Cell Metabolic Diagnosis and Control in CHO Fed-batch Process

Bingyu Kuang*, Duc Hoang*, Zhao Wang*, Seongkyu Yoon*

**Chemical Engineering Department, University of Massachusetts, Lowell, MA 01854, USA*

(e-mail: Seongkyu_Yoon@uml.edu).

Abstract: Therapeutic protein productivity and product quality highly rely on cell metabolism of the fed-batch process, which is a costly, time-consuming and lack of intracellular analytical diagnostic tools. Cell culture medium composition and feeding strategy is critical to regulate cell metabolism. In this study, we present an unorthodox approach to optimize CHO bioprocess by integrating conventional design-of-experiments (DOE) methodology with genome scale model (GEM) flux analysis. Generic CHO-K1 metabolic model was tailored and further integrated with CHO fed-batch metabolomic data to obtain a cell line- and process-specific model. *In silico* metabolic flux analysis was conducted via GEM to identify the critical medium components toward cellular growth and further evaluate their optimized flux values from thirty five simulated fed-batch DOE conditions. Glucose and valine were projected as the most critical nutrients in the process from the flux simulation analysis. Using this approach, previously identified metabolic inhibitor cytidine monophosphate accumulated in extracellular environment was found to be regulated by glucose, glutamine, aspartate, and alanine and further experimentally validated through dose-dependent amino acid spiking study. A process diagnostic and control model was constructed from network topology modeling constructed through GEM and pathway enrichment analysis, which allowed optimization of medium components utilized in a fed-batch feeding process to better support cell metabolism and mitigate accumulation of metabolic inhibitors.

Keywords: Therapeutic protein production, Fed-batch bioprocess, *in silico* model, FBA, Culture medium development, Metabolism shift

1. INTRODUCTION

Biotherapeutics have emerged as one of the most effective treatment options for many diseases including cancers and autoimmune disorders. Chinese hamster ovary (CHO) cells represent the most widely used host cell for therapeutic recombinant protein production (Aggarwal, 2011). CHO expression hosts are characterized by their high nutrients uptake rate during cellular expansion and stationary phase. However, inefficient intracellular metabolism often prevents cells to fully utilize nutrients to support growth and protein production. Instead, significant fraction of fed glucose and amino acids are diverted into generation of toxic metabolites (Cruz et al., 2000, Lao and Toth, 1997). Owing to this, tandem liquid chromatography-mass spectroscopy (LC-MS/MS) has emerged as one of the most powerful set of tools for various metabolomic studies due to its capability of analyzing countless of metabolites and medium additives from a single sample (Mohmad-Saberi et al., 2013, Hoang et al., 2021). These advances in metabolomics have enabled identification of additional cell generated inhibitory metabolites, other than lactate and ammonia (Mulukutla et al., 2017). In our most recent work, it has been shown that untargeted global metabolomics coupled LC-MS/MS can be applied to batch and fed-batch CHO bioprocess to identify six growth and titer production metabolic inhibitors generated from CHO metabolism (Kuang et al., 2021).

Constraint-based flux balance analysis (FBA) is a computational approach to study metabolic networks with a well-established literature (Feist et al., 2009, Bordbar et al., 2014). Coupled with a genome-scale model (GEM) network comprising of biochemical reactions capable of composing cellular metabolism, FBA can be employed to examine metabolic systems of various organisms. As such, a recent study showed the capability of tailoring the generic CHO GEM into host and recombinant cell-specific model to understand the genotypic and phenotypic traits differences between wild-type and recombinant cells (Yusufi et al., 2017). Others successfully combined *in silico* modeling and metabolomic analysis to characterize fed-batch of CHO cultures (Selvarasu et al., 2012). Despite contributing to the overall fundamental understanding of CHO cell metabolism, these studies often remain singular, and the mechanisms underlying cellular metabolic shift under different process conditions, as well as the metabolic relationships between toxic by-products and their nutrient precursors in CHO intracellular network are not yet fully understood. In this study, we present an unorthodox approach to optimize CHO bioprocess by integrating conventional design-of-experiments (DOE) methodology with GEM flux. Generic CHO-K1 metabolic model was tailored and further integrated with CHO fed-batch metabolomic data to obtain a cell line- and process-specific model. The study successfully establishes an *in silico* metabolomic platform relating metabolic inhibitors with their nutrient precursors which are experimentally validated based on amino acid dose-dependent spiking study. The data obtained from the study altogether enables visualization of metabolic relationships between nutrients in the networks, providing a deeper mechanistic understanding into different CHO physiologies

due to cellular metabolic shift with applications spanning from cell line evaluation, metabolic engineering to media optimization and biomanufacturing control.

2. METHOD

2.1 Flux balance analysis and genome-scale modeling

The foundation of FBA assumes the metabolism of a single cell is defined by a system of n reactions. Assuming $[C_i]$ denotes the concentration of metabolite i. The limiting steady state assumption of FBA at equilibrium condition ($t \to \infty$) constrains the fluxes within an average, single cell $\hat{c}$ consumed by FBA so that:

$$\frac{d[C_i^{\hat{c}}]}{dt} = \sum_{j=1}^{n} S_{ij}^{\hat{c}} v_j^{\hat{c}} = 0, \forall i \tag{1}$$

Here, S_{ij} is the stoichiometric coefficient for metabolite i in reaction j and the flux of reaction j is v_j. FBA is predicated on the assumption that cells have been tuned through a biochemical exchanging process to a stage where they "optimally" utilize their resources, where "optimal" is measured by a function of the fluxes, $g(\mathrm{v})$. Hence, FBA studies metabolic processes through optimization problems and adaptations of the form:

$$\max \left\{ g\left(v^{\hat{c}}\right) : S^{\hat{c}} v^{\hat{c}} = 0, L^{\hat{c}} \le v^{\hat{c}} \le U^{\hat{c}} \right\} \tag{2}$$

where $S^{\hat{c}}$ is the stoichiometric matrix whose components are S_{ij}. The objective function $g(\mathrm{v})$ is defined to be the rate at which biomass is created. For convenience, $g(\mathrm{v})$ is assumed to only contain the sole flux of the growth reaction (i.e., $g\left(v^{\hat{c}}\right) = \mathrm{v}_{growth}$). The vectors of lower bounds, $L^{\hat{c}}$, and upper bounds, $U^{\hat{c}}$, may contain fluxes of value $\pm\infty$, or some suitably large value, to indicate that a flux is unbounded. If either $L^{\hat{c}} < 0$ or $U^{\hat{c}} > 0$, then the biochemical reaction is said to be reversible.

2.2 Growth and exchange rate calculation

Growth rates of CHO cells were assumed to follow exponential growth behavior:

$$N_x = N_{x,0} \cdot e^{\mu t} \tag{3}$$

Here $N_{x,0}$ ($\times 10^6$ cells·mL^{-1}) is the number of cells at time 0 (hr), N_x ($\times 10^6$ cells·mL^{-1}) is the number of cells after culture time t (hr). The $IVCD$ profile of cell at time t_n can be calculated as followed:

$$IVCD_{t_n} = IVCD_{t_{n-1}} + \frac{(VCD_{t_n} + VCD_{t_{n-1}})/2}{\Delta t} \tag{4}$$

Rearranging (1), cellular growth rate μ (1·hr^{-1}) therefore can be expressed as a log-based growth model:

$$\mu = \ln (N_x/N_{x,0})/t \tag{5}$$

Assuming CHO cells growing exponentially, the metabolite exchange rates r (mmol·hr^{-1}) can be estimated by evaluating the change in measured concentration C_i (mmol) of metabolite i over time:

$$r = \frac{dC_i}{dt} \approx \frac{\Delta C_i}{\Delta t} \tag{6}$$

2.3 Boolean network topology logic model

Since metabolites and their corresponding pathways were reported in various mammalian systems other than CHO, a multi-stage *in silico* genome-scale network topology analysis was designed to explore and study metabolic by-products and their related pathways in CHO expression host. In this topology model, a set of logic rules was constructed to correlate a metabolite, through genome-scale metabolic model of cells, to upstream nutrient precursors. The main idea here is that, when applying this set of logic to the metabolite and the identified precursor, increasing the exchanged flux of upstream intake nutrient would also increase the generation flux of the downstream metabolite. Using this approach, the following operator notation was utilized: ¬ (not) and ∪ (or). For any two metabolites 1 and 2 in a biochemical reaction, the applied Boolean logic (B) must follow:

$$\neg(B_1 \cup B_2) = -1 \tag{7}$$

For generalization, a series of chemical reactions containing metabolite 1 to metabolite i must satisfy the logic of Boolean function f:

$$\prod_{i=1}^{n} f\left(B_1^i + B_2^i\right) = -1 \tag{8}$$

2.4 Process control and optimization

Briefly stated, GEM is a reconstruction of metabolic network, represented by a stoichiometric matrix of metabolites and reactions. Intracellular metabolic fluxes can be estimated by FBA with the optimization of a defined biological objective function based on constraints using a linear programming approach. The GEM of three different CHO cell lines including CHO-K1, CHO-S and CHO-DG44 were previously published (Hefzi et al., 2016). As the constraints of the FBA approach of the model, raw inputs of the model are the consumption or accumulation rate of amino acids, metabolites. Biomass growth rate in exponential phase and mAb productivity in stationary phase of CHO cell culture are the two objectives of optimization for solving the model using the linear programming approach. With the raw inputs as the constrains for both upper and lower boundaries, and the optimization objective specified, the biomass growth rate or mAb productivity can be estimated.

2.5 Metabolic flux analysis medium optimization

One of the controlling and optimization approaches in cell culture is to optimize the components of amino acids in the medium. The range of modification and the number of simulations within the range can be specified for different amino acids of interest. FBA can then be computed for all the conditions within the range of selected objective function targeted for optimization. For example, alanine and arginine are selected from the list with the range specified from 0.1 to 2.0 for 10 steps. The biomass in exponential phase is selected as the objective of optimization. FBA can then estimate different biomass growth rates based on the raw inputs while

modifying the flux of alanine and arginine in the range of 90% to 200%. The variation of each amino acids is calculated independently, therefore in total there are twenty estimations performed. Ten estimations from 90% to 200% are simulated for alanine and so is for arginine.

3. RESULTS AND DISCUSSION

3.1 CHO fed-batch process understanding and control

To understand and control CHO cell metabolism, a high cell density (HCD) fed-batch process with nine different feeding conditions using a proprietary medium A was conducted, from which metabolomic data obtained from cells were obtained, as illustrated in **Table 1**.

Table 1. Culture feeding strategy for high cell density process development. Feeding strategy was designed and referenced to the inoculation seeding volume (30 mL).

Condition	Feed Day	Feed Level (% Seed Volume)
Low Feed	2	Low – 5% (1.5 mL feeding)
High Feed	2	High – 15% (4.5 mL feeding)
Early Feed	2	Medium – 10% (3 mL feeding)
Late Feed	4	Medium – 10% (3 mL feeding)
Control	3	Medium – 10% (3 mL feeding)

To study the effect of different feeding amount and timing on cell culture performance, the results from the study showed that low feed (1.5 mL feeding, starting on Day 2) and late feed (3 mL feeding, starting on Day 4) resulted in the highest peak VCD attained on Day 7 and Day 9, respectively. On the other hand, high feed (1.5 mL feeding, starting on Day 2) and early feed (3 mL feeding, starting on Day 4), despite showing the longest culture duration, generally suffered from relatively low peak VCD (see **Figure 1** A). A similar trend was also observed from the IVCD profile of cells, as conditions with either exceedingly high or low growth rate due to employment of different feeding strategies altogether resulted in subpar IVCD profile (see **Figure 1** B).

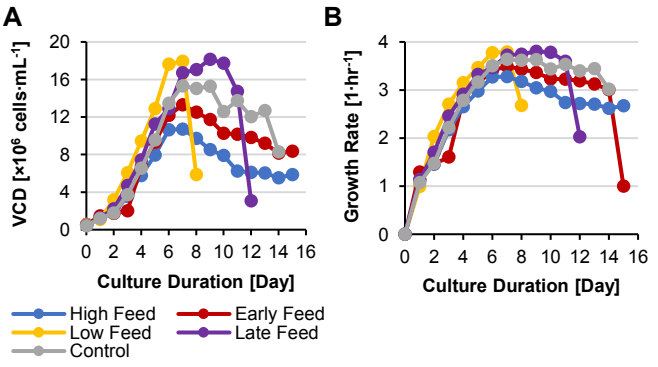


Figure 1. Cellular phenotype characteristics of CHO-K1 cells cultivated at different culture conditions employed in HCD fed-batch process. (A) VCD profile. (B) Growth rate profile. For (A), fed-batch process was conducted at 0.5×10^6 cells·mL^{-1} seeding density in 30 mL culture volume. For (B), the growth rate was calculated from a natural log growth model, as described in Eq.1.

3.2 Flux modeling analysis of CHO-K1 cells in HCD fed-batch process

To further understand the metabolism shift in different process, consumption fluxes of essential nutrients were incorporated as upper and lower bounds constraints for flux modeling analysis. All metabolite data were collected in the HCD fed-batch process, of which corresponding uptake fluxes were used to constraint CHO-K1 metabolic model for flux balance analysis (FBA). In the first stage of the study, experimental data obtained from the growth phase of HCD cell culture experiments were used as steady state inputs for the model with the objective function of maximizing biomass during linear growth phase. The simulation data were then compared against the experimental growth rate data as calculated via a log-based growth model, as shown in **Figure 2**. In general, the percentage differences of the result obtained by constraining the metabolic model with experimental uptake rates and the result calculated from experimental VCD data were ranging from 2% to 14% difference, further validating the predictability of the model when coupling with experimentally measured LC-MS metabolomic data.

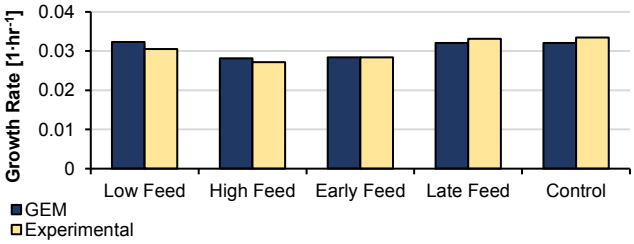


Figure 2. Growth profile of CHO-K1 cells cultivated in HCD fed-batch process. Cellular growth rate as calculated from performing flux balance analysis using experimentally measured metabolic uptake rate with the objective of maximizing biomass using genome scale modeling (GEM) calculated via FBA was compared against experimentally growth rate (Experimental) calculated via a log-based growth model.

3.4 Development of inhibitory metabolites control strategy via metabolic network topology

CHO cells secrete inhibitory metabolites during growth and production phase which hampers cellular performance and negatively impacts titer productivity and various product quality attributes. Previous works have reported by-products and their amino acid inputs in various mammalian cell lines other than CHO (see Table 2). Recent study conducted in our group has identified seven different by-product metabolites and verified their inhibitory impact on cellular performance across different modes of CHO bioprocess (Kuang et al., 2021). Thus, identification of their pathway of generation as well as their nutrient precursors is critical with regards to develop proper control strategy to minimize the accumulated concentration in CHO extracellular environment.

Table 2. Inhibitory metabolites as described in this study and their respective nutrient precursors. Here, downstream metabolite by-products as identified in other mammalian cell lines from the literature are mapped to their nutrient precursors. Shown here: *, essential amino acids.

Substrates	Metabolites
Arg	CMP (Brosnan and Brosnan, 2007), GMP (Brosnan and Brosnan, 2007), TAA (Saas et al., 2000)
Gln	MSA(Hu et al., 2017), CMP (Lane and Fan, 2015), GMP (Lane and Fan, 2015), TAA (Chen et al., 2018), TRI (Perchat et al., 2018)
Asn	NAP (Bocca et al., 2018), TAA (Kim et al., 2011)
Leu*	HICA (Ojala et al., 2013)
Ile*	HICA (Boebel and Baker, 1982), MSA (Nowaczyk et al., 1998)
Lys*	MSA (Nowaczyk et al., 1998)
Pro	NAP (Wu et al., 2008)
Asp	TRI (Ashihara, 2008), CMP (Gagne, 2014), GMP (Gagne, 2014)
Trp	ICA(Mulukutla et al., 2017), TRI (Åkesson et al., 2018)

Since metabolites and their corresponding pathways were reported in various mammalian systems other than CHO, a multi-stage *in silico* genome-scale network topology analysis was designed to explore and study metabolic by-products and their related pathways in CHO expression host. In this topology model, a set of logic rules was constructed to correlate a metabolite, through genome-scale metabolic model of cells, to upstream nutrient precursors. The main goal is that, when applying this set of logic to the metabolite and the identified precursor, increasing the exchanged flux of upstream intake nutrient would also increase the generation flux of the downstream metabolite. In this network topology, only reactions located at most three tiers away from the target inhibitor were considered. For a metabolic pathway to be considered critical in the network topology strategy, each metabolite must satisfy a rigorous rule of logic ordering at both two and three tiers of reactions, as illustrated in **Figure 3** A and C. To illustrate the application of our network topology model, CMP – a metabolite previously identified to be growth and productivity inhibitor in CHO bioprocess – was explored in this study. Here, CMP was identified in the cytoplasm compartment of CHO cellular metabolic network. *In silico* modeling FBA was conducted based on metabolomic uptake rate dataset of cells under different feeding strategies, after which a list of reactions with active metabolic fluxes was extracted. Network topology was then applied to identify the critical amino acid related pathways, where a rigorous set of logic strategy was applied to CMP and any subsequent metabolite in a backward searching fashion starting from a metabolite of interest back to the elementary nutrient precursors upstream to its formation.

Overall, the metabolic network topology strategy identified four unique logic strategies and thirty unique hits when searching for two tiers of reactions, and subsequently sixteen unique strategies with over seven hundred hits for three tiers of reactions (see B and D). Additionally, pathway analysis study of potential causation pathway leading to the formation of CMP is illustrated in **Figure 4** A and B. Pathway enrichment was further performed to evaluate generation of CMP from different feeding conditions with respect to input amino acids. The simulation data revealed high access to nutrients (early feed and high feed) predicted a higher accumulation of CMP when searching at two tiers (**Figure 4** C) and three tiers (**Figure 4** D) of reactions. Overall, the data obtained from the conducted analysis elucidated different fed-batch feeding strategies can have an impact on the downstream accumulation profile of inhibitory metabolites, further suggesting optimization over the concentration range of supplemented medium constituents can better improve cellular phenotype and overall culture performance.

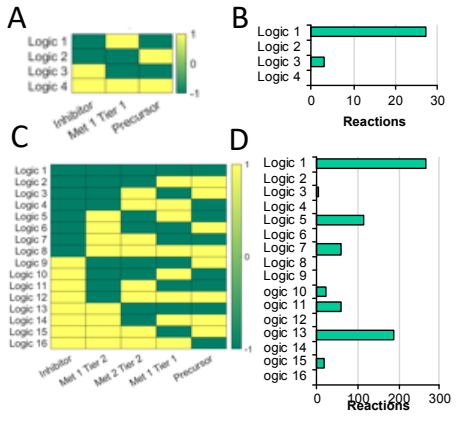


Figure 3. Boolean logic matrix for metabolic network topology. (A) Network logic matrix and (B) Numbers of reactions identified for 2 layers of reactions. Similarly, (C) Network logic matrix and (D) Numbers of reactions identified for 3 layers of reactions. Here, a network topology strategy was developed to correlate target inhibitory metabolite to their nutrient input precursor through CHO metabolic network. In this topology network, green cells (■) indicate a metabolite being consumed in a biochemical reaction, whereas yellow cells (□) indicate a metabolite being generated.

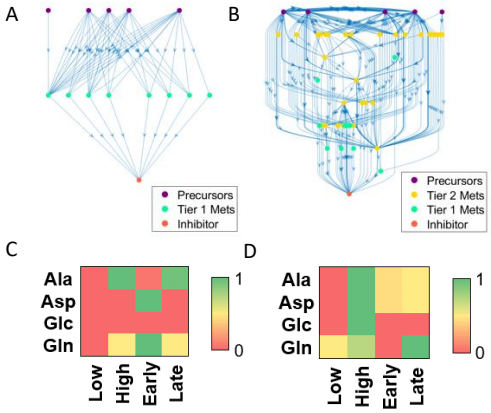


Figure 4. *In silico* metabolic network mapping and pathway analysis of inhibitory metabolite. Metabolic network mapping of downstream inhibitory metabolite to upstream nutrient precursors at two tiers (A) and three tiers (B) of reactions. Pathway analysis via CHO genome scale modeling at two tiers (C) and three tiers (D) of reactions predicting generation of inhibitory metabolite. Here, previously identified growth and productivity inhibitor CMP was identified in the cytoplasm [c] of CHO metabolic network. CMP was further mapped to input precursors including aspartate (asp), glutamine (gln), alanine (ala) and glucose (glc) to identify critical inhibitory metabolite-related pathways. Flux balance analysis was conducted based on metabolomic dataset, after which pathway enrichment was conducted on active reactions to evaluate generation of waste inhibitors under different fed-batch feeding strategies.

3.5 Validation of metabolites pathway via amino acids spiking study

As an effort to control and optimize the cellular phenotype in a standard CHO bioprocess, a dose-dependent nutrients spiking study was developed to identify key medium

constituents precursory to the formation metabolites. In this study, amino acid inputs including glutamine, alanine, aspartate, and glucose – previously explored in other mammalian systems (see **Table 2**) and also predicted in the simulation analysis (Figure 3) – were added to a pool of studied substrates and spiked into CHO-K1 fed-batch cultures on Day 0 at different doses, as described in **Table 3**.

The results of the dose-dependent nutrients spiking study are shown in **Figure 5**. When glutamine, glucose, aspartate, and glucose were spiked into CHO cultures at two- and three-factors higher than the control, a higher concentration of CMP and GMP released into extracellular environment was measured on Day 9, Day 12, and Day 14. Interestingly, the concentration of CMP and GMP from the three-factors condition as measured on Day 14 when compared against the control increased by 1.7-fold and 1.9-fold, respectively, further confirming the strong correlation between the supplemented pool of nutrients to the downstream formation of metabolites as previously explored in the *in silico* FBA study.

Table 3. Dose-dependent nutrients spiking study in CHO-K1 fed- batch process. In this analysis, each condition was spiked with different levels of substrates, after which downstream accumulated concentration of corresponding by-product metabolite was measured via LC-MS.

Substrates	Control	CMP/GMP Two-Factors	CMP/GMP Three-Factors
Ala [mM]	0.14	0.28	0.42
Asp [mM]	3.9	7.8	11.7
Gln [mM]	6	12	18
Glc [g·L^{-1}]	5	7.5	10

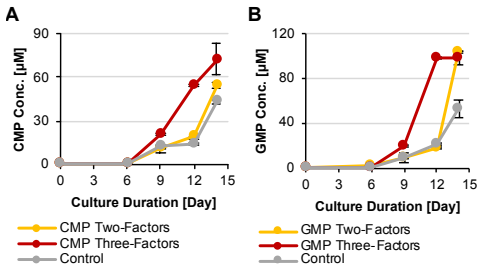


Figure 5. Metabolite analysis results from dose-dependent nutrients spiking study. The accumulated concentration of (A) CMP and (B) GMP with supplemented nutrients spiked into CHO-K1 fed-batch process increased by two- and three-factors (see **Table 3**) was measured.

Downstream metabolic genes involved in CMP-related metabolic pathway nested in the pyrimidine metabolism as well as their respective enzyme expression level evaluated via qPCR were also studied (see **Figure 6** B, C and D). Fed nutrients including glutamine generated from alanine and aspartate metabolism acts as substrate to generate n-carbamoyl-l-aspartate (CBAP), which is further metabolized into CMP. Uridine-cytidine kinase 2 (*Uck2*) facilitates the conversion of CMP to cytidine which is further metabolized into β-alanine (β-Ala). Ultimately, β-Ala is converted to acetyl coenzyme A (AcCoA) towards energy production in the citric acid cycle via 4-aminobutyrate aminotransferase (*Abat*). Downstream metabolite cytidine can also be reversely synthesized to CMP through 5'-nucleotidase (*Nt5*). Thus, elevated enzymatic expression level of *Uck2*, *Nt5* and *Abat* can serve as strong indication of excess CMP accumulated as response to different factors imposed by the dose-depending nutrients spiking study.

The gene expression levels of *Uck2*, *Nt5* and *Abat* as expressed on Day 3, Day 12 and Day 14 as quantified via qPCR are shown in Figure 6. In general, the results showed a higher expression level towards the later date of the culture, suggesting an increasing accumulation pattern of CMP as cells continuously metabolizing fed substrates. Unsurprisingly, the gene expression level of *Uck2*, *Nt5* and *Abat* obtained from the CMP two-factors conditions were all higher when compared against the control, further agrees with the metabolite analysis results, and therefore successfully validates the simulated result obtained from the flux modeling analysis study.

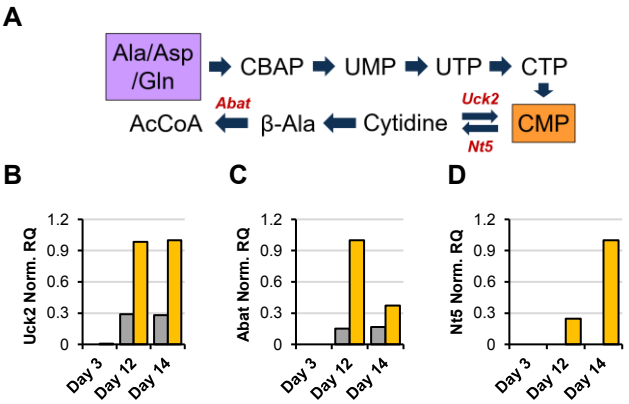


Figure 6. Gene expression analysis results from dose-dependent nutrients spiking study. (A) Schematic diagram of CMP-related metabolic pathway nested in the pyrimidine metabolism. The gene expression level of metabolic genes downstream to the formation of CMP were also evaluated via qPCR and presented as relative quantification (RQ) values. Shown here: (B) Uridine-cytidine kinase 2 (*Uck2*), (C) 4-Aminobutyrate aminotransferase (*Abat*), and (D) 5'-Nucleotidase (*Nt5*). Ala, alanine; CBAP, n-carbamoyl-l-aspartate; UMP, uridine monophosphate; CTP, cytidine triphosphate; CMP cytidine monophosphate; β-Ala, β-alanine; AcCoA, acetyl coenzyme A; ▫, amino acid nutrient inputs; ▫, metabolic inhibitor.

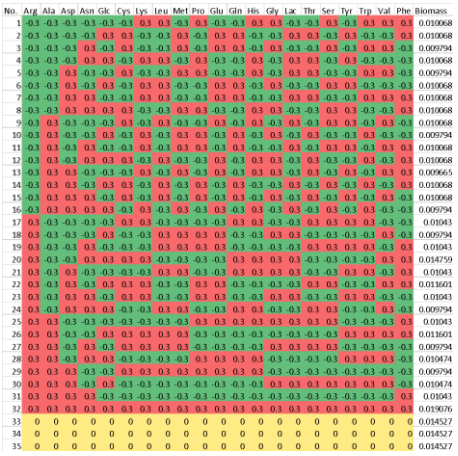
3 Medium components analysis via in silico process DOE

The low feed and late feed condition in **Figure 1** were selected to build the FBA and understand the process. The media compositions are optimized through *in silico* process DoE. For the screening design of amino acids at these two conditions, the fractional factorial designs were constructed with all the amino acids as the factors and the biomass growth rate as the response *in silico*. Two level of factors were 70% and 130%, meaning the variation of the raw input flux data. Three central points (100%) were also included in the fractional factorial design. In total thirty-five experimental runs were designed. With all the modified inputs, the biomass growth rates were estimated by FBA (see **Figure 7**). Each row represented the DOE condition of each run of experiment. The estimated growth rates were listed at the last column of the table.

However, it was not straightforward to observe the relationship between the estimated growth rate and the DOE conditions of amino acids directly from the table. Therefore, the Partial least squares regression (PLS) approach was introduced to abstract the relationship. The DOE condition data and the estimated growth rates generated from the model were fitted to a PLS model to abstract the relationship between the output data (growth rate) and the input data (DOE condition data). The variable importance in projection (VIP) plot was generated to understand how each amino acid was affecting the variation of biomass growth rate. The VIP value of each amino acid indicated its importance in controlling the growth rate at that condition (see **Figure 8**).

For the low feed and late feed condition, the most important amino acids components estimated were glucose, glutamine, lactate, asparagine, lysine, and serine, while that at high feed and early feed condition were valine, lysine, phenylalanine, cystine, arginine, and glucose. The difference in the top ranked amino acids and metabolites reflected the difference in cell metabolism at different conditions. When the low nutrient feed is applied at late stage of the cell culture, the metabolites that have a higher consumption (or accumulation rate) were at low level or depleted (or at high level), such as the consumption of glucose and glutamine (and accumulation of lactate). However, an early-stage feed with rich nutrient components in the media avoided such depletion or accumulation of metabolites. As the result, the variation in the media will not affect the cell growth as much as that in the former case. Instead, the valine plays an important role in the feed. The conclusion can be media ingredient and cell line specified. Valine is known as essential amino acids for cell culture and plays significant role in transport system in cell metabolism.

The top ranked amino acids and metabolites in VIP plots were selected for further analysis in two experiment conditions. Optimization on the composition of these selected amino acids for maximum biomass growth rate was estimated by optimization design with respond surface methodology approach. The optimization design was constructed with the factors being selected from the top ranked amino acids and metabolites on the VIP plot of the screening process. Three-level full factorial designs were established for two experiment conditions with the top six amino acids and metabolites mentioned in the previous paragraph as factors. In total, 732 runs of experiment designed including three central points are constructed for each condition. FBA was applied to GEM to perform *in silico* experiments based on the designed conditions to estimate the biomass growth rate. The relationship between the top six metabolites and the calculated biomass growth rate was studied by response surface methodology to find the optimum composition of the selected six metabolites for maximum biomass growth rate. The design conditions and the estimated biomass growth rate were fitted to a quadratic model using least regression. The maximum biomass growth rates were found based on the quadratic model. The range of amino acids inputs were found based on the model to achieve the maximum biomass growth rate.



No.	Arg	Ala	Asp	Asn	Glc	Cys	Lys	Leu	Met	Pro	Glu	Gln	His	Gly	Lac	Thr	Ser	Tyr	Trp	Val	Phe	Biomass
1	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	0.3	0.3	-0.3	0.3	-0.3	-0.3	0.3	0.3	-0.3	-0.3	0.3	-0.3	0.3	0.3	0.3	0.010068
2	-0.3	-0.3	-0.3	-0.3	0.3	0.3	-0.3	-0.3	0.3	-0.3	0.3	0.3	-0.3	-0.3	0.3	0.3	-0.3	0.3	-0.3	-0.3	0.3	0.010068
3	-0.3	-0.3	-0.3	0.3	-0.3	0.3	-0.3	-0.3	0.3	-0.3	0.3	0.3	-0.3	0.3	-0.3	-0.3	0.3	0.3	0.3	-0.3	-0.3	0.009794
4	-0.3	-0.3	-0.3	0.3	0.3	-0.3	0.3	0.3	-0.3	0.3	-0.3	-0.3	0.3	-0.3	0.3	0.3	-0.3	0.3	-0.3	-0.3	-0.3	0.010068
5	-0.3	-0.3	0.3	-0.3	-0.3	0.3	-0.3	-0.3	0.3	0.3	-0.3	0.3	0.3	0.3	-0.3	-0.3	0.3	-0.3	0.3	0.3	-0.3	0.009794
6	-0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	0.3	-0.3	-0.3	0.3	0.3	-0.3	0.3	-0.3	-0.3	0.3	-0.3	-0.3	-0.3	-0.3	0.010068
7	-0.3	-0.3	0.3	0.3	-0.3	-0.3	0.3	0.3	-0.3	-0.3	0.3	0.3	-0.3	-0.3	0.3	0.3	-0.3	-0.3	0.3	0.3	0.3	0.010068
8	-0.3	-0.3	0.3	0.3	0.3	0.3	-0.3	-0.3	0.3	0.3	-0.3	-0.3	0.3	0.3	-0.3	-0.3	0.3	0.3	-0.3	-0.3	0.3	0.010068
9	-0.3	0.3	-0.3	-0.3	-0.3	0.3	-0.3	-0.3	0.3	-0.3	0.3	0.3	-0.3	0.3	-0.3	-0.3	0.3	0.3	-0.3	-0.3	-0.3	0.010068
10	-0.3	0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	-0.3	-0.3	0.3	-0.3	-0.3	0.3	-0.3	0.009794
11	-0.3	0.3	-0.3	0.3	0.3	-0.3	0.3	-0.3	0.3	0.3	-0.3	0.3	-0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	0.010068
12	-0.3	0.3	-0.3	0.3	0.3	0.3	-0.3	-0.3	0.3	-0.3	0.3	0.3	-0.3	0.3	-0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	0.010068
13	-0.3	0.3	0.3	-0.3	-0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	0.3	-0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	0.009665
14	-0.3	0.3	0.3	-0.3	0.3	0.3	-0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	0.010068
15	-0.3	0.3	0.3	0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	-0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	0.010068
16	-0.3	0.3	0.3	0.3	0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	0.3	-0.3	0.3	-0.3	-0.3	0.3	-0.3	-0.3	-0.3	-0.3	0.009794
17	0.3	-0.3	-0.3	-0.3	-0.3	0.3	0.3	-0.3	-0.3	-0.3	-0.3	0.3	0.3	-0.3	-0.3	0.3	0.3	0.3	0.3	-0.3	-0.3	0.01043
18	0.3	-0.3	-0.3	-0.3	0.3	-0.3	-0.3	0.3	0.3	0.3	-0.3	0.3	-0.3	0.3	-0.3	-0.3	-0.3	0.3	-0.3	-0.3	-0.3	0.009794
19	0.3	-0.3	-0.3	0.3	-0.3	-0.3	0.3	0.3	-0.3	0.3	-0.3	-0.3	0.3	-0.3	-0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	0.01043
20	0.3	-0.3	-0.3	0.3	0.3	0.3	-0.3	-0.3	-0.3	-0.3	0.3	0.3	-0.3	0.3	-0.3	-0.3	0.3	-0.3	-0.3	-0.3	0.3	0.014759
21	0.3	-0.3	0.3	-0.3	-0.3	-0.3	0.3	0.3	-0.3	-0.3	0.3	0.3	-0.3	-0.3	0.3	0.3	-0.3	-0.3	0.3	-0.3	0.3	0.01043
22	0.3	-0.3	0.3	-0.3	0.3	0.3	-0.3	0.3	-0.3	0.3	0.3	-0.3	-0.3	-0.3	-0.3	0.3	0.3	-0.3	0.3	-0.3	0.3	0.011601
23	0.3	-0.3	0.3	0.3	-0.3	0.3	-0.3	-0.3	0.3	0.3	-0.3	0.3	-0.3	-0.3	0.3	0.3	-0.3	0.3	-0.3	-0.3	-0.3	0.01043
24	0.3	-0.3	0.3	0.3	0.3	-0.3	0.3	-0.3	0.3	-0.3	0.3	0.3	-0.3	0.3	-0.3	-0.3	0.3	-0.3	0.3	-0.3	-0.3	0.009794
25	0.3	0.3	-0.3	-0.3	-0.3	-0.3	-0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	-0.3	-0.3	-0.3	0.3	0.01043
26	0.3	0.3	-0.3	-0.3	0.3	0.3	0.3	0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	0.3	0.3	0.3	0.3	0.011601
27	0.3	0.3	-0.3	0.3	-0.3	0.3	0.3	0.3	-0.3	-0.3	0.3	0.3	-0.3	0.3	0.3	0.3	-0.3	0.3	-0.3	-0.3	-0.3	0.009794
28	0.3	0.3	-0.3	0.3	0.3	-0.3	-0.3	-0.3	0.3	0.3	0.3	0.3	-0.3	-0.3	-0.3	-0.3	-0.3	0.3	0.3	0.3	-0.3	0.010474
29	0.3	0.3	0.3	-0.3	-0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	0.009794
30	0.3	0.3	0.3	-0.3	0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	-0.3	0.010474
31	0.3	0.3	0.3	0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	0.3	0.01043
32	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.019076
33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.014527
34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.014527
35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.014527

Figure 7. Fractional factorial design with twenty-one factors at two levels (70% and 130%)..

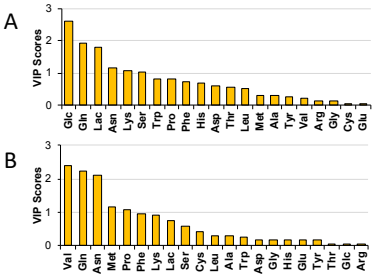


Figure 8. Variable importance in projection (VIP) plot of amino acids to biomass in fitted PLS model. (A) Condition 7: late feed and low feed. (B) Condition 3: early feed and high feed.

The maximum biomass growth rate estimated for condition 3 is higher than that in condition 7, which indicates the high feed and early feed indeed has potential for a higher biomass growth rate with proper optimization to the metabolism of cell culture. The optimized modification on metabolites in media is different for different experiment conditions as the cell metabolism and media composition are both different. The optimized values of metabolites being at around 0.3 is because of the optimization range set for the study. The maximum increase fraction of flux is set to 30% in this study since the design of experiments range is within 30%. The optimization design space is therefore within such boundary. The simulated optimization results suggested that the metabolites shall increase/reduce by the fraction specified in Table 1 to achieve the optimized combination of metabolites flux, which resulted the maximum biomass growth rate. Although the actual concentration of those metabolites may not be able to achieve the specific number indicated by the model due to the limitation in their solubility, the model provided a direction on optimizing the media at different feeding strategies. In this approach, we demonstrated that incorporating FBA and *in silico* DOE simulations is demonstrated can provide reveals insights concerning the study in optimizing medium components to achieve the maximum growth rate and productivity with simple experiments to be performed for the controlled conditions as raw input flux.

3.6 Medium components analysis via impact screening on biomass

As briefly mentioned, providing CHO culture with nutrient rich medium can have a direct negative impact on cellular performance, mostly due to diversion of excess nutrients available in extracellular environment into inefficient pathways which produces metabolic inhibitory by-products. In this study, the relationship between the supplemented concentration of the amino acid precursors and their impact on growth was further explored via in silico sensitivity analysis of medium components. Here, FBA simulation was performed based on previously obtained metabolomic flux data. Golden-section search is applied on cellular uptake flux of each amino acid to determine the lowest achievable down-regulated factor without compromising on the generation of cellular biomass, as shown in **Figure 9**. The final lower limit threshold at which amino acid flux can be reduced while still maintaining comparable level of growth rate to the control (no down-regulated factor applied) was recorded and pooled towards subsequent medium optimization study.

The results of the sensitivity analysis study are shown in **Figure 9**. The down-regulated factors of nine different amino acid precursors to the identified inhibitory metabolites – as previously shown in **Table 2** – were further considered for the following medium optimization study. Since the utilized CHO culture medium as used throughout this study was intensified with concentrated medium components to maximize resource use, the final down-regulated factors as studied under this approach were slightly adjusted to allow complete dissolution of supplementing amino acid without requiring the adjustment of medium pH which could affect osmolality and further introducing another variable in play.

3.7 Plackett-Burman design to optimize medium components

In the following study, ten different amino acids as previously identified (**Table 4**) were added to a pool of components targeted for medium optimization strategy. A Plackett-Burman DOE design was incorporated to optimize medium components prioritizing the support of cellular expansion rate in a CHO fed-batch process. The amino acid concentrations were simultaneously varied to either high level (+1, same concentration as found in the cultivated medium) or low level (-1, with a down-regulated factor applied) and supplemented to the seed medium on Day 0. Overall, the design presented a screening matrix with twelve unique conditions along with a control (condition 13, no down-regulated factor applied).

Critical process performance in terms of cellular phenotypic behavior such as VCD profile, IVCD profile and growth rate obtained from the conducted batch process were systematically evaluated to assess the impact of each variable across all tested conditions. In general, all conditions studied with at least four or more amino acid components subjected to decreased concentration exhibited improved cellular phenotypes. When comparing the best performing condition at each CPP against the control, the peak VCD profile of cells realized a 40% increase on Day 4 at 13.1×10^6 cells·mL^{-1} (condition 3) versus 8.8×10^6 cells·mL^{-1} (control). Similarly, the cumulative VCD profile also realized a 15% increase on Day 6 at 6.1×10^6 cells·mL^{-1} (condition 9) versus 36.1×10^6 cells·mL^{-1} (control). Interestingly, the effect of down-supplementing amino acids realized the best improvement in cellular growth rate profile calculated by a log-based growth model at 2.9-folds increase on Day 1 with the highest growing condition (condition 1) comparing against the control. Overall, improvement in cellular performance throughout the entire culture duration (Day 0 to Day 6) with respect to the target product CPP was realized across all studied conditions. The results from the study altogether confirmed significant fraction of fed nutrients was diverted into non-regulated metabolic pathways generating inhibitory by-products; and thus, by effectively limiting the amount of supplemented amino acid nutrients in the cultivated medium, the impact of these identified rate-limiting factors was effectively mitigated, allowing cells to achieve a higher growth rate and better overall peak and cumulative cell densities.

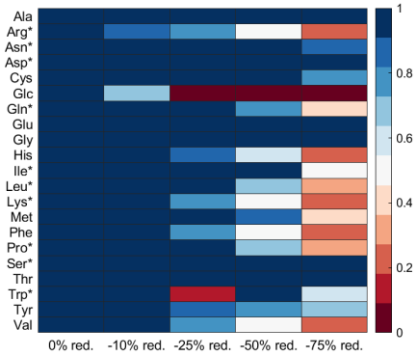


Figure 9. *In silico* impact simulation of reducing amino acid uptake fluxes on generation of cellular biomass. Shown here: *, amino acid precursor of identified inhibitory metabolites.

Table 4. Target identified amino acids and their corresponding down-regulated factors applied.

Asn	Asp	Gln	Ile	Leu	Pro	Trp	Ser	Glu	Met
0.44	0.25	0.72	0.49	0.69	0.67	0.39	0.25	0.25	0.59

4. CONCLUSIONS

We present an unorthodox approach to optimize CHO bioprocess by integrating conventional design-of-experiments methodology with genome scale model flux analysis. Generic CHO-K1 metabolic model was tailored and further integrated with CHO fed-batch metabolomic data to obtain a cell line- and process-specific model. In silico metabolic flux analysis was conducted via GEM to identify the critical medium components toward cellular growth and further evaluate their optimized flux values from thirty-five simulated fed-batch DOE conditions. Glucose and valine were projected as the most critical nutrients in the process from the flux simulation analysis. Previously identified metabolic inhibitor cytidine monophosphate accumulated in extracellular environment was found to be regulated by glucose, glutamine, aspartate, and alanine and further experimentally validated through dose-dependent amino acid spiking study. A process diagnostic and control model was constructed from network topology modeling constructed through GEM and pathway enrichment analysis, which allowed optimization of medium components utilized in a fed-batch feeding process to better support cell metabolism and mitigate accumulation of metabolic inhibitors.

5. REFERENCES

Aggarwal, S. 2011. What's fueling the biotech engine--2010 to 2011. *Nat Biotechnol,* 29**,** 1083-9 DOI: 10.1038/nbt.2060.

Åkesson, K., Pettersson, S., Ståhl, S., et al. 2018. Kynurenine pathway is altered in patients with SLE and associated with severe fatigue. *Lupus science & medicine,* 5**,** e000254.

Ashihara, H. 2008. Trigonelline (N-methylnicotinic acid) biosynthesis and its biological role in plants. *Natural Product Communications,* 3**,** 1934578X0800300906.

Bocca, C., Kane, M. S., Veyrat-Durebex, C., et al. 2018. The Metabolomic Bioenergetic Signature of Opa1-Disrupted Mouse Embryonic Fibroblasts Highlights Aspartate Deficiency. *Sci Rep,* 8**,** 11528 DOI: 10.1038/s41598-018-29972-9.

Boebel, K. P. & Baker, D. H. 1982. Comparative utilization of the alpha-keto and D- and L-alpha-hydroxy analogs of leucine, isoleucine and valine by chicks and rats. *J Nutr,* 112**,** 1929-39 DOI: 10.1093/jn/112.10.1929.

Bordbar, A., Monk, J. M., King, Z. A., et al. 2014. Constraint-based models predict metabolic and associated cellular functions. *Nat Rev Genet,* 15**,** 107-20 DOI: 10.1038/nrg3643.

Brosnan, M. E. & Brosnan, J. T. 2007. Orotic acid excretion and arginine metabolism. *J Nutr,* 137**,** 1656S-1661S DOI: 10.1093/jn/137.6.1656S.

Chen, Q., Kirk, K., Shurubor, Y. I., et al. 2018. Rewiring of glutamine metabolism is a bioenergetic adaptation of human cells with mitochondrial DNA mutations. *Cell metabolism,* 27**,** 1007-1025. e5.

Cruz, H. J., Freitas, C. M., Alves, P. M., et al. 2000. Effects of ammonia and lactate on growth, metabolism, and productivity of BHK cells. *Enzyme Microb Technol,* 27**,** 43-52 DOI: 10.1016/s0141-0229(00)00151-4.

Feist, A. M., Herrgard, M. J., Thiele, I., et al. 2009. Reconstruction of biochemical networks in microorganisms. *Nat Rev Microbiol,* 7**,** 129-43 DOI: 10.1038/nrmicro1949.

Gagne, F. 2014. *Biochemical ecotoxicology: principles and methods*, Elsevier.

Hefzi, H., Ang, K. S., Hanscho, M., et al. 2016. A Consensus Genome-scale Reconstruction of Chinese Hamster Ovary Cell Metabolism. *Cell Syst,* 3**,** 434-443 e8 DOI: 10.1016/j.cels.2016.10.020.

Hoang, D., Galbraith, S., Kuang, B., et al. 2021. Characterization of Chinese hamster ovary cell culture feed media precipitate. *Biotechnol Prog,* 37**,** e3188 DOI: 10.1002/btpr.3188.

Hu, X., Shen, J., Pu, X., et al. 2017. Urinary Time- or Dose-Dependent Metabolic Biomarkers of Aristolochic Acid-Induced Nephrotoxicity in Rats. *Toxicol Sci,* 156**,** 123-132 DOI: 10.1093/toxsci/kfw244.

Kim, S. H., Cho, S. K., Hyun, S. H., et al. 2011. Metabolic profiling and predicting the free radical scavenging activity of guava (Psidium guajava L.) leaves according to harvest time by 1H-nuclear magnetic resonance spectroscopy. *Biosci Biotechnol Biochem,* 75**,** 1090-7 DOI: 10.1271/bbb.100908.

Kuang, B., Dhara, V. G., Hoang, D., et al. 2021. Identification of novel inhibitory metabolites and impact verification on growth and protein synthesis in mammalian cells. *Metab Eng Commun,* 13**,** e00182 DOI: 10.1016/j.mec.2021.e00182.

Lane, A. N. & Fan, T. W. 2015. Regulation of mammalian nucleotide metabolism and biosynthesis. *Nucleic Acids Res,* 43**,** 2466-85 DOI: 10.1093/nar/gkv047.

Lao, M. S. & Toth, D. 1997. Effects of ammonium and lactate on growth and metabolism of a recombinant Chinese hamster ovary cell culture. *Biotechnol Prog,* 13**,** 688-91 DOI: 10.1021/bp9602360.

Mohmad-Saberi, S. E., Hashim, Y. Z., Mel, M., et al. 2013. Metabolomics profiling of extracellular metabolites in CHO-K1 cells cultured in different types of growth media. *Cytotechnology,* 65**,** 577-86 DOI: 10.1007/s10616-012-9508-4.

Mulukutla, B. C., Kale, J., Kalomeris, T., et al. 2017. Identification and control of novel growth inhibitors in fed-batch cultures of Chinese hamster ovary cells. *Biotechnol Bioeng,* 114**,** 1779-1790 DOI: 10.1002/bit.26313.

Nowaczyk, M. J., Lehotay, D. C., Platt, B. A., et al. 1998. Ethylmalonic and methylsuccinic aciduria in ethylmalonic encephalopathy arise from abnormal isoleucine metabolism. *Metabolism,* 47**,** 836-9 DOI: 10.1016/s0026-0495(98)90122-6.

Ojala, T., Wilson, J. M., Hulmi, J. J., et al. 2013. α-Hydroxy-Isocaproic Acid (HICA)—Effects on Body Composition, Muscle Soreness and Athletic Performance. *Nutrition and Enhanced Sports Performance.* Elsevier.

Perchat, N., Saaidi, P. L., Darii, E., et al. 2018. Elucidation of the trigonelline degradation pathway reveals previously undescribed enzymes and metabolites. *Proc Natl Acad Sci U S A,* 115**,** E4358-E4367 DOI: 10.1073/pnas.1722368115.

Saas, J., Ziegelbauer, K., Von Haeseler, A., et al. 2000. A developmentally regulated aconitase related to iron-regulatory protein-1 is localized in the cytoplasm and in the mitochondrion of Trypanosoma brucei. *Journal of Biological Chemistry,* 275**,** 2745-2755.

Selvarasu, S., Ho, Y. S., Chong, W. P., et al. 2012. Combined in silico modeling and metabolomics analysis to characterize fed-batch CHO cell culture. *Biotechnol Bioeng,* 109**,** 1415-29 DOI: 10.1002/bit.24445.

Wu, G., Bazer, F. W., Datta, S., et al. 2008. Proline metabolism in the conceptus: implications for fetal growth and development. *Amino Acids,* 35**,** 691-702 DOI: 10.1007/s00726-008-0052-7.

Yusufi, F. N. K., Lakshmanan, M., Ho, Y. S., et al. 2017. Mammalian Systems Biotechnology Reveals Global Cellular Adaptations in a Recombinant CHO Cell Line. *Cell Syst,* 4**,** 530-542 e6 DOI: 10.1016/j.cels.2017.04.009.