economic aspects of continuous mammalian cell culture. *Journal of Biotechnology*, *22*(1-2), 51–68. doi: 10.1016/0168-1656(92)90132-s