Trends in Biotechnology

CellPress

Review

Driving towards digital biomanufacturing by CHO genome-scale models

Seo-Young Park [1,7], Dong-Hyuk Choi [1,7], Jinsung Song [1], Meiyappan Lakshmanan [2], Anne Richelle [3], Seongkyu Yoon [4], Cleo Kontoravdi [5], Nathan E. Lewis [6], and Dong-Yup Lee [1,*]

Genome-scale metabolic models (GEMs) of Chinese hamster ovary (CHO) cells are valuable for gaining mechanistic understanding of mammalian cell metabolism and cultures. We provide a comprehensive overview of past and present developments of CHO-GEMs and *in silico* methods within the flux balance analysis (FBA) framework, focusing on their practical utility in rational cell line development and bioprocess improvements. There are many opportunities for further augmenting the model coverage and establishing integrative models that account for different cellular processes and data for future applications. With supportive collaborative efforts by the research community, we envisage that CHO-GEMs will be crucial for the increasingly digitized and dynamically controlled bioprocessing pipelines, especially because they can be successfully deployed in conjunction with artificial intelligence (AI) and systems engineering algorithms.

Highlights

The reliability and methodology of genome-scale metabolic models (GEMs) of Chinese hamster ovary (CHO) cells have advanced.

CHO-GEMs have aided in cell line and process development, thus impacting on biomanufacturing efficiency.

An integrative model structure can incorporate multiple layers and capture condition-specific cell regulation.

Integration of CHO-GEMs with artificial intelligence (AI) and advanced algorithms will enable autonomous bioreactor management for digital biomanufacturing.

Unleashing the potential: refining biomanufacturing with GEMs of CHO cells

CHO cells have been the dominant host for producing **recombinant therapeutic proteins** (**RTPs**; see Glossary) in biomanufacturing [1]. Given their proficiency in generating human-compatible glycosylation patterns and continuous progress in both cell line and bioprocess development, they have been very valuable for enhancing the production of RTPs [2]. These improvements have relied on empirical methods that demand extensive time and resources. However, the complex and non-linear nature of biological systems pose a significant challenge, impeding the development of precise control systems that are essential for consistent product quality and intensified productivity. In response, the biopharmaceutical industry is pivoting from these empirical methods to a new era of process digitalization and automation. This paradigm shift aims to forge a more streamlined, reliable, and cost-efficient bioprocessing platform to reduce deviations and accelerate the manufacturing of high-quality RTPs [3]. A comprehensive mechanistic understanding of cellular characteristics and their dynamic behaviors during cell culture will help the industry to achieve a practical digital representation of bioprocesses.

A key research milestone in the cell culture field was the sequencing of the CHO cell genome [4] which provided a valuable resource of the myriad genetic elements associated with culture phenotypes, including transfection efficiency, genetic stability, growth rates, and productivity. These data also enabled the reconstruction of **GEMs** which mathematically link genotype with phenotype. These CHO-GEMs can act as digital replica of CHO cells, providing a lens through which to interpret cellular behaviors and metabolic states in biomanufacturing [3]. However, despite their great potential for the rational design of CHO cell lines and associated cell culture processes, we continue to face challenges and limitations in terms of model reliability, method development, and practical application. Hence, we present the evolution and current status of CHO-GEMs, and explore their industrial or industrially relevant applications. We also propose future directions for

[1]School of Chemical Engineering, Sungkyunkwan University, Suwon, Gyeonggi-do 16419, Republic of Korea
[2]Department of Biotechnology, Bhupat and Jyoti Mehta School of Biosciences, and Centre for Integrative Biology and Systems Medicine (IBSE), Indian Institute of Technology Madras, Chennai 600036, Tamil Nadu, India
[3]Sartorius Corporate Research, Avenue Ariane 5, 1200 Brussels, Belgium
[4]Department of Chemical Engineering, University of Massachusetts Lowell, Lowell, MA 01850, USA
[5]Department of Chemical Engineering and Chemical Technology, Imperial College London, South Kensington Campus, London SW7 2AZ, UK
[6]Departments of Pediatrics and Bioengineering, University of California San Diego, La Jolla, CA 92093, USA
[7]Equal contributions

*Correspondence: dongyuplee@skku.edu (D.-Y. Lee).

model enhancement, focusing on overcoming existing hurdles and moving towards a fully virtual CHO cell system.

Mapping: decoding mammalian cell metabolism with CHO genome-based models

The initial CHO-GEM, designated *i*CHO1766 [5], emerged from a synergistic collaboration within the scientific community and was built from the CHO cell genome sequence unveiled in 2011 [6]. This model enabled subsequent advances where its applicability, quality, and reliability saw continuous improvement. Furthermore, researchers have been able to further incorporate high-throughput omics data (e.g., transcriptomics, proteomics, metabolomics) to model different culture conditions using diverse *in silico* methods within the framework of FBA (Box 1). Consequently, this comprehensive approach allowed a deeper mechanistic understanding of CHO cell metabolism, paving the way for systematic optimization of cell cultures and pinpointing potential engineering targets.

Onward and upward of CHO-GEMs: model applicability and quality in focus

The *i*CHO1766 reconstruction provided a genetic and metabolic backbone for deriving cell line-specific models (Figure 1). For example, *i*CHO1766 has been tailored using relevant experimental and omics data to reflect the metabolism of parental CHO cell lines such as CHO-K1, -S, and -DG44, thereby predicting their growth phenotype, auxotrophy, and potential productivity upon various biochemical treatments [5]. Similarly, a GEM for a CHO-K1-derived producer cell (i.e., CHO-SH87) was generated with the necessary amino acid composition for antibody synthesis and served as a scaffold for combining multi-omics profiles (e.g., transcriptomics, metabolomics, lipidomics, and glycomics). This study elucidated the genomic rearrangements and metabolic adjustments occur in CHO cells following transgene integration [7]. Likewise, time-series transcriptomic and exometabolomic data were used to refine parental and producer GEMs to understand their metabolic behaviors under various culture conditions [8–10]. Recently, the cell line specificity of CHO-GEMs was further assessed by exploring their compositional variations in biomass constituents such as DNA, RNA, proteins, lipids, and other relevant components [11,12].

The quality of the CHO-GEMs has been persistently enhanced by refining genomic contents (i.e., annotated sets of genes, reactions, and metabolic gaps) and extending the coverage by including functional processes and regulatory elements based on the secretome and **kinetome**. For instance, *i*CHO2048s captured the protein post-translational biosynthetic machinery and secretory pathways spanning across multiple cellular organelles such as endoplasmic reticulum (ER), the Golgi apparatus, and the endomembrane system onto the genome-scale metabolic network, thus allowing the prediction of metabolic costs and bottlenecks linked with various RTPs [13]. Comparably, *i*CHO2291 incorporated publicly available kinetic information such as the turnover numbers (k_{cat}) and molecular weights of all the metabolic enzymes accounted for in the CHO-GEM to estimate a more plausible internal flux distribution by properly allocating enzymatic

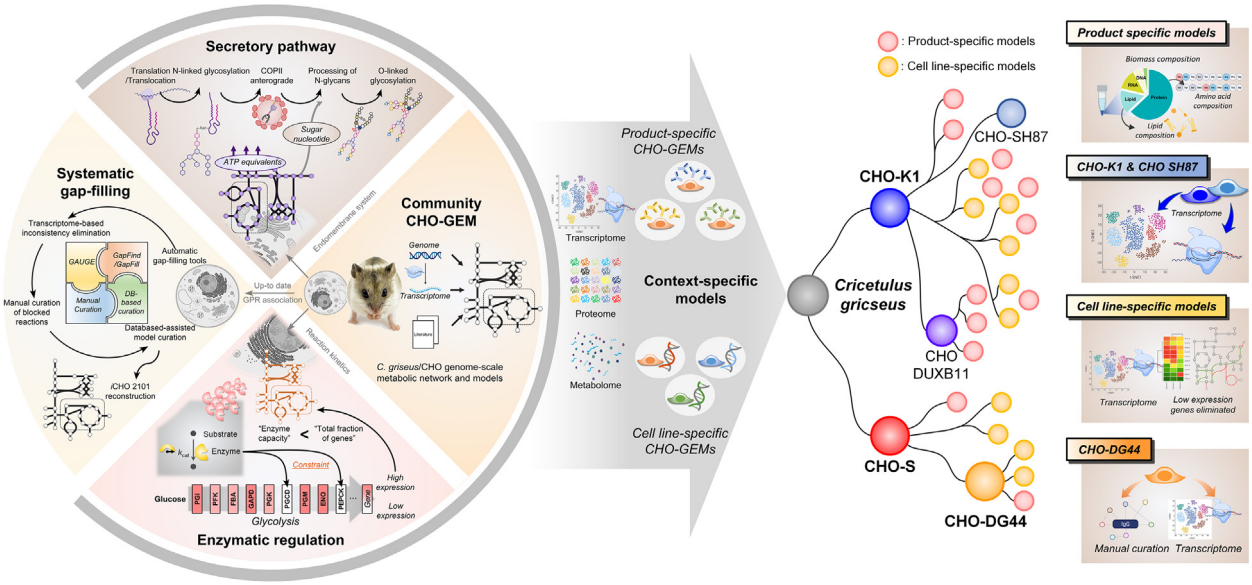
Box 1. Flux balance analysis (FBA)

FBA is the most basic **constraint-based method** used for modeling which enables the representation of intracellular biochemical and transport reactions, as well as the characterization of cell state, by using flux values with the steady-state assumption [60]. The stoichiometric matrix (*S*) represents the relationships between the various metabolites (*m*) involved in the reactions (*n*) of the system. Thus, the stoichiometric matrix captures the conversion of reactants into products in each reaction. Each row corresponds to a metabolite, and each column corresponds to a reaction, with entries in the matrix indicating the stoichiometric coefficients. The central idea of FBA is to optimize a cellular objective, often the production of biomass or another cellular component (e.g., ATP), subject to the constraints imposed by the stoichiometric matrix and other environmental conditions. The optimization problem is formulated as a linear programming problem that aims to find a set of fluxes that maximize or minimize the objective function. For example, the FBA framework idealizes cellular metabolism through the optimization of a specific metabolic objective, such as maximizing the growth rate, as with the **biomass objective function (BOF)** [60].

Glossary

Biomass objective function (BOF): a mathematical equation used in constraint-based models to represent the biomass components necessary to meet the growth requirements of a cell.
Constraint-based methods: computational techniques utilize genetic, biochemical, and physiological constraints to model, simulate, and analyze biological systems; these are especially applied to GEMs for studying the behavior of metabolic networks under various conditions.
Enzyme capacity: the kinetic parameters, in terms of molecular weight and turnover number of each enzyme and the available intracellular protein content, are obtained from publicly available databases (e.g., BRENDA, UNIPROT, and KEGG) and relevant studies.
Gene set enrichment analysis: a computational method used to determine whether predefined sets of genes are differentially expressed under different biological states or conditions.
Genome-scale metabolic models (GEMs): comprehensive mechanistic and mathematical representations that integrate information from the genome of an organism to predict and analyze the complex network of biochemical reactions occurring in its metabolism.
Kinetome: the overall profile of kinetic information such as kinetic parameters (turnover number) and molecular weights with gene-protein-reaction (GPR) information.
Post-translational modifications (PTMs): protein modifications occur after protein synthesis. These modifications, including phosphorylation, glycosylation, acetylation, methylation, and ubiquitination, play vital roles in protein maturation and functionality.
Recombinant therapeutic proteins (RTPs): non-native proteins that are used to treat various diseases and disorders such as cancer, inflammation, and genetic disorders. CHO cells are the primary producers of RTPs, particularly for monoclonal and bispecific antibodies (mAbs and bsAbs, respectively).
Uptake rate objective function (UOF): a mathematical equation that is used to predict the optimal nutrient uptake rates for cellular growth and function.

(A)



(B)

Type of model	Generic CHO-GEMs					Contextualized CHO-GEMs			
Name	*i*CHO1766	*i*CHO2291	*i*CHO2048s	*i*CHO2101	*i*CHO2441	CHO-K1	CHO-DG44	CHO-biomass	CHO-parameter
Genes	1766	2291	2048	2101	2441	1250	1188	1132-1766	1043
Reactions	3229	5587	6770	7346	6337	4710	3942	6648	2375
Metabolites	3034	3972	4663	4527	4174	2769	2750	4473	587
Year	2016	2020	2020	2021	2023	2017	2019	2020	2021
Ref.	Hefzi *et al.*	Yeo *et al.*	Gutierrez *et al.*	Fouladiha *et al.*	Strain *et al.*	Yusufi *et al.*	Calmels *et al.*	Széliová *et al.*	Schinn *et al.*

Trends in Biotechnology

Figure 1. Development and improvement of Chinese hamster ovary (CHO) cell genome-scale metabolic models (GEMs) for deciphering metabolic complexity. (A) The *i*CHO1766 model serves as a genetic and metabolic backbone for further deriving cell line- and product-specific models by incorporating secretory pathways, enzyme capacity constraints, and a systematic gap-filling algorithm and combining omics datasets. (B) Various versions of reconstructed CHO-GEMs are presented in chronological order, providing information on the model components for generic models (*i*CHO1766 [5], *i*CHO2291 [14], *i*CHO2048s [13], *i*CHO2101 [15], and *i*CHO2441 [16]) and contextualized models (represented as CHO-K1 [7], CHO-DG44 [9], CHO-biomass [12], and CHO-parameters [8]). Abbreviations: COPII, coat protein complex II; GAUGE, gap analysis using gene expression data; GPR, gene-protein-reaction association.

costs and constraining the total **enzyme capacity** within the cells [14]. In parallel with the model expansions, additional efforts improved the metabolic coverage of GEMs by augmenting the lumped and multistep expression of biosynthetic pathways (e.g., lipid metabolism and glycan biosynthesis) [14] and by resolving the gaps and dead-ends present within the CHO metabolic network using various algorithms [15,16].

Forward with prediction fidelity: improving reliability of CHO-GEMs

One of the well-known challenges in implementing FBA of large-scale GEMs such as CHO cells is the difficulty of reliably predicting their cell culture behavior and interpreting the metabolic activity based on the resultant intracellular flux distributions. These challenges stem from the complexity and regulatory uncertainty in relating multiple inputs (e.g., glucose and amino acids) to multiple outputs (e.g., biomass, toxic metabolites, antibody proteins, etc.), whereby numerous alternative feasible solutions are possible for the same phenotype. This inherent limitation is due to the underdetermined nature of the models, which contain more flux variables than metabolite balance equations, as well as the unknown regulation of metabolic enzymes. Thus, it is desirable to narrow down the solution space and determine more plausible metabolic states [17], and

this can be achieved by imposing additional constraints and by using different objective functions within the FBA framework [8].

Implementation of improved constraints, based on condition-dependent empirically derived assumptions, has yielded better predictions on cellular phenotypes (e.g., growth and nutrient uptake rates). For example, variations in measurements of nutrients and byproducts in cell cultures were deduced by principal component analysis of 21 datasets of reactor experiments. There were used to relax the constraints on their exchange rates (HybridFBA), thus effectively dealing with the potential technical errors in analytical measurements [18]. In addition, the essentiality of key nutritional components was predicted in three CHO cell lines (CHO-K1, -S, -DG44) by minimizing their uptake rates (**uptake rate objective function, UOF**) instead of maximizing the cell growth – which is routinely used as the objective function in FBA [19]. Experimentally determined and computationally estimated non-growth-associated maintenance energy (i.e., maintenance ATP) was also parameterized in CHO-GEMs to reflect the conditional energy requirements of producers and non-producers [20].

The reliability of intracellular flux predictions can be enhanced by FBA variants. Flux sampling can capture a broader spectrum of flux distributions, including suboptimal states, without considering any cellular objectives [16,21], while parsimonious FBA (pFBA) predicts the most plausible flux distribution among multiple others based on the assumption that cells minimize their overall metabolic activity during resource utilization [22]. In carbon constraint FBA (ccFBA) [23] and nitrogen/carbon constraints FBA (nccFBA) [16], the total amounts of carbon/nitrogen in the culture media were added to the elemental balance constraints to remove unrealistic internal loops or futile cycles, thus constraining the bounds of model-predicted flux to a more realistic range. Therefore, the utilization of various carbon and nitrogen sources (glucose, lactate, amino acids) required for cell growth and RTP production was better estimated, and could possibly identify desirable media conditions for the enhanced antibody productivity. Notably, enzyme capacity-constrained FBA (ecFBA) was applied to *i*CHO2291 by integrating kinetic parameters (the kinetome) to improve its prediction fidelity, thus more realistically portraying resource allocation and regulatory processes in CHO cell cultures [14]. This approach successfully demonstrated the reliability of flux predictions in *i*CHO2291 by using ^{13}C isotope-labeled measurements, and this enabled a flux-based understanding of lactate overflow metabolism which is rarely captured by conventional approaches.

Searching and navigating: leveraging CHO-GEMs in bioprocess development

With growing interest in utilizing CHO-GEMs for bioprocess development, their practical applications have primarily focused on upstream processes; only a limited number of studies have explored their potential in cell line development. We review here how CHO-GEMs can guide the systematic and rational development of both cell lines and culture processes (Figure 2). The insights from these models are advancing our mechanistic understanding and proficiency in bioprocessing.

Cell line development and engineering

The conventional procedure for CHO cell line development requires considerable time and resource investments. Indeed, it involves multiple steps, ranging from screening and isolation of cells to clone selection and characterization. Thus, we need methods for more affordable and efficient development of high-yielding expression hosts; such methods can be based on GEM-guided analysis to help understand the mechanisms underlying clonal variation, stability, productivity, and product quality, as well as to identify cell engineering targets. To this end, recent studies have combined enzyme assays [24] and time-series transcriptomics [10] with CHO-GEMs. In such studies, a comparative flux analysis between high and low producers identified metabolic bottlenecks in the

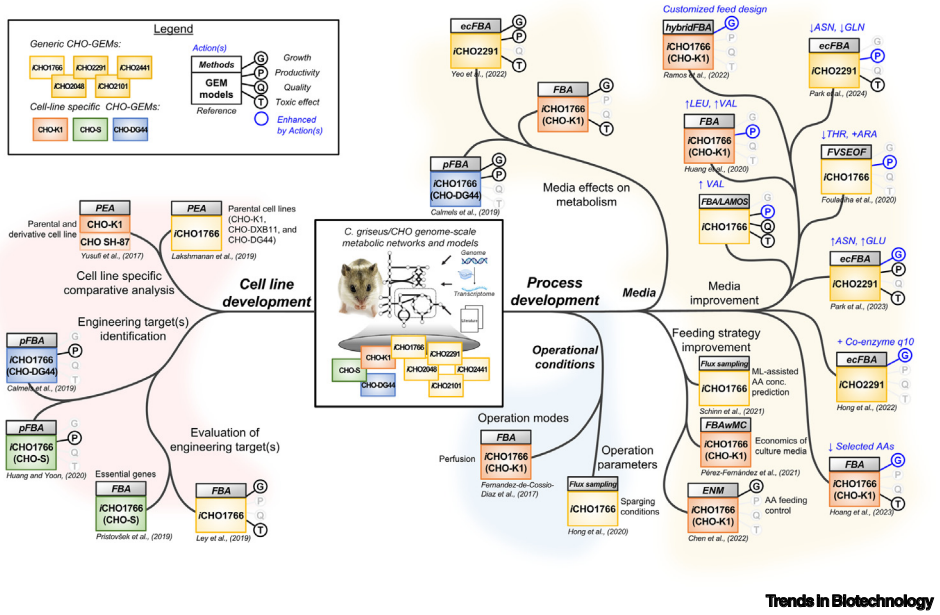


Figure 2. Key applications of Chinese hamster ovary (CHO) cell genome-scale metabolic models (GEMs). A phylogenetic tree is used to depict these publications where each box corresponds to a specific FBA method (upper smaller segment), a GEM (lower larger segment), and research target(s) such as growth, productivity, product quality, and toxic effects. The generic models are represented in yellow, while the cell line-specific models are depicted using different colors: CHO-K1 (orange), CHO-S (green), and CHO-DG44 (blue). In the box with branching arm(s), research targets such as growth, productivity, product quality, and toxic effects are indicated. Depending on the specific focus of each study, the research targets are represented by larger circles. In addition, for cases involving improvements in media, instances of target enhancements are denoted by blue circles, with the action(s) at the top of the box. The arrows above represent an increase in the concentration of the mentioned component(s), whereas the arrows below indicate a decrease; the plus sign (+) denotes the addition of new additives to the media. Abbreviations: AA, amino acid; ARA, arachidonate; ASN, asparagine; ASP, aspartate; conc., concentration; *C. griseus, Cricetulus griseus*; ecFBA, enzyme capacity-constrained flux balance analysis; ENM, essential nutrient minimization; FBA, flux balance analysis; FBAwMC, FBA with molecular crowding; FVSEOF, flux variability scanning based on enforced objective flux; GLN, glutamine; GLU, glutamate; LAMOS, linear algorithm for finding multiple optimal solutions; LEU, leucine; ML, machine learning; PEA, pathway enrichment analysis; pFBA, parsimonious FBA; THR, threonine; VAL, valine. See [7,9,10,18,24–27,29,31,33–39,49].

pentose phosphate pathway (PPP) and tricarboxylic acid (TCA) cycle, and also suggested possible engineering strategies for enhanced growth and productivity. Similarly, metabolic connectivity information derived from CHO-GEMs has been combined with multi-omics profiles for **gene set enrichment analysis** to characterize various parental CHO cells including CHO-K1, -DG44 and -DXB11. This analysis revealed variations in cell-cycle progression and protein processing induced by the absence of the dihydrofolate reductase (DHFR) [25]. Another study demonstrated the utility of CHO-GEM for engineering CHO-S cell-derived producers by computationally evaluating the lethality of gene candidates for integration sites [26] and CRISPR/Cas9-based gene editing [27]. For next-generation cell line development and engineering, omics-based model-guided strain design algorithms can be applied to systematically identify and prioritize context-specific targets for genome editing [28].

Process development

CHO-GEM-based FBA was successfully demonstrated for process development in multiple studies. Mechanistic insights into phenotypic responses from metabolic changes during cell cultures were revealed under various environmental conditions, specifically changes in the nutritional composition of the media, process parameters [e.g., dissolved oxygen (DO), temperature, pH,

aeration, etc.] and operational modes. For model-driven media development, phenotypic specificities and metabolic bottlenecks in CHO cells could be characterized to more systematically develop strategies to improve their culture performance. For example, a flux scanning approach was applied to shortlist candidate nutrients that may increase monoclonal antibody (mAb) productivity, thus identifying 15 metabolites for a subsequent step in the design of experiment (DoE) [29]. In addition, their combinations were further explored by incorporating dietary reactions into the FBA framework, showing their optimal and synergistic sets, which await experimental validation [30]. Another study aimed to predict optimal nutrient levels in the media. The UOF approach was used to simulate the uptake rates of three essential amino acids (leucine, lysine, and valine) and guide their supplementation strategies [31]. Similarly, the growth-inhibiting effects of vital nutrients (e.g., arginine, glucose, lysine, phenylalanine, and valine) were quantified through flux sensitivity analysis, suggesting subsequent basal media designs [32]. A data-driven and CHO-GEM-guided systematic framework was recently developed to evaluate metabolic bottlenecks, and identified insufficient nutrients and metabolic wastes in antitrypsin-producing CHO-DG44 cells [33]. As part of the industrial applications of the framework for rational media improvement, nutritional additives for basal media (i.e., coenzyme Q10) [34], feed media (i.e., glutamate and asparagine) [35], and dipeptide feed media (i.e., glutamine and asparagine) [36] were targeted to debottleneck energy imbalance, byproduct accumulation, and excessive utilization of TCA cycle intermediates, respectively. These interventions were validated for their ability to increase cell growth and final titer. Concurrently, the physiological and metabolic effects of operating conditions were also explored using GEM-based approaches. For instance, considering the balance between cell density and dilution rate, the cell-specific perfusion rate was parameterized into the FBA framework to represent multiple metabolic phases (i.e., excessive nutrient, limited nutrient, and maximized yield) during continuous CHO cell culture [37]. Interestingly, this approach was further exploited to reduce the accumulation of unused nutrients or toxic byproducts, thus indicating the requirements for potential perfusion media [38]. The impact of gas sparging, one of the most important operating parameters, was also investigated by simulating the metabolic response to sparging stress under fed-batch bioreactor cultures. It revealed that elevated H_2O_2 turnover rates as a result of increased uptake of amino acids under higher sparging conditions leads to oxidative stress which could be restored by manipulating redox homeostasis [39].

Driving: propelling CHO-GEMs towards digital bioprocessing

To date, a variety of CHO-GEMs and algorithmic innovations have facilitated practical implementation for cell line and bioprocess improvements. Looking ahead, we anticipate a shift towards more digitized and dynamically managed culture processes. This evolution is expected to generate a wealth of high-throughput omics and real-time culture data, supported by state-of-the-art analytical technologies ranging from next-generation sequencing to non-invasive optical sensors. Despite these advances, the industrial application of CHO-GEMs is still at its nascent stages because the current models do not fully capture condition-specific cellular regulations that affect culture performance [40]. Thus, the community needs to (i) expand and refine the models by covering and integrating more regulatory elements spanning multiple layers and scales, and (ii) relate them to key process variables by formulating the dependency of process parameters within the GEM-FBA framework.

CHO-GEMs also need to more accurately emulate and decipher phenotypic responses to genetic and environmental shifts during cell culture. For this, further expansions of the existing CHO-GEMs will increase their comprehensiveness by including more metabolic and regulatory mechanisms. In this regard, more contextualized models have been developed based on a variety of omics datasets (e.g., genomics, transcriptomics, proteomics, and metabolomics) as outlined in Table 1. In addition to the gene-protein-reaction (GPR) and their metabolic associations, the

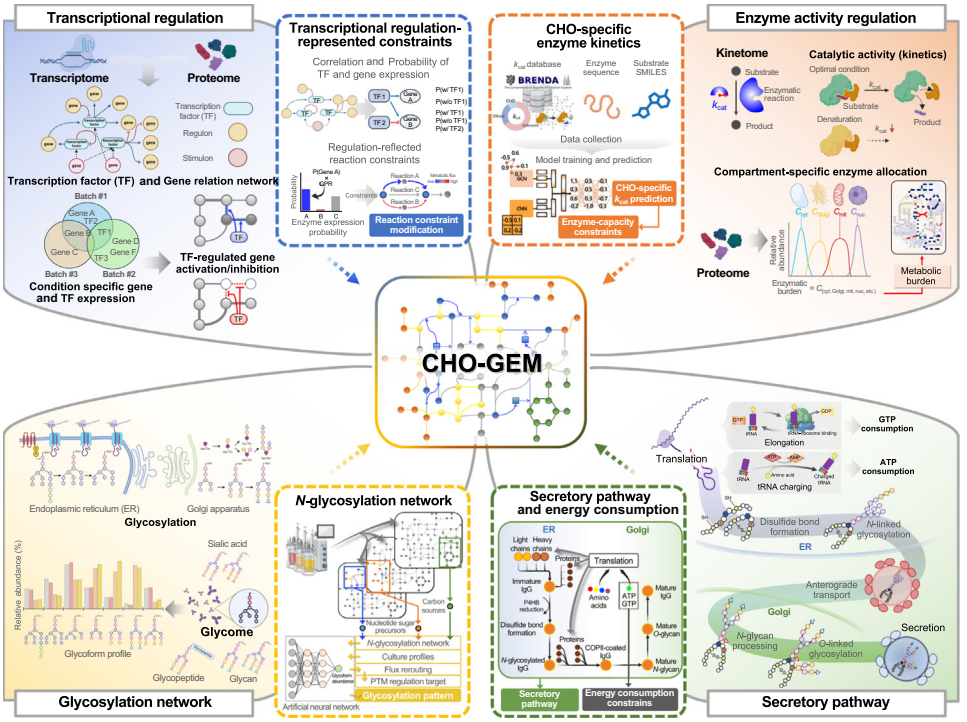
Table 1. Examples of omic integration algorithms already used with CHO-GEMs for different cell lines[a]

Algorithm	Input omic dataset	Description	Target cell line and Refs		Method Refs
GIMME	Transcriptome, proteome	• Non-classifying method • Determine active reactions above the transcriptome threshold • Calculate the penalty score based on their expression deficit between minimum flux and threshold • Minimize the total sum of (flux) × (penalty score) to constrain low-expression reactions with the estimated flux value	CHO-K1	[5,7,16]	[55]
			CHO-S	[5,7,16]	
			CHO-DG44	[9]	
			CHO-SH-87	[7]	
			CHO-GS	[10]	
iMAT	Transcriptome, proteome	• Classify reactions into high and low expression with the transcriptome or proteome threshold • Determine active/inactive reactions by comparing expression levels and possible fluxes based on MILP optimization • Remove inactive reactions while retaining low-expression reactions which are required for feasible solutions	CHO-S	[56]	[57]
mCADRE	Transcriptome, exometabolome	• Classify reactions into high and low expression using a transcriptome threshold • Define high-expression reactions as conserved core reactions • Remove unnecessary reactions that are unassociated with core reactions based on reaction connectivity	CHO-S	[56]	[58]
			CHO-DG44	[13]	
CORDA	Proteome, exometabolome	• Classify reactions into high- and low-expression reactions based on proteome abundancy • Assign a zero-reaction cost to highly expressed reactions and a higher cost to low-expression reactions • Minimize the total sum of (flux) × (cost) to define inactive reactions with flux-based threshold	CHO-K1	[16]	[59]
			CHO-S	[16]	

[a]Abbreviations: CORDA, cost optimization reaction dependency assessment; GIMME, gene inactivity moderated by metabolism and expression; iMAT, integrative metabolic analysis tool; mCADRE, metabolic context-specificity assessed by deterministic reaction evaluation; MILP, mixed integer linear programming.

models could integrate the mechanistic details of transcriptional regulation, protein secretion, **post-translational modifications (PTMs)**, and enzyme catalytic capacity (Figure 3). For example, a detailed representation of transcription factors (TFs) and other pathway machinery (e.g., protein transport, unfolded protein response, chaperones) can be achieved by expanding the scope of CHO-GEMs; meanwhile, logical and probabilistic integration of transcriptional interactions and major features of protein synthesis, folding, trafficking, and eventual release were partially presented in the models [13,41]. Among the PTMs, *N*-linked glycosylation is an important process connected to metabolism that affects the critical quality attributes (CQAs) of secreted RTPs [42,43]. To modularize this regulatory component at the molecular scale, multistep glycosylation of *N*-linked glycans in ER and the Golgi apparatus could be mechanistically described to relate to the abundance of nucleotide sugar precursors and glycan profiles. A hybrid CHO-GEM and glycosylation model can then link nutritional environment and critical process parameters (CPPs) with product quality during production culture via nucleotide sugars [44] which are differentially synthesized through cellular metabolism in a condition-dependent manner. Thus, adjustable and actionable targets can be identified to maintain desirable glycoforms of RTPs [45]. At the reaction scale, kinetic information (i.e., enzyme capacity constraints with k_{cat} values) was incorporated into the ecFBA to obtain more plausible metabolic states. However, the predictive accuracy of the model largely depends on the CHO-specific k_{cat} values, which can be further updated by resorting to machine learning/deep learning algorithms on the basis of protein sequences and/or substrate structures for a diverse array of enzymes, as similarly has been done in *Escherichia coli* [46] and fungal species [47].

With advanced multiple-layered and -scaled CHO-GEMs, condition-dependent cellular behaviors can be better predicted for more reliable industrial applications by empirically or mechanistically relating nutritional environment and CPPs (e.g., pH, temperature, DO levels, and agitation rates) with model components (e.g., reaction and enzyme activity) through mathematical or



Trends in Biotechnology

Figure 3. Multiple layers and scales of Chinese hamster ovary (CHO) cell genome-scale metabolic models (GEMs) to capture condition-specific cellular regulation that affects culture performance. This figure illustrates how to expand and enhance models by integrating regulatory components such as transcriptional regulation, enzyme activity regulation, secretory pathways, and glycosylation networks within a CHO-GEM, beyond gene-protein-reaction (GPR) associations. The concept of each content (in clockwise order starting from the upper left box) are introduced: transcriptional regulation (blue), enzyme activity regulation (orange), secretory pathway (green), and glycosylation network (yellow). In addition to the concepts, broken boxes illustrate how the relevant constraint and network can be incorporated into a CHO-GEM. Transcriptional regulation (upper left): in response to a stimulus, transcription factors (TFs) are either expressed or inhibited, leading to changed expression levels of subsequent genes associated with metabolic enzymes. Probability calculation of gene expression under particular conditions, along with consequent enzyme expression levels, and their formulation as reaction constraints facilitate the representation of regulatory mechanisms in metabolic models. Enzyme activity regulation (upper right): enzyme activity (kinetome) determines the metabolic state of a cell, making it a prerequisite for the construction of a precise GEM. However, the kinetome is rarely measured for mammalian cells. AI-driven k_{cat} prediction models utilize databases containing enzyme sequences and substrate structures to estimate enzyme activity based on an organism or cellular compartment. Enzyme capacity constraints can then be integrated into model in combination with compartment-specific enzyme allocation to construct an advanced CHO-GEM. Secretory pathway (lower right): translation and glycosylation processes consume energy cofactors and proteins, and inclusion of these components is therefore necessary. Secretory pathway-involved subsystems and their metabolic enzymes can be formulated into stepwise reactions [13] where their consumption of energy, metabolites, and proteins can be added to the reaction formula. Glycosylation network (lower left): the glycosylation pathway from the ER to the Golgi apparatus. Through advanced analyses of glycome, the glycosylation patterns can be investigated, which facilitates the optimization of production processes for desired mAb quality. The *N*-linked glycan, which is determined by a sequence of interconnected pathways, can vary depending on intracellular metabolism and the levels of nucleotide sugar precursors and carbon sources. The CHO-GEM, expanded with a glycosylation network through a hybrid approach, can identify rerouted fluxes and possible PTM targets for a desired *N*-glycosylated pattern. Abbreviations: AI, artificial intelligence; CNN, convolutional neural network; COPII, coat protein complex II; ER, endoplasmic reticulum; GCN, graph convolutional network; GPR, gene-protein-reaction; GalNAc, *N*-acetylgalactosamine; GlcNAc, *N*-acetylglucosamine; IgG, immunoglobulin G; mAb, monoclonal antibody; NeuAc, *N*-acetylneuraminic acid; P, probability; P4HB, prolyl 4-hydroxylase β-polypeptide; PTM, post-translational modification; SMILES, simplified molecular input line entry system.

logical representations. When the specific mechanisms underlying the dynamics of nuanced and fluctuating conditions in bioreactors are unclear, the FBA framework could be expanded by effectively adopting machine learning approaches, thus better predicting cellular phenotypes

or metabolite concentrations [48,49] and their future trajectories [50]. Recently, major operational variables (e.g., pH and temperature) were parameterized into GEMs for microbial systems. Specifically, the flux variables related to redox cofactor (e.g., NADH), proton transport, enzymatic turnover rate (k_{cat}), and maintenance energy were expressed as canonical functions of pH and temperature, respectively, reflecting their dependency in *E. coli* [51] and *Saccharomyces cerevisiae* [52]. Accordingly, the relevant formulations can be derived from multi-omics and culture profiles of CHO cells under various pH and temperature conditions (Box 2), thereby guiding experimental design for cell line engineering and improving process efficiency, adaptability, and operability under dynamic environments for digital bioprocessing.

Box 2. Case study: integrating operational parameters and CHO-GEMs

GEMs poses can integrate operational parameters, such as pH and temperature, to mechanistically predict CHO cell growth, productivity, and product quality (Figure I). In CHO cell culture, pH has been reported as a key parameter that affects cell growth, viability, productivity, and product quality [61]. Mechanisms associated with pH can be illustrated as multiple layers of advanced GEM, specifically transcriptional regulation, post-translational modifications, and enzyme kinetics, and can be represented as pH-dependent components. Especially regarding enzyme kinetics, pH acts as a critical parameter that drives structural deformation. Thus, enzymatic activity can be represented as a function of environmental pH and the enzymatically optimal pH.

Combined with enzyme capacity-constrained flux balance analysis (ecFBA), pH-dependent k_{cat} estimates enable the simulation of cellular metabolic states under varying pH conditions. Moreover, adopting kinetic equations for acidic or basic metabolite transportation rates between compartments, such as the cytoplasm, extracellular space, and organelles, will facilitate the calculation of compartment-specific pH and resultant enzymatic activity. Predicting k_{cat} for mitochondria, ER, and the Golgi apparatus will provide accurate predictions of fluxes related to cell proliferation and product quality which are prone to perturbations by culture phase or pH [62,63]. Similarly, temperature affects enzyme kinetics, thereby perturbing cellular systems with respect to viable cell density (VCD) and productivity [64]. However, the lack of mechanistic relationships between the temperature and the cellular response hinders the integration of temperature into GEMs. Similarly to pH, enzyme kinetics can be formulated as a function of temperature. Furthermore, the relationship between temperature- or pH-dependent effects and cell metabolism can be explored by incorporating comparative analysis of multi-omics data obtained from high-throughput condition-varying cultures.

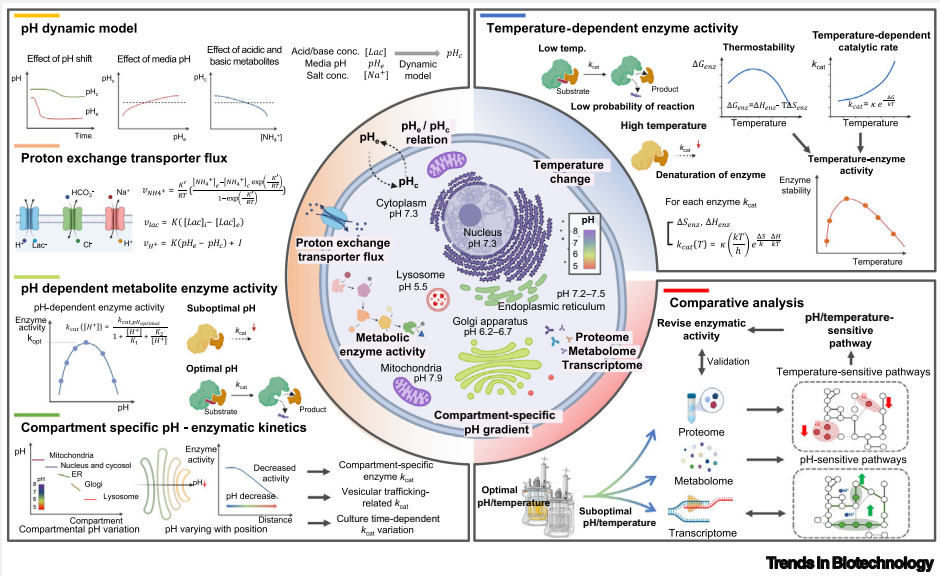


Figure I. Detailed structure of pH and temperature integration into a GEM. The model of pH dynamics and transport, which provides compartment-specific pH differences in a cell, is applicable to changes in enzymatic activity. Temperature changes are used to calculate alterations in the enzyme activity. Experimental measurements made under pH or temperature shifts can be utilized for comparative analysis to validate the integration process. Abbreviations: conc., concentration; pH_c, cytoplasmic pH; pH_e, extracellular pH.

Concluding remarks and future perspectives

Advances in CHO-GEMs and relevant analytical methods are already showing utility in several practical applications including rational cell line and process development. Further capabilities of these models will emerge following a sustained focus on augmenting model coverage and creating an integrative model structure with multiple layers, including additional cellular processes other than metabolism (e.g., enzyme and transcriptional regulation).

As the field continues to grow and evolve, we foresee enhanced predictive prowess from GEMs that will be catalyzed by sustained collaboration and dedication within the biomanufacturing community such as the International Biomanufacturing Network (IBioNe) and the European Society for Animal Cell Technology (ESACT). The community is envisaged to actively foster education through organized courses and workshops to provide valuable insights into the modeling of animal cell culture. This educational initiative not only contributes to academia but meets the growing demand for both CHO-GEMs and FBA applications within the industry (see Outstanding questions).

CHO-GEMs, renowned for their adaptability across various mammalian cell lines, are poised to revolutionize fields beyond their current scope. From recombinant adeno-associated virus (rAAV) production for gene therapy using HEK293 cells [53] to cultured meat production with animal tissues [54], they are poised to propel applications in biopharmaceutical and food production because such models mechanistically encode the processes involved in biomass and macromolecule production. In addition, GEMs are pivotal in shaping the digital twins of bioprocesses as instrumental tools for accurately tracking, forecasting, and navigating cellular behaviors and metabolic states with timely monitored and digitalized culture data. Hence, we envisage a future where GEMs, seamlessly integrated with data-driven AI models and sophisticated control algorithms, autonomously orchestrate bioreactor runs and meticulously guide the optimal culture trajectory. In this context, AI and machine learning strategies can be applied to expand the adoption of GEMs across the biopharmaceutical industry for more innovative and efficient RTP production [40]. This represents a profound culmination of scientific progress and innovation towards fully digital biomanufacturing.

Outstanding questions

Which advanced methodologies can be integrated with CHO-GEMs to enhance their predictive accuracy of cellular behaviors under diverse conditions, especially considering complex biological phenomena?

Is it feasible to develop a universal framework or set of guidelines for the creation and application of GEMs across various CHO cell lines in industrial contexts?

How can we evolve CHO-GEMs to more precisely simulate dynamic cellular responses under fluctuating bioprocessing conditions?

Which strategies can be used to fuse CHO-GEMs with regulatory components and process modifications, aiming to achieve a more holistic representation of biological interactions with the cell culture environment?

Which innovative models or methodologies hold potential to be synergized with CHO-GEMs for fully digitalizing bioprocesses?

What specific roles do academic and industrial partnerships play in propelling the advancement of biomanufacturing via CHO-GEMs, and how can these collaborations be further optimized?

How can CHO-GEMs be used for different modality cell lines, such as HEK293 for gene therapy?

Acknowledgments

D-Y.L. received a Korea Innovation Foundation grant (2021-DD-UP-0369) from the Ministry of Science and ICT; D-Y.L. and S-Y.P. received a National Research Foundation of Korea grant (2022R1A4A5032720); N.E.L. was supported by the Novo Nordisk Foundation (NNF20SA0066621), National Institute of General Medical Sciences (NIGMS, grant R35 GM119850), and Sartorius AG; S.Y. was supported by the Advanced Mammalian Biomanufacturing Innovation Center (AMBIC) by means of the Industry–University Cooperative Research Center Program under the US National Science Foundation grants (1624684 and 2100075), and the National Institute for Innovation in Manufacturing Biopharmaceuticals (NIIMBL, grant 70NANB17H002). Some components of Figures 1 and 3 and Figure I in Box 2 were created with BioRender.

Declaration of interests

The authors declare no conflicts of interests.

References

1. Walsh, G. and Walsh, E. (2022) Biopharmaceutical benchmarks 2022. *Nat. Biotechnol.* 40, 1722–1760
2. Tihanyi, B. and Nyitray, L. (2020) Recent advances in CHO cell line development for recombinant protein production. *Drug Discov. Today Technol.* 38, 25–34
3. Park, S.-Y. *et al.* (2021) Bioprocess digital twins of mammalian cell culture for advanced biomanufacturing. *Curr. Opin. Chem. Eng.* 33, 100702
4. Wurm, F.M. and Hacker, D. (2011) First CHO genome. *Nat. Biotechnol.* 29, 718–720
5. Hefzi, H. *et al.* (2016) A consensus genome-scale reconstruction of Chinese hamster ovary cell metabolism. *Cell Syst.* 3, 434–443
6. Xu, X. *et al.* (2011) The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. *Nat. Biotechnol.* 29, 735–741
7. Yusufi, F.N.K. *et al.* (2017) Mammalian systems biotechnology reveals global cellular adaptations in a recombinant CHO cell line. *Cell Syst.* 4, 530–542
8. Schinn, S.-M. *et al.* (2021) Systematic evaluation of parameters for genome-scale metabolic models of cultured mammalian cells. *Metab. Eng.* 66, 21–30

9. Calmels, C. *et al.* (2019) Application of a curated genome-scale metabolic model of CHO DG44 to an industrial fed-batch process. *Metab. Eng.* 51, 9–19
10. Huang, Z. and Yoon, S. (2020) Integration of time-series transcriptomic data with genome-scale CHO metabolic models for mAb engineering. *Processes* 8, 331
11. Choi, Y.-M. *et al.* (2023) Mitigating biomass composition uncertainties in flux balance analysis using ensemble representations. *Comput. Struct. Biotechnol. J.* 21, 3736–3745
12. Széliová, D. *et al.* (2020) What CHO is made of: variations in the biomass composition of Chinese hamster ovary cell lines. *Metab. Eng.* 61, 288–300
13. Gutierrez, J.M. *et al.* (2020) Genome-scale reconstructions of the mammalian secretory pathway predict metabolic costs and limitations of protein secretion. *Nat. Commun.* 11, 68
14. Yeo, H.C. *et al.* (2020) Enzyme capacity-based genome scale modelling of CHO cells. *Metab. Eng.* 60, 138–147
15. Fouladiha, H. *et al.* (2021) Systematically gap-filling the genome-scale metabolic model of CHO cells. *Biotechnol. Lett.* 43, 73–87
16. Strain, B. *et al.* (2023) How reliable are Chinese hamster ovary (CHO) cell genome-scale metabolic models? *Biotechnol. Bioeng.* 120, 2460–2478
17. Reed, J.L. (2012) Shrinking the metabolic solution space using experimental datasets. *PLoS Comput. Biol.* 8, e1002662
18. Ramos, J.R.C. *et al.* (2022) Genome-scale modeling of Chinese hamster ovary cells by hybrid semi-parametric flux balance analysis. *Bioprocess Biosyst. Eng.* 45, 1889–1904
19. Chen, Y. *et al.* (2019) An unconventional uptake rate objective function approach enhances applicability of genome-scale models for mammalian cells. *NPJ Syst. Biol. Appl.* 5, 25
20. Széliová, D. *et al.* (2021) Inclusion of maintenance energy improves the intracellular flux predictions of CHO. *PLoS Comput. Biol.* 17, e1009022
21. Park, J.M. *et al.* (2012) Flux variability scanning based on enforced objective flux for identifying gene amplification targets. *BMC Syst. Biol.* 6, 106
22. Lewis, N.E. *et al.* (2010) Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models. *Mol. Syst. Biol.* 6, 390
23. Lularevic, M. *et al.* (2019) Improving the accuracy of flux balance analysis through the implementation of carbon availability constraints for intracellular reactions. *Biotechnol. Bioeng.* 116, 2339–2352
24. Calmels, C. *et al.* (2019) Application of a genome-scale model in tandem with enzyme assays for identification of metabolic signatures of high and low CHO cell producers. *Metab. Eng. Commun.* 9, e2300126
25. Lakshmanan, M. *et al.* (2019) Multi-omics profiling of CHO parental hosts reveals cell line-specific variations in bioprocessing traits. *Biotechnol. Bioeng.* 116, 2117–2129
26. Pristovšek, N. *et al.* (2019) Systematic evaluation of site-specific recombinant gene expression for programmable mammalian cell engineering. *ACS Synth. Biol.* 8, 757–774
27. Ley, D. *et al.* (2019) Reprogramming AA catabolism in CHO cells with CRISPR/Cas9 genome editing improves cell growth and reduces byproduct secretion. *Metab. Eng.* 56, 120–129
28. Hong, J.K. *et al.* (2018) Towards next generation CHO cell line development and engineering by systems approaches. *Curr. Opin. Chem. Eng.* 22, 1–10
29. Fouladiha, H. *et al.* (2020) A metabolic network-based approach for developing feeding strategies for CHO cells to increase monoclonal antibody production. *Bioprocess Biosyst. Eng.* 43, 1381–1389
30. Weston, B.R. and Thiele, I. (2023) A nutrition algorithm to optimize feed and medium composition using genome-scale metabolic models. *Metab. Eng.* 76, 167–178
31. Chen, Y. *et al.* (2022) A genome-scale nutrient minimization forecast algorithm for controlling essential amino acid levels in CHO cell cultures. *Biotechnol. Bioeng.* 119, 435–451
32. Hoang, D. *et al.* (2023) Modulation of nutrient precursors for controlling metabolic inhibitors by genome-scale flux balance analysis. *Biotechnol. Prog.* 39, e3313
33. Yeo, H.C. *et al.* (2022) Combined multivariate statistical and flux balance analyses uncover media bottlenecks to the growth and productivity of Chinese hamster ovary cell cultures. *Biotechnol. Bioeng.* 119, 1740–1754
34. Hong, J.K. *et al.* (2022) Data-driven and model-guided systematic framework for media development in CHO cell culture. *Metab. Eng.* 73, 114–123
35. Park, S.-Y. *et al.* (2023) Debottlenecking and reformulating feed media for improved CHO cell growth and titer by data-driven and model-guided analyses. *Biotechnol. J.* 18, e2300126
36. Park, S.-Y. *et al.* (2024) Exploring metabolic effects of dipeptide feed media on CHO cell cultures by in silico model-guided flux analysis. *Appl. Microbiol. Biotechnol.* 108, 123
37. Fernandez-de-Cossio-Diaz, J. *et al.* (2017) Characterizing steady states of genome-scale metabolic networks in continuous cell cultures. *PLoS Comput. Biol.* 13, e1005835
38. Pérez-Fernández, B.A. *et al.* (2021) In-silico media optimization for continuous cultures using genome scale metabolic networks: the case of CHO-K1. *Biotechnol. Bioeng.* 118, 1884–1897
39. Hong, J.K. *et al.* (2020) In silico model-based characterization of metabolic response to harsh sparging stress in fed-batch CHO cell cultures. *J. Biotechnol.* 308, 10–20
40. Richelle, A. *et al.* (2020) Towards a widespread adoption of metabolic modeling tools in biopharmaceutical industry: a process systems biology engineering perspective. *NPJ Syst. Biol. Appl.* 6, 6
41. Strain, B. *et al.* (2023) Genome-scale models as a vehicle for knowledge transfer from microbial to mammalian cell systems. *Comput. Struct. Biotechnol. J.* 21, 1543–1549
42. Wei, B. *et al.* (2017) Glycation of antibodies: modification, methods and potential effects on biological functions. *mAbs* 9, 586–594
43. Sha, S. *et al.* (2020) At-line N-linked glycan profiling for monoclonal antibodies with advanced sample preparation and high-performance liquid chromatography. *J. Biosci. Bioeng.* 130, 327–333
44. Antonakoudis, A. *et al.* (2021) Synergising stoichiometric modelling with artificial neural networks to predict antibody glycosylation patterns in Chinese hamster ovary cells. *Comput. Chem. Eng.* 154, 107471
45. Sha, S. *et al.* (2016) N-glycosylation design and control of therapeutic monoclonal antibodies. *Trends Biotechnol.* 34, 835–846
46. Heckmann, D. *et al.* (2018) Machine learning applied to enzyme turnover numbers reveals protein structural correlates and improves metabolic models. *Nat. Commun.* 9, 5252
47. Li, F. *et al.* (2022) Deep learning-based kcat prediction enables improved enzyme-constrained model reconstruction. *Nat. Catal.* 5, 662–672
48. Gopalakrishnan, S. *et al.* (2024) COSMIC-dFBA: a novel multi-scale hybrid framework for bioprocess modeling. *Metab. Eng.* 82, 183–192
49. Schinn, S.-M. *et al.* (2021) A genome-scale metabolic network model and machine learning predict amino acid concentrations in Chinese hamster ovary cell cultures. *Biotechnol. Bioeng.* 118, 2118–2123
50. Park, S.-Y. *et al.* (2023) Data-driven prediction models for forecasting multistep ahead profiles of mammalian cell culture toward bioprocess digital twins. *Biotechnol. Bioeng.* 120, 2494–2508
51. Du, B. *et al.* (2019) Genome-scale model of metabolism and gene expression provides a multi-scale description of acid stress responses in *Escherichia coli*. *PLoS Comput. Biol.* 15, e1007525
52. Li, G. *et al.* (2021) Bayesian genome scale modelling identifies thermal determinants of yeast metabolism. *Nat. Commun.* 12, 190
53. Fu, Q. *et al.* (2023) Critical challenges and advances in recombinant adeno-associated virus (rAAV) biomanufacturing. *Biotechnol. Bioeng.* 120, 2601–2621
54. Gomez Romero, S. and Boyle, N. (2023) Systems biology and metabolic modeling for cultivated meat: a promising approach for cell culture media optimization and cost reduction. *Compr. Rev. Food Sci. Food Saf.* 22, 3422–3443
55. Becker, S.A. and Palsson, B.O. (2008) Context-specific metabolic networks are consistent with experiments. *PLoS Comput. Biol.* 4, e1000082
56. Gopalakrishnan, S. *et al.* (2023) Guidelines for extracting biologically relevant context-specific metabolic models using gene expression data. *Metab. Eng.* 75, 181–191

57. Zur, H. *et al.* (2010) iMAT: an integrative metabolic analysis tool. *Bioinformatics* 26, 3140–3142
58. Wang, Y. *et al.* (2012) Reconstruction of genome-scale metabolic models for 126 human tissues using mCADRE. *BMC Syst. Biol.* 6, 153
59. Schultz, A. and Qutub, A.A. (2016) Reconstruction of tissue-specific metabolic networks using CORDA. *PLoS Comput. Biol.* 12, e1004808
60. Orth, J.D. *et al.* (2010) What is flux balance analysis? *Nat. Biotechnol.* 28, 245–248
61. Grilo, A.L. and Mantalaris, A. (2019) Apoptosis: a mammalian cell bioprocessing perspective. *Biotechnol. Adv.* 37, 459–475
62. Grabe, M. and Oster, G. (2001) Regulation of organelle acidity. *J. Gen. Physiol.* 117, 329–343
63. Lee, J.H. *et al.* (2019) Analysis of Golgi pH in Chinese hamster ovary cells using ratiometric pH-sensitive fluorescent proteins. *Biotechnol. Bioeng.* 116, 1006–1016
64. Xu, J. *et al.* (2019) Systematic development of temperature shift strategies for Chinese hamster ovary cells based on short duration cultures and kinetic modeling. *mAbs* 11, 191–204