Biotechnology Journal

WILEY-VCH

Biotechnology Journal

RESEARCH ARTICLE OPEN ACCESS

Time-Resolved Hierarchical Modeling Highlights Metabolites Influencing Productivity and Cell Death in Chinese Hamster Ovary Cells

Andreas Eriksson[1] | Anne Richelle[2] | Johan Trygg[1,3] | Steffi Scholze[4] | Shanti Pijeaud[5] | Henrik Antti[1] | Christoph Zehe[5] | Izabella Surowiec[6] | Pär Jonsson[1,3]

[1]Department of Chemistry, Umeå University, Umeå, Sweden | [2]Sartorius Corporate Research, Brussels, Belgium | [3]Sartorius Corporate Research, Umeå, Sweden | [4]Sartorius Stedim Biotech GmbH, Göttingen, Germany | [5]Sartorius Corporate Research, Ulm, Germany | [6]Sartorius Stedim Data Analytics AB, Umeå, Sweden

Correspondence: Pär Jonsson (paer.jonsson@sartorius.com)

Received: 16 October 2024 | **Revised:** 19 February 2025 | **Accepted:** 20 February 2025

Keywords: bioprocess data | Chinese hamster ovary (CHO) cells | death rate | hierarchical modeling | metabolomics | orthogonal partial least squares (OPLS) | productivity

ABSTRACT

Biopharmaceuticals are medical compounds derived from biological sources and are often manufactured by living cells, primarily Chinese hamster ovary (CHO) cells. CHO cells display variation among cell clones, leading to growth and productivity differences that influence the product's quantity and quality. The biological and environmental factors behind these differences are not fully understood. To identify metabolites with a consistent relationship to productivity or cell death over time, we analyzed the extracellular metabolome of 11 CHO clones with different growth and productivity characteristics over 14 days. However, in bioreactor processes, metabolic profiles and process variables are both strongly time-dependent, confounding the metabolite-process variable relationship. To address this, we customized an existing hierarchical approach for handling time dependency to highlight metabolites with a consistent correlation to a process variable over a selected timeframe. We benchmarked this new method against conventional orthogonal partial least squares (OPLS) models. Our hierarchical method highlighted several metabolites consistently related to productivity or cell death that the conventional method missed. These metabolites were biologically relevant; most were known already, but some that had not been reported in CHO literature before, such as 3-methoxytyrosine and succinyladenosine, had ties to cell death in studies with other cell types. The metabolites showed an inverse relationship with the response variables: those positively correlated with productivity were typically negatively correlated with the death rate, or vice versa. For both productivity and cell death, the citrate cycle and adjacent pathways (pyruvate, glyoxylate, pantothenate) were among the most important. In summary, we have proposed a new method to analyze time-dependent omics data in bioprocess production. This approach allowed us to identify metabolites tied to cell death and productivity that were not detected with traditional models.

Abbreviations: μ_d, specific death rate; μ_g, specific growth rate; ABC, ATP-binding cassette; CHO, Chinese hamster ovary; CoA, coenzyme A; DA, discriminant analysis; EP, effect projection; GC, gas chromatography; GSH, reduced glutathione; GSSG, oxidized glutathione; LC, liquid chromatography; mAB, monoclonal antibody; MS, mass spectrometry; MVDA, multivariate data analysis; NMR, nuclear magnetic resonance; OPLS, orthogonal partial least squares; p(corr), correlation-scaled loadings; PCA, principal component analysis; PCC, Pearson correlation coefficient; Q^2, coefficient of prediction; qP, specific productivity; TCA, tricarboxylic acid; UV, unit variance; VCD, viable cell density.

1 | Introduction

Biopharmaceuticals, or biologics, are pharmaceuticals derived from biological sources. Biopharmaceuticals include vaccines, monoclonal antibodies (mAbs), and other recombinant proteins. Mammalian cells are the leading expression systems in the biologics industry [1], comprising 2/3 of newly approved drugs from 2018 to 2022, of which Chinese hamster ovary (CHO) cells accounted for 89% [2]. CHO cells are good production systems, reaching high densities in chemically defined media and suspension cultures [3]. CHO cells can also make human-compatible post-translational modifications like protein folding and glycosylation [1] and are less susceptible to human viruses than other cells [3].

The rising demand for biologics necessitates higher production. Currently, protein titers around 10 g/L are possible in prevailing fed-batch cultures [4]. However, continuous cultures are emerging due to advances in upstream bioprocessing. Continuous cultures cycle waste with perfusion feeding systems, improving the idle reactor time and viable cell density (VCD), in turn enhancing the volumetric productivity (g/L/day). Productivity can also be improved on a cell-specific (qP) level via cell line engineering or optimized culture conditions [5]. Twenty million cells/mL and 50 pg/cell/day [6] have been achieved in fed-batch cultures. However, VCDs cannot increase indefinitely, and higher VCDs may not lead to higher titers if qP is adversely affected [4]. Thus, optimizing not only VCD but also qP is crucial to increasing volumetric productivity further.

Cell proliferation is governed by two complementary processes: cell division, measured by the specific growth rate (μ_g, 1/day), and cell death, measured by the specific death rate (μ_d, 1/day) [7]. These rates inform how VCD evolves: increasing in the exponential phase ($\mu_g > \mu_d$), peaking in the stationary phase ($\mu_g \approx \mu_d$), and declining in the death phase ($\mu_g < \mu_d$) [8]. Most CHO studies target the exponential and stationary phases. Fewer examine the death phase [9], indicating a need to study CHO metabolism late in production further. Cell death is detrimental to product quality and titer. Dying cells release enzymes that degrade the recombinant proteins and alter the glycan structures. Cellular debris also creates impurities, complicating downstream bioprocessing [10]. μ_d increases over time, accelerating from the stationary phase, likely due to toxic byproducts, not nutrient exhaustion [4]. Removing waste metabolites via perfusion could improve viability and productivity. Lactate and ammonia, the primary waste metabolites from glucose and amino acids, respectively, can be limited with optimized feeding strategies [9], but other toxic metabolites must also be considered to enhance culture productivity [3, 9, 11].

Multidimensional omics techniques aid in investigating biological systems. Metabolomics, in particular, links snapshots of cellular physiology to important phenotypes like high productivity or low cell death [3, 12]. Chemometric techniques help interpret metabolomics data. Principal component analysis (PCA) identifies overall trends and outliers, while orthogonal partial least squares (OPLS) relates metabolites to process variables, or distinguishes between sample classes [13]. These techniques have been applied to CHO metabolic growth phases [14], cell densities [15], and media compositions [16], identifying targets for cell line and medium development, such as metabolites that stifle growth or induce apoptosis [4, 17, 18]. Even under identical growth conditions, CHO cells show clonal variation, leading to growth and productivity differences between production batches that affect the titer and critical quality attributes [12]. However, the underlying biology is underexplored.

The obstacle here is time. Metabolomics samples from the same batch are not independent over time, creating artificially high metabolite-process variable correlations. PCA struggles when the greatest variation is time-related [19]. Various (O)PLS-based techniques have been used with time series, each with its strengths and drawbacks: PLS on batch-level models can provide metabolic correlations per time point [20], but monotonic or cumulative response variables are necessarily time-dependent [21]. OPLS-discriminant analysis (DA) can compare time points pair-wise [14], but the differences might not relate to a process variable of interest. Multiway PLS-DA [22] can compare groups of cell lines for a single time point or overall, but may not capture changes in the response variable over time.

To overcome the time dependence, in this study, we devised a hierarchical method based on OPLS regression to analyze metabolites at each time point separately. We compared our method against conventional, global OPLS models that consider the metabolites at all time points simultaneously. Seeking metabolites linked to productivity and cell death, we analyzed the extracellular metabolome of CHO cells with varying growth and productivity properties. The new method pinpointed several metabolites with consistent relations to qP and/or μ_d in the stationary and death phases that the global method missed. Many metabolites were biologically relevant to CHO cell death and/or productivity, demonstrating the new method's effectiveness in analyzing time-dependent bioprocess omics data.

2 | Materials and Methods

2.1 | Cell Line and Culture Conditions

We grew eleven clones of the same CHO-DG44 parental cell line (Sartorius, Germany) as fed-batch cultures in 0.25 L automated parallel Ambr 250 High Throughput bioreactors (Sartorius, Germany) as previously described [23]. Each clone had distinct growth and productivity properties, but all expressed the same IgG1 mAb. The reactors ran for 14 days (one was halted on day 12 after contamination). Samples were automatically drawn daily by an Ambr 250 liquid handler (Sartorius, Germany).

2.2 | Extracellular Metabolomics Analysis

We conducted untargeted metabolomics analysis of extracellular metabolites using tandem liquid and gas chromatography-mass spectrometry (LC-MS/MS, GC-MS), and nuclear magnetic resonance (NMR) spectroscopy. The Supporting Information details sample preparation, quantitation, and data pre-processing. We removed metabolites with relative standard deviations over 30% in the quality control samples, signal-to-noise ratios below 3, and duplicates across LC, GC, and/or NMR. In all, 109 unique metabolites were annotated.

2.3 | Bioprocess Variables

Cell culture variables (Figure S1) were measured with a BioProfile FLEX2 (Nova Biomedical, USA), as was the titer using an Octet R8 system (Sartorius, Germany). After removing outliers, we calculated two process variables): the specific death rate (μ_d, 1/day) and the specific productivity (qP, pg/cell/day). For qP, the titer data were smoothed using a generic logistic function, and its derivative was normalized to the VCD. μ_d and the specific growth rate (μ_g, 1/day) were calculated as described previously [24]. We assigned cell growth phases as follows: exponential: $\mu_g > \mu_d$ ($\approx$ days 1–6); stationary: $\mu_g \approx \mu_d$ ($\approx$ days 7–8); death: $\mu_g < \mu_d$, ($\approx$ days 9–14).

2.4 | Multivariate Data Analysis (MVDA)

We performed MVDA using SIMCA, version 18.0.0 (Sartorius, Sweden). All data were mean-centered and scaled to unit variance (UV) unless noted otherwise. PCA provided a data overview and outlier detection. OPLS regression linked the relative metabolite levels (X) to μ_d or qP as the response (y) variable. SIMCA automatically estimated the number of model components from the coefficient of prediction (Q^2) using cross-validation. Models exceeding two components were manually reduced to 1 predictive + 1 orthogonal to avoid the risk of overfitting [25].

To find the timeframe where the metabolite concentrations correlated with a response (μ_d or qP), we created one model per day and response, yielding 14 μ_d and 14 qP OPLS models. Similar to other studies [26], we used $Q^2 > 0.25$ as the threshold for acceptable models, retaining only contiguous days for cohesion, meaning days 8–14 for qP and 7–14 for μ_d, covering the stationary and death phases. We assessed the metabolites' relationship to the bioprocess variables with two methods. In the global method, we created a single OPLS model for each response (μ_d or qP) using samples from the retained days. In the hierarchical method, we used the 8 μ_d- and 7 qP-retained OPLS models and calculated the metabolites' correlation-scaled loadings (p(corr)) [21]. The p(corr) values were then used as input in OPLS effect projection (EP) [27] models without mean-centering. OPLS-EP used p(corr) as X data against a $y = 1$ column vector to identify common trends between time points.

We compared which metabolites were highlighted as significant in the global and hierarchical methods. Significance was estimated using p(corr) values from the two global OPLS models and the two OPLS-EP models. p(corr) is functionally equivalent to the Pearson correlation coefficient (PCC or Mr) [28] with significance based on Student's two-tailed t-distribution:

$$t = r\sqrt{\frac{n-2}{1-r^2}} \tag{1}$$

where t is the t-statistic and n is the number of independent samples. The global method set the number of batches as n since samples from different days were dependent. If n were the number of samples, nearly all metabolites would be significant. The hierarchical method set the number of days as n since all batches were condensed to one value [21]. The global method's p(corr) describes the metabolite-response correlation, while the hierarchical method's p(corr) describes consistently sized correlations between time points. To avoid falsely flagging consistent but negligibly correlated metabolites as significant, we set the mean p(corr) threshold to 0.30 [29] (Figure S2). The Supporting Information contains pathway overrepresentation analysis of the significant metabolites.

3 | Results and Discussion

3.1 | Benchmarking Hierarchical Modeling with OPLS-EP Against Global OPLS Modeling

The method we present here rests on a hierarchical approach to model bioprocess data by Alinaghi et al. [21]. Their approach captures the relationship between metabolites and response variables with separate OPLS regression models for each time point, then utilizes the metabolites' p(corr) values, which summarize each batch into one value per time point, visualizing them in a PCA model [21]. However, PCA models show overall differences, such as which metabolic correlations change over time. In this study, we focused on similarities instead: metabolites with consistent, strong relationships to a response at each time point. Thus, we replaced the PCA model with an OPLS-EP model (Figure 1). OPLS-EP [27] was originally developed for dependent samples where within-group variation can exceed between-group variation. OPLS-EP captures the difference X_{after}—X_{before} as X_{effect} against a $y = 1$ column vector, highlighting variables that change in the same direction for most observations [27]. Here, we used p(corr) as a stand-in for X_{effect}, revealing metabolites that correlated consistently with a response variable across most days without the time dependence.

We benchmarked integrating hierarchical modeling with OPLS-EP against conventional, global OPLS with all time points simultaneously. We created one model per method (global/hierarchical) and response (qP/μ_d), covering the stationary and death phases. We assessed the metabolite-response correlations' significance using shared and unique structures (SUS)-plots [25]. Table S1 lists all correlations and trajectories over time.

For both methods (Figure 2A,B), most metabolites were associated negatively with qP and positively with μ_d, whether we considered all metabolites or only significant ones. Most metabolites also had the same correlation sign ($\pm$) in both methods for a given response. However, the hierarchical method generally showed higher p(corr) values (closer to $\pm$ 1) and more significant metabolites, meaning more metabolites to investigate further compared with the global method. Interestingly, the responses were inversely related: metabolites correlated with μ_d were generally anti-correlated with qP, and vice versa. As living cells synthesize mAbs, factors benefiting productivity should not lead to cell death simultaneously. A tradeoff between cell growth and productivity has been reported before [4, 30]. This inverse relationship was clearer in the global models (Figure 2C) as both μ_d and qP co-varied with time. Without time as a confounder, the metabolites were more spread out in the hierarchical models (Figure 2D). Thus, the hierarchical approach more accurately reflected the metabolite-process variable relationship.

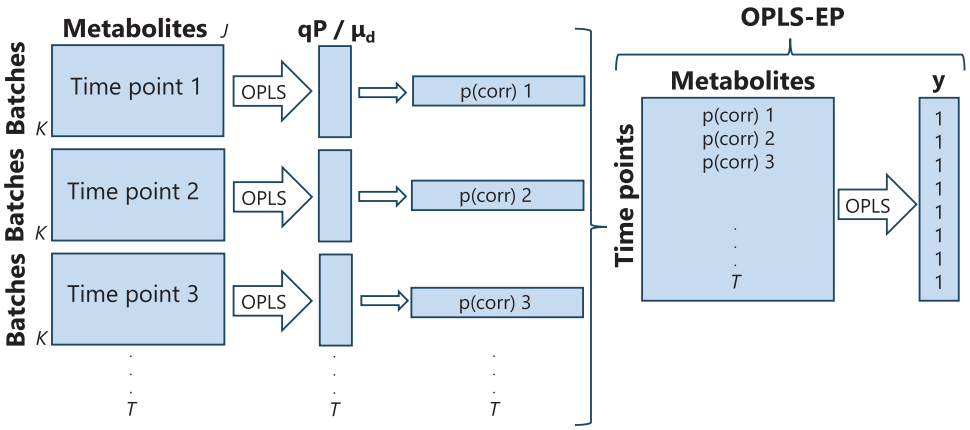

FIGURE 1 | Schematic of hierarchical modeling with OPLS-EP. To separate time- and batch-related changes, each time point is analyzed separately by regressing the metabolite levels (X) against qP or μ_d (Y) using OPLS. p(corr) loadings denote the correlation between X and the predicted y, all batches summarized into one value per time point. Stacking p(corr) vectors reveals the metabolites' relation to qP or μ_d at each time point, rather than overall. Lastly, OPLS-EP captures differences in X against a column vector of $y = 1$, pinpointing metabolites consistently related to qP or μ_d at most time points. Adapted from Alinaghi et al. [21].

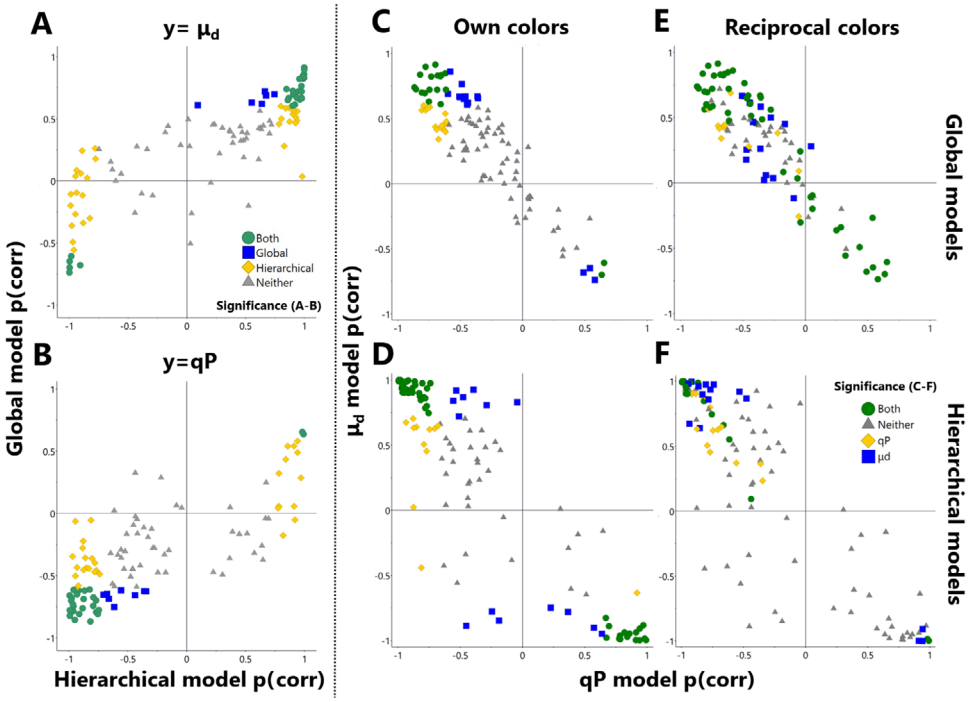

FIGURE 2 | SUS-plots comparing the metabolites' correlations (p(corr)) between models and responses. (A), (B) Global (Y-axis) versus hierarchical (x-axis) models with $y = \mu_d$ (A) or qP (B), coded by significance ($\alpha = 0.05$): both models (green circles), the global model (blue squares), the hierarchical model (yellow diamonds), or neither (gray triangles). (C)–(F) μ_d (Y-axis) versus qP (X-axis) correlations for global (C), (E) or hierarchical (D), (F) models, colored by that model's significance (C), (D) by the other model's reciprocally (E), (F): both responses (green circles); μ_d (blue squares); qP (yellow diamonds); or neither (gray triangles). This visualizes one method's significant metabolites according to the other: the global method's significant metabolites (C) in the hierarchical SUS-plot (F), and the hierarchical method's significant metabolites (D) in the global SUS-plot (E).

In total, 109 metabolites were detected. For 66, the outcome was consistent across methods: 26 correlated significantly with the same response(s) in both methods, while 40 did not correlate significantly with any response. The method affected the remaining 43 metabolites, correlating significantly with qP and/or μ_d only in one method. The hierarchical method had more unique metabolites: 30 versus 8 for qP; and 32 versus 6 for μ_d, as shown by the hierarchical method's significant metabolites (Figure 2D) appearing both near the corners (significant) and the center (non-significant) of the global SUS-plot (Figure 2E). Conversely, the global method's significant metabolites (Figure 2C) mostly appear in the hierarchical SUS-plot's corners (Figure 2F). Most unique metabolites correlated moderately with the response(s) ($r \approx \pm 0.5$) in the other method, though some were biologically negligible

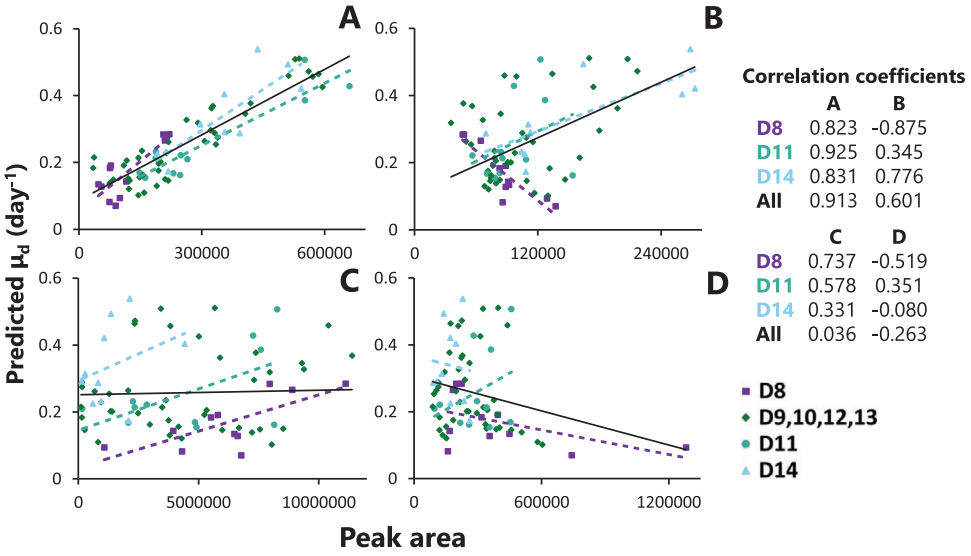

FIGURE 3 | Predicted μ_d versus peak area for four sample metabolites. Day (D) 8 (purple squares), 11 (teal circles), or 14 (blue triangles) in isolation, and all combined (solid line) for a metabolite highlighted as significant by (A) both methods (guanosine monophosphate), (B) the global method exclusively (N-acetylserine), (C) the hierarchical method exclusively (GSSG), or (D) neither method (γ-glutamylalanine).

($r < \pm 0.3$). The unique metabolites were interesting as they differentiated the methods, highlighting separate metabolites depending on their relationship with the response variables over time.

We illustrate the methods' differences using four sample metabolites representing different categories: significant in both methods, only in the global, only in the hierarchical, or in neither.

I. Guanosine monophosphate (Figure 3A) was highlighted as significantly correlated across all days and on each day separately.

II. N-acetylserine (Figure 3B) correlated positively across all days. However, by splitting the data by time point, the correlation switched signs from negative (day 8) to positive (day 11, 14) Thus, while significant in the global approach, N-acetylserine was nonsignificant in the hierarchical method due to day-to-day inconsistency.

III. Oxidized glutathione (GSSG) (Figure 3C) was nonsignificant for μ_d across all days, but splitting the data by day revealed a consistent correlation over time: moderate (day 8, 11) but decreasing (day 14). The consistent day-to-day sign made GSSG significant for μ_d in the hierarchical method.

IV. γ-glutamylalanine (Figure 3D) was not significant, being weakly correlated both overall and on each day.

Each category (Figure 3A–D) contained multiple metabolites, so the methods targeted different metabolites rather than just having sharper cutoffs. The global method strictly pinpointed metabolites with overall correlations, which the hierarchical method did not do if the sign ($\pm$) was inconsistent over time. Conversely, hierarchical modeling pinpointed consistent metabolites, even if the size varied so the overall correlation was weak. Thus, the hierarchical method found metabolites whose consistent response relationships were obscured by time in the global models.

3.2 | Metabolites Significantly Correlated with qP and μ_d Are Enriched in Similar Pathways

After comparing the methods, we explored whether certain sets of the hierarchical models' significantly correlated metabolites were enriched in specific pathways relative to all known compounds in CHO-related pathways for days 7–14 (μ_d, Figure S3A) and 8–14 (qP, Figure S3B). Table S1 lists pathways per metabolite. Because most metabolites correlated with both responses, their enrichment profiles overlapped, suggesting that many pathways are related to both responses:

- Both responses: ATP-binding cassette (ABC) transporters; the tricarboxylic acid (TCA) cycle; pantothenate and coenzyme A (CoA) biosynthesis; beta-alanine metabolism; alanine, aspartate, and glutamate metabolism; glyoxylate and dicarboxylate metabolism; glycine, serine, and threonine metabolism; and glutathione metabolism.
- Unique to μ_d: ferroptosis (cell death from accumulated iron and lipid peroxides).
- Unique to qP: aminoacyl-tRNA biosynthesis, glycerophospholipid metabolism, D-amino acid metabolism, and pyruvate metabolism.

In previous studies, the TCA cycle was enriched for qP [4]. Pantothenate and CoA metabolism was notable in the stationary phase, as was pyruvate, glyoxylate, and dicarboxylate metabolism in the death phase [14]. Glycerophospholipid metabolism has been linked to growth limitation [18]. Glutathione metabolism is tied to the oxidative stress response [31] and to productivity [32]. While ABC transporters and aminoacyl-tRNA biosynthesis were

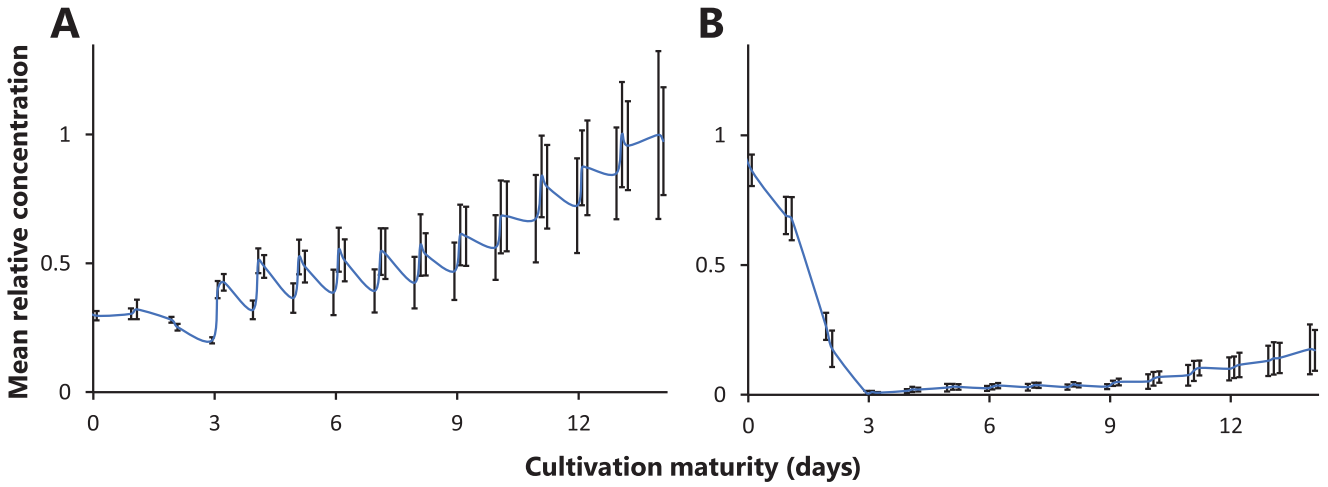

FIGURE 4 | Mean relative concentrations of amino acids over time. (A) Phenylalanine and (B) glutamine. Three samples were drawn daily: just before feeding (node), immediately after feeding (peak), and slightly later (post-peak) ($n = 11$ bioreactors, black error bars = 1 standard deviation).

significant, they are not specific and could reflect any individual amino acid.

We explored the response-amino acid correlations further. Many amino acids were present in the feed and/or medium (Table S1), affecting their levels over time and limiting how they reflected CHO metabolic changes. Most amino acids were detected by NMR, which had three daily samples, but we only modeled with one to work with the LC/GC-MS data. The NMR data showed that each bolus raised the amino acid levels (Figure 4A) more than the cells consumed between feeds, for a net increase over time, giving spurious correlations. Only glutamine (Figure 4B) did not display this behavior, potentially due to its early depletion [33]. While beyond this study's scope, studying production and consumption rates instead of concentrations might better characterize how amino acids impact productivity and cell death. For now, we suggest users monitor feed metabolites online or draw multiple samples per feed interval in fed-batch and continuous cultures to ensure that observed concentration changes stem from cellular metabolism, not feeding.

3.3 | Metabolites Unique to Hierarchical Modelling Are Biologically Relevant

The pathway analysis included all significant metabolites from the hierarchical method, but some were also significant in the global method. Excluding amino acids, 28 metabolites were unique to the hierarchical method (Table 1). Next, we examined their biological relevance to the responses.

Carbohydrate metabolism is crucial for cell growth and productivity. CHO cells preferably convert glucose to lactate via aerobic glycolysis, even with ample oxygen. Lactate reduces cell growth and productivity and can induce apoptosis [33]. We found that lactate anti-correlated significantly with qP. If glucose is depleted, CHO cells consume lactate instead [33], converting lactate to pyruvate and acetyl-CoA. Acetyl-CoA can be made into acetate, which we observed was significantly correlated with μ_d and anti-correlated with qP, accruing extracellularly in the stationary phase. Still, acetyl-CoA mainly enters the TCA cycle and oxidative phosphorylation, which are highly active in productive CHO cells [33]. TCA intermediates can escape into the cytosol [34], so replenishing them upholds TCA cycle activity.

We detected several TCA intermediates. Cis- and trans-aconitate and citrate/isocitrate (indistinguishable) were significantly correlated with qP and anti-correlated with μ_d. In one study, supplementation with succinate, malate, and α-ketoglutarate improved qP by up to 35% without viability loss [35]. Succinate addition yielded positive growth into the stationary phase and the highest qP among all tested intermediates [34]. Citrate, cis-aconitate, fumarate, malate, and succinate correlated with qP and anti-correlated with VCD in the stationary phase. Adding citrate in the exponential phase reduced peak VCD but increased qP [4]. Citrate in the feed decreased the VCD earlier than in controls, possibly due to apoptosis [34].

TCA intermediates are replenished using amino acids from the medium and cellular biosynthesis [36], first with glutamine, then other amino acids [33]. Excessive or catabolized amino acids can create growth-inhibiting derivatives that may accumulate extracellularly [3]. Ammonia, the primary byproduct, negatively affects productivity and cell growth [37]. Other derivatives include 4-hydroxyphenyllactate, a growth inhibitor derived from phenylalanine and tryptophan [38], which we found was significantly correlated with μ_d and anti-correlated with qP. N-formylmethionine, a cysteine and methionine product, accumulates in the exponential and stationary phases [18]. Here, N-formylmethionine declined in the death phase, significantly correlating with qP and anti-correlating with μ_d. Formate, a product of serine, glycine, and threonine, negatively impacts growth [38]. In our study, formate was significantly correlated with qP and anti-correlated with μ_d.

Fatty acid production keeps pace with cellular demand during growth but remains active even during nongrowth, causing build-up over time [18]. Glycerol-3-phosphocholine and phosphocholine decrease inside cells and accumulate outside in the exponential phase, correlating with apoptosis [17, 37]. In our study, glycerol-3-phosphocholine was significantly correlated with μ_d. Glycerol-3-phosphocholine and phosphocholine

TABLE 1 | Metabolites unique to hierarchical modeling.

Name	Trajectory	H	G	H	G	Ties to cell death, productivity, and/or extracellular accumulation	Reference
2-Hydroxy-4-(methylthio)butyrate		−		+			
2-Hydroxyglutarate/3-Hydroxyglutarate		−		+			
3-Methoxytyrosine		−					
3'-O-Methylguanosine		+		−		Correlated with apoptosis.	[31]
4-Hydroxyphenyllactate		+		−		Accumulation negatively impacts growth.	[38]
7-Methylguanine		+		−		Correlated with apoptosis.	[31]
Acetate		+	+	−		Accumulates extracellularly from the stationary phase.	[37]
Cholesterol		+		−	−	High synthesis improves productivity by enhancing secretion capacity.	[40]
cis-Aconitate		−	−	+		Correlated with qP, anti-correlated with VCD.	[4]
Citrate/Isocitrate		−	−	+		Correlated with qP, anti-correlated with VCD.	[4]
Formate		−	−	+		Accumulation negatively impacts growth.	[38]
gamma-Glutamylleucine				−			
gamma-Glutamyltryptophan		+		−	−		
Glutathione		−		+		Detoxifies reactive oxygen species by being oxidized. Correlated with productivity.	[31, 32]
Glycero-3-phosphocholine		+	+	−		Accumulates extracellularly over time. Correlated with apoptosis.	[17, 37]
Glycolate			+	−		Not known to be toxic; CHO cells lack the enzyme to convert to glyoxylate.	[44]
Inosine monophosphate		+	+	−			
Isoleucyl-Leucine/Isoleucyl-Isoleucine		−					
Lactate				−		Negatively impacts growth and productivity. Can induce apoptosis.	[33]
N2,N2-Dimethylguanosine		+	+	−		Correlated with apoptosis.	[31]
N-Dimethylglycine		−		+			
N-Formyl-methionine		−		+		Accumulates extracellularly over time.	[18]
Oxidized glutathione (GSSG)		+		−		Correlates with caspase activity. Can induce apoptosis in fresh cultures.	[17]
Phosphocholine				−		Accumulates extracellularly over time. Correlated with apoptosis.	[17, 37]
Spermidine		+		−	−	Depletion leads to growth arrest and eventually cell death.	[45]
Succinyladenosine		+	+	−			
Threonate		−					
trans-Aconitate		−		+			

Note: Metabolites with significant positive (+) or negative (−) correlations to μ_d and/or qP in the hierarchical model (H), but not to the same response(s) in the global model (G). Feed and medium metabolites were excluded. Concentration trajectories for days 1–14 and known ties to μ_d, qP, or extracellular accumulation in CHO literature are listed.

were anti-correlated with qP and accumulated in the death phase. Phospholipid build-up may indicate poor control of lipid metabolism and membrane composition [3]. Conversely, intracellular plasma membrane precursors decrease over time [18], so lipid metabolism and cell growth limitations are tightly linked.

Energy production is vital in mAb synthesis as cells need ATP to form peptide bonds [39]. High mAb producers have elevated levels of electron carriers in oxidative phosphorylation [31]. However, this process creates reactive oxygen species, which can be detoxified by reduced glutathione (GSH) through GSH's oxidation to GSSG. GSH synthesis is upregulated in productive CHO cells [32], while GSSG correlates with caspase activity and can induce apoptosis in fresh cultures [31]. Curiously, GSH depletion leads to reduced cholesterol synthesis. Since high cholesterol synthesis enhances productivity by improving secretion capacity, GSH's productivity effects may be tied to cholesterol [40]. In our study, cholesterol was significantly correlated with μ_d and anti-correlated with qP.

To our knowledge, several metabolites exclusively significant in the hierarchical models have not been linked to productivity or cell death in CHO cells before, making them interesting for future research. Some metabolites have ties in other cell lines, though. For example, 3-methoxytyrosine was significantly correlated with qP, anti-correlated with μ_d, and accumulated in the death phase. In humans, pathogenic 3-methoxytyrosine levels can lead to oxidative stress and possibly cell death [41]. γ-glutamyltryptophan inhibits cell growth in tumor cell lines [42]. In our study, γ-glutamyltryptophan significantly correlated with qP and anti-correlated with μ_d in the death phase. Succinyladenosine was significantly correlated with μ_d, anti-correlated with qP, and accumulated over time in our study. In adipocytes, succinyladenosine was significantly enriched after apoptosis [43].

In summary, most metabolites highlighted only by the hierarchical method were biologically relevant to productivity and/or cell death in CHO cells. Some correlated metabolites have, to our knowledge, not been reported in CHO cells before. While several metabolites were already known from other studies, they often concern the exponential or stationary phases, and either productivity or cell death. We report correlations for the death phase and both responses for most metabolites. Although we have metabolite levels for the entire batch (14 days), we only obtained functional class models in the stationary and death phases. This might stem from insufficient time for the cells to impact the extracellular medium in the exponential phase [21]. Alternatively, μ_d and qP may not have been relevant initially; for instance, μ_d remained low throughout the exponential phase.

4 | Conclusions

Bioprocess time series data are difficult to analyze. Both metabolic profiles and process variables are time-dependent, confounding how the metabolites relate to process variables. Traditional correlation-based statistical methods cannot resolve this time dependence. Here, we adopted a novel hierarchical method to model time-dependent data at each time point separately based on p(corr), combining the results in an OPLS-EP model. We applied this method to the extracellular metabolome of 11 CHO clones with different growth and productivity properties over 14 days, seeking metabolites with similar correlations with μ_d and/or qP between time points.

The hierarchical method pinpointed many metabolites consistently related to μ_d and/or qP that conventional, global OPLS modeling missed. The new approach also avoided flagging metabolites with inconsistent relations as significant, unlike the global method. μ_d and qP were inversely related: most metabolites were positively associated with one and negatively with the other, reflecting the relation between cell viability and mAb production. The key pathways concerned carbohydrate metabolism (TCA cycle, pyruvate, glyoxylate, pantothenate). Amino acid metabolism was highlighted as important too, but the correlations were spurious, mainly caused by amino acids' being in the feed, necessitating caution when analyzing changes in amino acid levels in fed-batch cultures.

Our study focused on the extracellular metabolome, which admittedly has limited relevance to intracellular processes. Likewise, intracellular metabolomics cannot adequately describe extracellular metabolic dynamics. Most studies explore either intra- or extracellular metabolomics, so integrating both sides could provide a more holistic understanding of cellular metabolism during bioprocess production. Importantly, the metabolites in our study that only the hierarchical method highlighted were biologically relevant. Some were described previously as relevant for cell death or productivity, but others have not been reported in CHO literature before, showing the new method's benefits. We hope our findings contribute to a better understanding of biopharmaceutical production and to improving bioprocesses for cheaper and safer medicines.

Author Contributions

Andreas Eriksson: Data curation (lead); investigation (lead); methodology (equal); visualization (lead); writing–original draft (lead); writing–review and editing (equal). **Anne Richelle**: Data curation (supporting); formal analysis (equal); writing–review and editing (equal). **Johan Trygg**: Funding acquisition (equal); writing–review and editing (supporting). **Steffi Scholze**: Investigation (supporting); resources (lead); writing–review and editing (supporting). **Shanti Pijeaud**: Formal analysis (equal); writing–review and editing (supporting). **Henrik Antti**: Funding acquisition (equal); supervision (equal); writing–review and editing (supporting). **Christoph Zehe**: Conceptualization (equal); writing–review and editing (supporting). **Izabella Surowiec**: Conceptualization (equal); project administration (lead); supervision (equal); writing–review and editing (equal). **Pär Jonsson**: Formal analysis (equal); methodology (equal); supervision (equal); writing–review and editing (equal).

Acknowledgments

The authors thank Kate Bennett, Ilona Dudka, and the Swedish Metabolomics Centre's staff (all Umeå University) for supporting the metabolomics analysis; Merle Rattay (Sartorius) for the process variable analysis; Lennart Eriksson (Sartorius) for valuable feedback. The study was co-funded in equal measure by Umeå University's Industrial Doctoral School and Sartorius.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The processed dataset can be obtained from the corresponding author on reasonable request.

References

1. Z.-M. Li, Z.-L. Fan, X.-Y. Wang, and T.-Y. Wang, "Factors Affecting the Expression of Recombinant Protein and Improvement Strategies in Chinese Hamster Ovary Cells," *Frontiers in Bioengineering and Biotechnology* 10 (2022): 880155, https://doi.org/10.3389/fbioe.2022.880155.

2. G. Walsh and E. Walsh, "Biopharmaceutical Benchmarks 2022," *Nature Biotechnology* 40, no. 12 (2022): 1722–1760, https://doi.org/10.1038/s41587-022-01582-x.

3. S. Pereira, H. F. Kildegaard, and M. R. Andersen, "Impact of CHO Metabolism on Cell Growth and Protein Production: An Overview of Toxic and Inhibiting Metabolites and Nutrients," *Biotechnology Journal* 13, no. 3 (2018): 1700499, https://doi.org/10.1002/biot.201700499.

4. G. Yao, K. Aron, M. Borys, Z. Li, G. Pendse, and K. Lee, "A Metabolomics Approach to Increasing Chinese Hamster Ovary (CHO)," *Cell Productivity Metabolites* 11, no. 12 (2021): 823, https://doi.org/10.3390/metabo11120823.

5. C. Chen, H. E. Wong, and C. T. Goudar, "Upstream Process Intensification and Continuous Manufacturing," *Current Opinion in Chemical Engineering* 22 (2018): 191–198, https://doi.org/10.1016/j.coche.2018.10.006.

6. R. Kunert and D. Reinhart, "Advances in Recombinant Antibody Manufacturing," *Applied Microbiology and Biotechnology* 100, no. 8 (2016): 3451–3461, https://doi.org/10.1007/s00253-016-7388-9.

7. P. Kroll, K. Eilers, J. Fricke, and C. Herwig, "Impact of Cell Lysis on the Description of Cell Growth and Death in Cell Culture," *Engineering in Life Sciences* 17, no. 4 (2017): 440–447, https://doi.org/10.1002/elsc.201600088.

8. A. J. Rish, J. K. Drennen, and C. A. Anderson, "Metabolic Trends of Chinese Hamster Ovary Cells in Biopharmaceutical Production Under Batch and Fed-Batch Conditions," *Biotechnology Progress* 38, no. 1 (2022): e3220, https://doi.org/10.1002/btpr.3220.

9. M. Coulet, O. Kepp, G. Kroemer, and S. Basmaciogullari, "Metabolic Profiling of CHO Cells During the Production of Biotherapeutics," *Cells* 11, no. 12 (2022): 1929, https://doi.org/10.3390/cells11121929.

10. M. J. Gramer, "Product Quality Considerations for Mammalian Cell Culture Process Development and Manufacturing," in *Mammalian Cell Cultures for Biologics Manufacturing*, ed. W. Zhou and A. Kantardjieff (Springer, 2013), 123–166, https://doi.org/10.1007/10_2013_214.

11. R. Singh, E. Fatima, L. Thakur, S. Singh, C. Ratan, and N. Kumar, "Advancements in CHO Metabolomics: Techniques, Current State and Evolving Methodologies," *Frontiers in Bioengineering and Biotechnology* 12 (2024): 1347138, https://doi.org/10.3389/fbioe.2024.1347138.

12. G. Barberi, A. Benedetti, P. Diaz-Fernandez, et al., "Integrating Metabolome Dynamics and Process Data to Guide Cell Line Selection in Biopharmaceutical Process Development," *Metabolic Engineering* 72 (2022): 353–364, https://doi.org/10.1016/j.ymben.2022.03.015.

13. B. Worley and R. Powers, "Multivariate Analysis in Metabolomics," *Current Metabolomics* 1, no. 1 (2012): 92–107, https://doi.org/10.2174/2213235x11301010092.

14. F. Zhao, Y. Wan, L. Nie, et al., "^{1}H NMR-Based Process Understanding and Biochemical Marker Identification Methodology for Monitoring CHO Cell Culture Process during Commercial-Scale Manufacturing," *Biotechnology Journal* 18, no. 7 (2023): 2200616, https://doi.org/10.1002/biot.202200616.

15. D. J. Karst, R. F. Steinhoff, M. R. G. Kopp, et al., "Intracellular CHO Cell Metabolite Profiling Reveals Steady-State Dependent Metabolic Fingerprints in Perfusion Culture," *Biotechnology Progress* 33, no. 4 (2017): 879–890, https://doi.org/10.1002/btpr.2421.

16. L. Christie, S. Rutherford, D. S. Palmer, M. J. Baker, and H. J. Butler, "Bioprocess Monitoring Applications of an Innovative ATR-FTIR Spectroscopy Platform," *Frontiers in Bioengineering and Biotechnology* 12 (2024), https://doi.org/10.3389/fbioe.2024.1349473.

17. W. P. K. Chong, F. N. K. Yusufi, D.-Y. Lee, et al., "Metabolomics-based Identification of Apoptosis-Inducing Metabolites in Recombinant Fed-Batch CHO Culture Media," *Journal of Biotechnology* 151, no. 2 (2011): 218–224, https://doi.org/10.1016/j.jbiotec.2010.12.010.

18. S. Selvarasu, Y. S. Ho, W. P. K. Chong, et al., "Combined in Silico Modeling and Metabolomics Analysis to Characterize Fed-Batch CHO Cell Culture," *Biotechnology and Bioengineering* 109, no. 6 (2012): 1415–1429, https://doi.org/10.1002/bit.24445.

19. S. Kang, Z. Zhang, J. Richardson, et al., "Metabolic Markers Associated With High Mannose Glycan Levels of Therapeutic Recombinant Monoclonal Antibodies," *Journal of Biotechnology* 203 (2015): 22–31, https://doi.org/10.1016/j.jbiotec.2015.03.002.

20. S.-Y. Park, J. Song, D.-H. Choi, et al., "Exploring Metabolic Effects of Dipeptide Feed Media on CHO Cell Cultures by in Silico Model-Guided Flux Analysis," *Applied Microbiology and Biotechnology* 108, no. 1 (2024): 123, https://doi.org/10.1007/s00253-023-12997-0.

21. M. Alinaghi, I. Surowiec, S. Scholze, et al., "Hierarchical Time-Series Analysis of Dynamic Bioprocess Systems," *Biotechnology Journal* 17, no. 12 (2022): e2200237, https://doi.org/10.1002/biot.202200237.

22. G. Barberi, A. Benedetti, P. Diaz-Fernandez, et al., "Productive CHO Cell Lines Selection in Biopharm Process Development Through Machine Learning on Metabolomic Dynamics," *AIChE Journal* 71, no. 2 (2025): e18602, https://doi.org/10.1002/aic.18602.

23. R. Machleid, M. Hoehse, S. Scholze, et al., "Feasibility and Performance of Cross-Clone Raman Calibration Models in CHO Cultivation," *Biotechnology Journal* 19, no. 1 (2024): 2300289, https://doi.org/10.1002/biot.202300289.

24. A. Richelle, B. Corbett, P. Agarwal, A. Vernersson, J. Trygg, and C. McCready, "Model-Based Intensification of CHO Cell Cultures: One-Step Strategy From Fed-Batch to Perfusion," *Frontiers in Bioengineering and Biotechnology* 10 (2022): 948905, https://doi.org/10.3389/fbioe.2022.948905.

25. S. Wiklund, E. Johansson, L. Sjöström, et al., "Visualization of GC/TOF-MS-Based Metabolomics Data for Identification of Biochemically Interesting Compounds Using OPLS Class Models," *Analytical Chemistry* 80, no. 1 (2008): 115–122, https://doi.org/10.1021/ac0713510.

26. M. Tarapoulouzi, I. Pashalidis, and C. R. Theocharis, "Discrimination of Cheese Products Regarding Milk Species' Origin Using FTIR, 1H-NMR, and Chemometrics," *Applied Sciences* 14, no. 6 (2024): 2584, https://doi.org/10.3390/app14062584.

27. P. Jonsson, A. Wuolikainen, E. Thysell, et al., "Constrained Randomization and Multivariate Effect Projections Improve Information Extraction and Biomarker Pattern Discovery in Metabolomics Studies Involving Dependent Samples," *Metabolomics* 11, no. 6 (2015): 1667–1678, https://doi.org/10.1007/s11306-015-0818-3.

28. V. Kotu and B. Deshpande, *Data Science—Concepts and Practice*, 2nd ed. (Morgan Kaufmann, 2018), https://doi.org/10.1016/B978-0-12-814761-0.00011-3.

29. H. A. Miot, "Análise de Correlação em Estudos Clínicos e Experimentais," *Jornal Vascular Brasileiro* 17, no. 4 (2018): 275–279, https://doi.org/10.1590/1677-5449.174118.

30. V. S. Martínez, K. Rodriguez, T. McCubbin, et al., "Amino Acid Degradation Pathway Inhibitory by-Products Trigger Apoptosis in CHO Cells," *Biotechnology Journal* 19, no. 2 (2024): 2300338, https://doi.org/10.1002/biot.202300338.

31. W. P. K. Chong, S. H. Thng, A. P. Hiu, D.-Y. Lee, E. C. Y. Chan, and Y. S. Ho, "LC-MS-Based Metabolic Characterization of High Monoclonal Antibody-Producing Chinese Hamster Ovary Cells," *Biotechnology and Bioengineering* 109, no. 12 (2012): 3103–3111, https://doi.org/10.1002/bit.24580.

32. C. A. Orellana, E. Marcellin, B. L. Schulz, A. S. Nouwens, P. P. Gray, and L. K. Nielsen, "High-Antibody-Producing Chinese Hamster Ovary Cells Up-Regulate Intracellular Protein Transport and Glutathione Synthesis," *Journal of Proteome Research* 14, no. 2 (2015): 609–618, https://doi.org/10.1021/pr501027c.

33. N. Templeton, J. Dean, P. Reddy, and J. D. Young, "Peak Antibody Production Is Associated With Increased Oxidative Metabolism in an Industrially Relevant Fed-Batch CHO Cell Culture," *Biotechnology and Bioengineering* 110, no. 7 (2013): 2013–2024, https://doi.org/10.1002/bit.24858.

34. M. Saldanha, K. Padhye, V. G. Warke, P. Dandekar, and R. Jain, "A Feed Enrichment Strategy Targeting the Tricarboxylic Acid Cycle for Increasing Monoclonal Antibody Production and Alleviating Ammonia Accumulation in Chinese Hamster Ovary Cell Culture," *Biochemical Engineering Journal* 192, (2023): 108836, https://doi.org/10.1016/j.bej.2023.108836

35. X. Zhang, R. Jiang, H. Lin, and S. Xu, "Feeding Tricarboxylic Acid Cycle Intermediates Improves Lactate Consumption and Antibody Production in Chinese Hamster Ovary Cell Cultures," *Biotechnology Progress* 36, no. 4 (2020): e2975, https://doi.org/10.1002/btpr.2975

36. L. M. Carrillo-Cocom, T. Genel-Rey, D. Araíz-Hernández, et al., "Amino Acid Consumption in Naïve and Recombinant CHO Cell Cultures: Producers of a Monoclonal Antibody," *Cytotechnology* 67, no. 5 (2015): 809–820, https://doi.org/10.1007/s10616-014-9720-5

37. N. Carinhas, T. M. Duarte, L. C. Barreiro, M. J. T. Carrondo, P. M. Alves, and A. P. Teixeira, "Metabolic Signatures of GS-CHO Cell Clones Associated With Butyrate Treatment and Culture Phase Transition," *Biotechnology and Bioengineering* 110, no. 12 (2013): 3244–3257, https://doi.org/10.1002/bit.24983

38. B. C. Mulukutla, J. Kale, T. Kalomeris, M. Jacobs, and G. W. Hiller, "Identification and Control of Novel Growth Inhibitors in Fed-Batch Cultures of Chinese Hamster Ovary Cells," *Biotechnology and Bioengineering* 114, no. 8 (2017): 1779–1790, https://doi.org/10.1002/bit.26313

39. J. Dean and P. Reddy, "Metabolic Analysis of Antibody Producing CHO Cells in Fed-Batch Production," *Biotechnology and Bioengineering* 110, no. 6 (2013): 1735–1747, https://doi.org/10.1002/bit.24826

40. V. Chevallier, E. M. Schoof, L. Malphettes, M. R. Andersen, and C. T. Workman, "Characterization of Glutathione Proteome in CHO Cells and Its Relationship With Productivity and Cholesterol Synthesis," *Biotechnology and Bioengineering* 117, no. 11 (2020): 3448–3458, https://doi.org/10.1002/bit.27495

41. P. Jenner, C. Falup-Pecurariu, V. Leta, et al., "Adopting the Rumsfeld Approach to Understanding the Action of Levodopa and Apomorphine in Parkinson's Disease," *Journal of Neural Transmission* 130, no. 11 (2023): 1337–1347, https://doi.org/10.1007/s00702-023-02655-0

42. C. W. Tuthill, J. Papkoff, B. A. Watkins, and S. T. Sonis, "The Effect of the Immunomodulatory Peptide Gamma-D-glutamyl-L-tryptophan in Leukemia, Lymphoma, and Head and Neck Cancer Xenograft Models," *Journal of Clinical Oncology* 29, no. 15 (2011): 5592–5592, https://doi.org/10.1200/jco.2011.29.15_suppl.5592

43. B. Niemann, S. Haufs-Brusberg, L. Puetz, et al., "Apoptotic Brown Adipocytes Enhance Energy Expenditure via Extracellular Inosine," *Nature* 609, no. 7926 (2022): 361–368, https://doi.org/10.1038/s41586-022-05041-0

44. M. Wang, M. Xu, Y. Long, et al., "High Throughput Cell-Based Assay for Identification of Glycolate Oxidase Inhibitors as a Potential Treatment for Primary Hyperoxaluria Type 1," *Scientific Reports* 6, no. 1 (2016): 34060, https://doi.org/10.1038/srep34060

45. B. Capella Roca, N. Lao, N. Barron, P. Doolan, and M. Clynes, "An Arginase-Based System for Selection of Transfected CHO Cells Without the Use of Toxic Chemicals," *The Journal of Biological Chemistry* 294, no. 49 (2019): 18756–18768, https://doi.org/10.1074/jbc.RA119.011162

Supporting Information

Additional supporting information can be found online in the Supporting Information section.