



**HAL**  
open science

# Building blocks needed for mechanistic modeling of bioprocesses: A critical review based on protein production by CHO cells

Yusmel González-Hernández, Patrick Perré

## ► To cite this version:

Yusmel González-Hernández, Patrick Perré. Building blocks needed for mechanistic modeling of bioprocesses: A critical review based on protein production by CHO cells. *Metabolic Engineering Communications*, 2024, 18, pp.e00232. 10.1016/j.mec.2024.e00232 . hal-04515543

**HAL Id: hal-04515543**

**<https://hal.science/hal-04515543>**

Submitted on 21 Mar 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

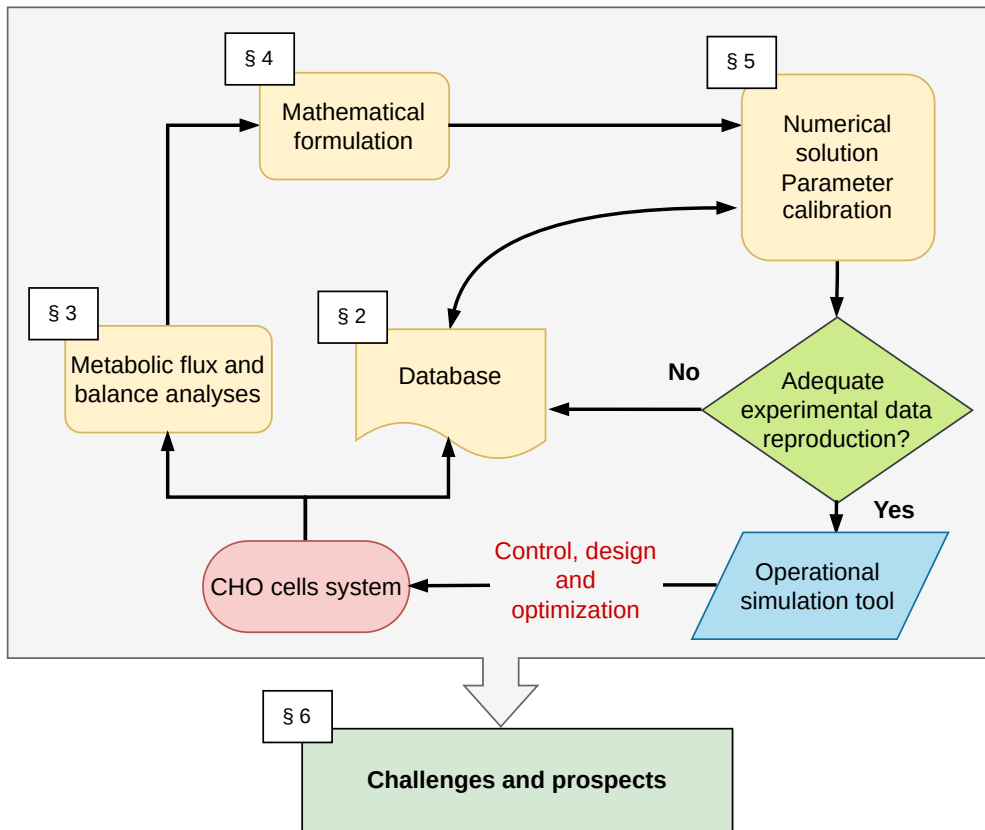
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 Graphical Abstract

2 Building Blocks Needed for Mechanistic Modeling of Bioprocesses: a

3 Critical Review Based on Protein Production by CHO Cells

4 Yusmel González-Hernández, Patrick Perré



5 Building Blocks Needed for Mechanistic Modeling of  
6 Bioprocesses: a Critical Review Based on Protein  
7 Production by CHO Cells

8 Yusmel González-Hernández<sup>a</sup>, Patrick Perré<sup>a</sup>,

9 <sup>a</sup> *Université Paris-Saclay, CentraleSupélec, Laboratoire de Génie des Procédés et Matériaux,*  
10 *Centre Européen de Biotechnologie et de Bioéconomie (CEBB), 3 rue des Rouges Terres 51110*  
11 *Pomacle, France*

---

12 **Abstract**

13 This paper reviews the key building blocks needed to develop a mechanistic model  
14 for use as an operational production tool. The Chinese Hamster Ovary (CHO) cell,  
15 one of the most widely used hosts for antibody production in the pharmaceutical  
16 industry, is considered as a case study. CHO cell metabolism is characterized by  
17 two main phases, exponential growth followed by a stationary phase with strong  
18 protein production. This process presents an appropriate degree of complexity to  
19 outline the modeling strategy. The paper is organized into four main steps: (1)  
20 CHO systems and data collection; (2) metabolic analysis; (3) formulation of the  
21 mathematical model; and finally, (4) numerical solution, calibration, and validation.  
22 The overall approach can build a predictive model of target variables. According  
23 to the literature, one of the main current modeling challenges lies in understanding  
24 and predicting the spontaneous metabolic shift. Possible candidates for the trigger  
25 of the metabolic shift include the concentration of lactate and carbon dioxide. In  
26 our opinion, ammonium, which is also an inhibiting product, should be further  
27 investigated. Finally, the expected progress in the emerging field of hybrid modeling,  
28 which combines the best of mechanistic modeling and machine learning, is presented  
29 as a fascinating breakthrough. Note that the modeling strategy discussed here is a  
30 general framework that can be applied to any bioprocess.

---

*Email address: patrick.perre@centralesupelec.fr (Patrick Perré)*

31 **Highlights**

- 32 • An overview of the main metabolic pathways in CHO cells is provided,
- 33 • A simple and intuitive strategy for modeling CHO cell metabolism is outlined,
- 34 • Inverse analysis using evolutionary algorithms is a powerful tool for model  
35 calibration,
- 36 • The main open questions and challenges are summarized,
- 37 • Recommendations and medium-term prospects are proposed for further work.

38 *Keywords:* CHO cells, model calibration, metabolic shift, activation-inhibition,  
39 digital twin

---

40 **Contents**

41	<b>1 Introduction</b>	<b>4</b>
42	<b>2 CHO cells system</b>	<b>7</b>
43	2.1 CHO culture conditions . . . . .	7
44	2.2 Gene amplification systems . . . . .	10
45	2.2.1 DHFR-based MTX selection . . . . .	10
46	2.2.2 GS-based MSX selection . . . . .	10
47	2.3 Measured data . . . . .	11
48	<b>3 Metabolic pathway identification</b>	<b>13</b>
49	3.1 Metabolic flux analyses . . . . .	13
50	3.2 Metabolic pathways description . . . . .	17
51	3.3 Metabolic shift . . . . .	18
52	3.4 Main acid contributors to pH . . . . .	21

53	<b>4 Mathematical formulation</b>	<b>22</b>
54	4.1 Kinetics of biological processes . . . . .	23
55	4.2 Kinetics of glutamine metabolism . . . . .	28
56	4.3 Kinetics of physical and chemical processes . . . . .	29
57	4.4 Critical analysis of the published CHO mechanistic models . . . . .	31
58	<b>5 Numerical solution and calibration</b>	<b>35</b>
59	5.1 Computational solutions of ODEs . . . . .	35
60	5.2 Calibration and validation . . . . .	38
61	5.2.1 Parameter estimation by inverse analysis method . . . . .	39
62	5.2.2 Calibration challenges . . . . .	41
63	5.3 Optimization methods for minimizing the objective function . . . . .	42
64	<b>6 Challenges and prospects of mechanistic modeling of CHO cells</b>	<b>47</b>
65	6.1 Needs, concerns and improvement gaps . . . . .	47
66	6.2 The digital twin at the crossroads of mechanistic modeling and data	
67	science . . . . .	50
68	<b>7 Conclusions</b>	<b>53</b>
69	<b>8 Acknowledgments</b>	<b>54</b>

70 **1. Introduction**

71 The operation of bioreactors in bioproduction still needs to be exploited empirically,  
72 with a small amount of information collected online by sensors. However, the po-  
73 tential of mechanistic models as operational production tools is likely to increase  
74 efficiency, for example, through the early detection and correction of issues. In this  
75 review paper, bioproduction by Chinese Hamster Ovary (CHO) cells is used as an  
76 example to detail the key bricks needed to build a mechanistic model and how they  
77 work together. We chose this because CHO cells are the most commonly mammalian  
78 host used for therapeutic protein production in the pharmaceutical industry (Yang

79 [et al., 2022](#)). The metabolism of CHO cells includes an initial growth phase, with  
80 a high level of lactate production, and a subsequent stationary phase, where cell  
81 growth has slowed or stopped, and recombinant protein production is high ([Dean  
82 and Reddy, 2013](#)). This succession of phases is efficient as it separates the growth  
83 phase, where resources are used to increase the population, from the stationary  
84 phase, where the cells use resources to produce the protein of interest ([Sengupta  
85 et al., 2011](#)). Nevertheless, the factors and mechanisms that trigger this metabolic  
86 shift between the exponential and stationary phases observed in CHO cultures re-  
87 main poorly understood ([Hartley et al., 2018](#); [Yahia et al., 2021](#)). Metabolic flux  
88 analysis (MFA) and flux balance analysis (FBA) are essential in mechanistic un-  
89 derstanding of cell metabolism for optimal production process planning and design  
90 ([Huang et al., 2017](#); [Sha et al., 2018](#)). Despite tremendous progress, much remains  
91 to be done to understand CHO cell metabolism ([Marx et al., 2022](#)).

92 This production process is complex, and combining productivity, product quality,  
93 efficiency, and consistency remains challenging ([Luo et al., 2021](#)). This complexity  
94 explains why many industrial strategies for protein production by CHO cells remain  
95 mainly based on empirical results ([Calmels et al., 2019](#)). Mechanistic modeling is a  
96 powerful tool that combines knowledge from different sciences (biology, chemistry,  
97 physics, maths, etc.) that can be used for assumption testing, experimental design,  
98 and control/optimization of bioprocesses. However, before mechanistic modeling  
99 can be used as an operational tool, developing such models requires a demand-  
100 ing experimental effort for calibration and validation. Over-parameterized models  
101 could affect this process, which could turn them into models lacking robustness and  
102 universality ([Tsopanoglou and del Val, 2021](#)). Several models have already been  
103 developed for simulating the CHO cell metabolism ([Xing et al., 2010](#); [Nolan and  
104 Lee, 2011](#); [López-Meza et al., 2016](#); [Jimenez del Val et al., 2016](#); [Kotidis et al., 2019](#);  
105 [Yahia et al., 2021](#)). But most models are poorly identified, and the description of

106 the metabolic pathways and indicators used for the prediction of metabolic shift is  
107 still inadequate. Sometimes, models are either too complex or straightforward for  
108 realistic industry use.

109 To further advance the field, this paper reviews the progress in the mathematical  
110 modeling of CHO cells. For clarity, the comprehensive work to be produced to ob-  
111 tain an efficient mechanistic model in bioprocesses is divided into a set of building  
112 blocks (Fig. 1).

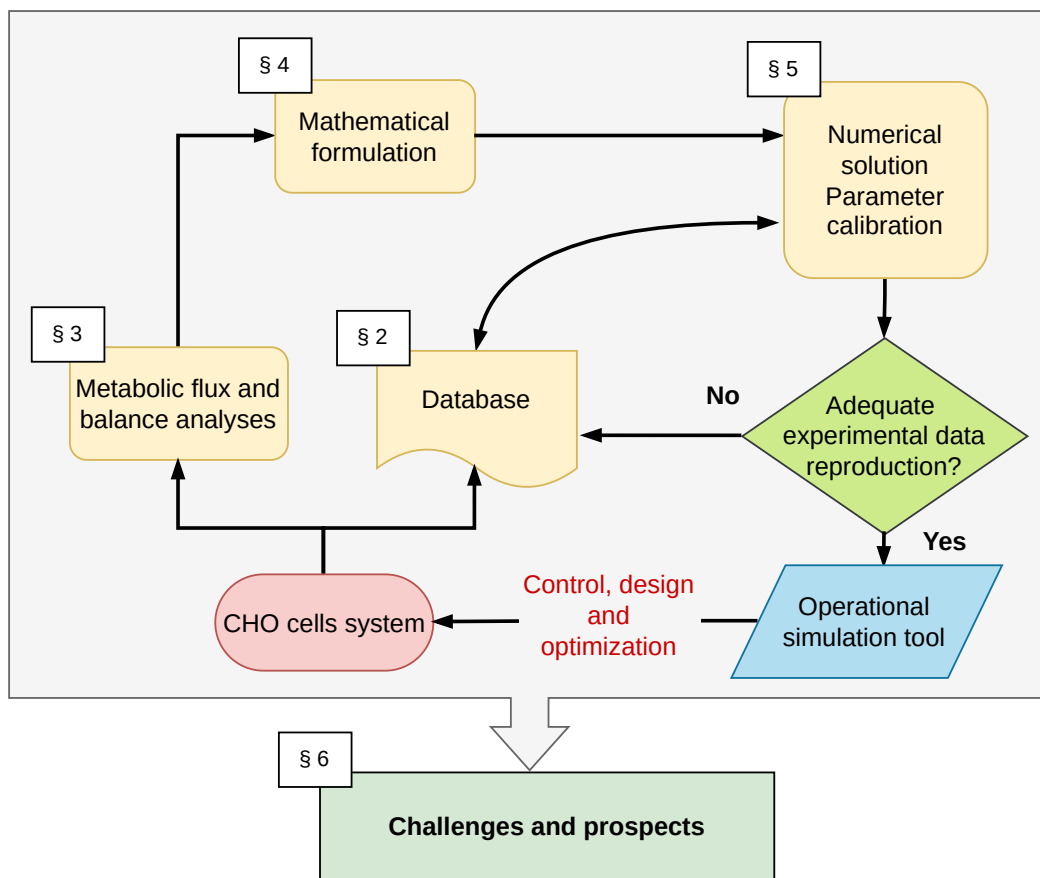


Fig. 1: Strategy for modeling CHO cell metabolism.

113 The paper is organized as follows. It is first necessary to understand the biological  
114 system (substrate, culture media, operating parameters, metabolic pathways, etc.).  
115 This should be done at the bioreactor level, at which cultures should be performed  
116 over a wide range of growing conditions with comprehensive instrumentation (online

117 and offline) to build a database (section 2). At the fundamental level, metabolic  
118 pathways need to be described, using metabolic flux analysis and flux balance anal-  
119 ysis (section 3). The metabolic pathways are then formulated to build a system  
120 of ordinary differential equations (ODEs) (section 4). With relevant applied math-  
121 ematical tools, the set of equations can finally be solved efficiently for predictive  
122 simulation (section 5). At this stage, the database of section 2 is crucial to define  
123 the model parameters by direct determination or inverse analysis. Once validated,  
124 the mechanistic model can be used as an operational control command, design, and  
125 optimization tool.

126 A final section 6 is devoted to the remaining challenges and prospects. After some  
127 application examples of mechanistic modeling to CHO cell systems, the main open  
128 questions and remaining challenges will be exposed. Then, as part of the answers  
129 to these challenges, we explain how the combination of mechanistic modeling and  
130 machine learning, together with database and online information, is about to change  
131 the vision of the sequential loop depicted in Fig. 1. Finally, recommendations are  
132 provided to the reader as a guide for future studies.

## 133 **2. CHO cells system**

### 134 *2.1. CHO culture conditions*

135 The production of recombinant proteins using CHO cells requires optimal culture  
136 media that support cell growth and productive yield, providing the necessary re-  
137 sources for high viable cell densities, stimulating synthesis processes, and extracellu-  
138 lar transport of biological products (Ritacco et al., 2018). Currently, several media  
139 that provide good yields for CHO cells (high density of viable cells, cell longevity,  
140 and increased product titers) in therapeutic protein production are commercially  
141 available (Pan et al., 2017b). The optimal composition of a basal medium for pro-  
142 tein production depends mainly on the CHO cells strain, specific characteristics of



143 the subclones, and the protein of interest (Rodrigues et al., 2012; Reinhart et al.,  
144 2015; Pan et al., 2017b).

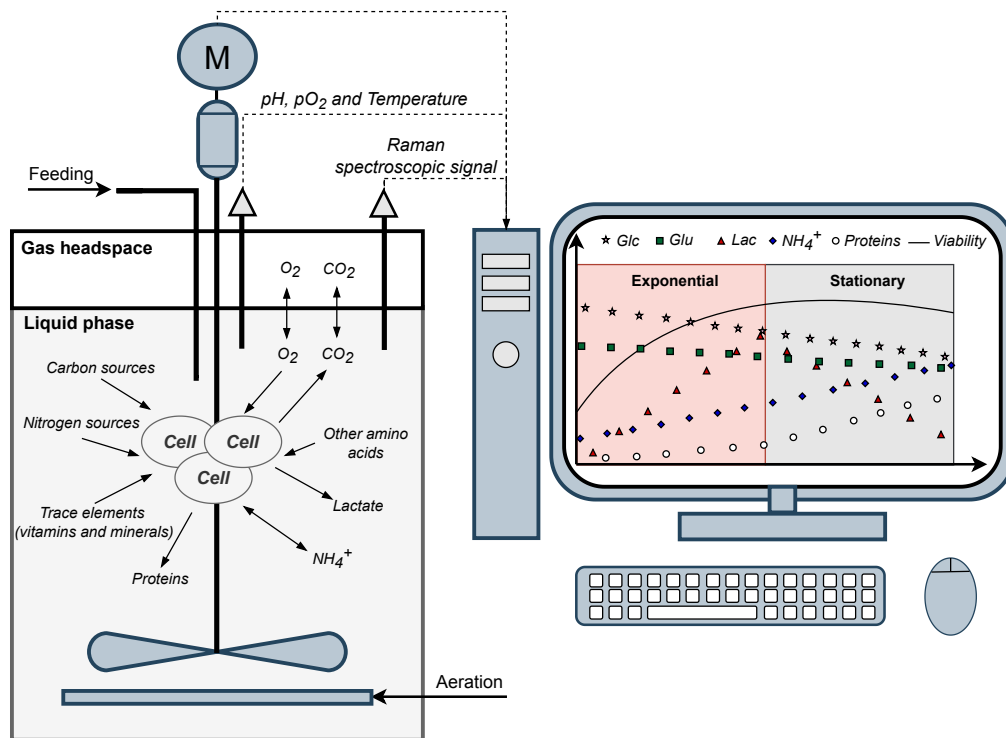


Fig. 2: Schematic representation of the bioreactor functioning during CHO cell culture (Glc: Glucose, Glu: Glutamate, and Lac: Lactate).

145 Whatever the CHO strain, the culture medium must be rich in carbon sources (sugar,  
146 mostly glucose), nitrogen sources (mainly glutamate or glutamine, and other amino  
147 acids, especially those that cannot be synthesized de novo by the CHO cells and  
148 known as essentials: histidine, phenylalanine, leucine, isoleucine, lysine, methionine,  
149 threonine, tryptophan, arginine, cysteine, proline, and valine (Carrillo-Cocom et al.,  
150 2015; Hefzi et al., 2016; Pan et al., 2017a,b)), and trace elements (vitamins and  
151 minerals), under aerobic conditions (Huang et al., 2017). This system is usually  
152 operated in fed-batch mode to maintain optimal nutrient concentration values and  
153 ensure maximum protein production (Fig. 2). In recent years, perfusion reactors  
154 have attracted a great deal of interest due to the advantages they offer over batch-  
155 feed systems: higher viable cell densities for extended periods of time, resulting

156 in increased volumetric titers and a smaller footprint ([MacDonald et al., 2022](#)).  
157 CHO cell metabolism is highly complex and divided into two fundamental phases.  
158 The first stage is exponential growth and low protein production, followed by a  
159 stationary phase, characterized by low cell growth and strong protein production  
160 using the lactate produced during the exponential phase as substrate. A final phase,  
161 the decline phase with lower production and higher mortality, could occur, but the  
162 industrial process usually stops before this decline phase. Lactate and ammonium  
163 are the most significant inhibitory by-products of CHO cell metabolism, but an  
164 excess of specific metabolites, including amino acids (e.g., phenylalanine, leucine,  
165 threonine, tryptophan, tyrosine, serine, methionine, etc.) and by-products (e.g.,  
166 formate, indolelactate, homocysteine, phenylacetate, etc.), can also impact CHO  
167 cell metabolism negatively ([Pereira et al., 2018](#)).

168 The stationary phase can be manually induced by a temperature shift or triggered  
169 spontaneously due to the accumulation of inhibiting secondary metabolites for cell  
170 metabolism, primarily lactate and ammonium. The temperature strategy is timed  
171 to increase longevity and improve final antibody yield ([Torres et al., 2018](#); [McHugh  
172 et al., 2020](#)).

173 Due to the great complexity of these systems, many studies are currently being  
174 carried out to optimize their operation by reducing the production of inhibiting  
175 secondary metabolites for cellular metabolism and increasing the yield in protein  
176 production. Note that experiments in these systems are costly, time-consuming,  
177 and prone to bacterial contamination. With recent developments in computing  
178 capabilities, mathematical modeling has become an undeniable ally in bioprocess  
179 research, enabling the validation of various hypotheses with considerable savings in  
180 resources and time.

## 181 2.2. Gene amplification systems

182 Stable CHO cell lines for recombinant protein production are obtained using two  
183 main gene amplification systems: dihydrofolate reductase (DHFR)-based methotrex-  
184 ate (MTX) selection or glutamine synthetase (GS)-based methionine sulfoximine  
185 (MSX) selection (Matasci et al., 2008; Costa et al., 2010; Fan et al., 2012; Budge  
186 et al., 2021; Yang et al., 2022).

### 187 2.2.1. DHFR-based MTX selection

188 The primary function of the DHFR enzyme is to catalyze the production of tetrahy-  
189 drofolate from folic acid (Costa et al., 2010), a process involved in the biosynthesis  
190 of glycine, purines, and thymidylic acid (Cacciatore et al., 2010; Budge et al., 2021;  
191 Yang et al., 2022). The DHFR gene is transfected into host cells in the same gene  
192 expression vector as the protein of interest, serving as a marker to select cells trans-  
193 fected with the protein gene of interest in a medium deficient in glycine, purines, and  
194 thymidylic acid (Noh et al., 2013). MTX, a DHFR inhibitor, is also used to create  
195 more pressure in the selection process until only cells with an elevated gene copy  
196 number prevail (Noh et al., 2013). Although this selection system has been the most  
197 widespread, as it allows for greater efficiency in gene amplification, it requires many  
198 rounds of selection involving considerable time consumption (Noh et al., 2018).

199 In modeling, this amplification method should be considered when working with  
200 systems with deficient glycine, purines, and thymidylic acid. However, these nutri-  
201 ents are usually fulfilled in the production process, and their consumption is less  
202 extensive than other nutrients like glucose, glutamate, and glutamine.

### 203 2.2.2. GS-based MSX selection

204 The primary function of the enzyme glutamine synthetase is to catalyze the syn-  
205 thesis of glutamine from glutamate and ammonium (Cacciatore et al., 2010; Yang  
206 et al., 2022). This amplification method is adapted to cells that do not survive in  
207 glutamine-poor media. Then, the GS gene is transfected into host cells in the same

208 expression vector as the protein gene of interest, serving as a marker for selecting  
209 cells transfected with the protein gene of interest in the glutamine-deficient medium  
210 (Noh et al., 2013). MSX, an enzyme inhibitor, is then used to increase the selec-  
211 tion pressure until only cells with elevated gene copy numbers prevail (Zhang et al.,  
212 2022). The GS system can achieve adequate expression levels through a single round  
213 of selection and amplification, thus significantly reducing the time required for cell  
214 line generation (Kingston et al., 2002; Noh et al., 2013). The shorter amplification  
215 time needed and its contribution to ammonium reduction by converting glutamate  
216 and ammonium to glutamine have increased its application in the pharmaceutical  
217 industry (Fan et al., 2013).

218 Unlike the previous amplification method, the GS-based MSX selection method has  
219 a significant impact during the production process. This method endows the cell  
220 with the ability to synthesize glutamine from glutamate, ammonium, and ATP,  
221 helping to counteract the inhibition of ammonium in CHO cell metabolism (Fan  
222 et al., 2013). In the case of glutamate limitation, the use of glutamine has a more  
223 pronounced negative effect on the cell than glutamate utilization, as it results in an  
224 increased release of ammonium (Dang, 2010). Therefore, when modeling a system  
225 in which this selection method has been applied, it is crucial to consider both the  
226 synthesis and utilization of glutamine. These processes involve extensively consumed  
227 substrates (glutamate and glutamine) and an inhibiting metabolite (ammonium),  
228 significantly impacting cell metabolism during protein production.

### 229 *2.3. Measured data*

230 In CHO systems, the experimental determination of variables such as total and  
231 viable cell number (or viability), offline pH, partial pressures of oxygen ( $pO_2$ ) and  
232 carbon dioxide ( $pCO_2$ ), osmolality, glucose, lactate, amino acids, ammonia, and  
233 mAbs concentrations are common (Yahia et al., 2021; Huang et al., 2017; Xing  
234 et al., 2010; Nolan and Lee, 2011, 2012). Determining these parameters during the

235 process enables its performance to be assessed and the operating conditions to be  
236 adjusted by manipulating parameters such as the feeding of substrates or media. In  
237 most cases, the sampling and subsequent analytical techniques used to determine  
238 the variables require considerable resources and time. This is a lock that prevents  
239 the operator from reacting in time to ensure optimum performance.

240 Recently, probes have been developed that allow online monitoring of most variables,  
241 as is the case with probes designed using spectroscopy. Raman spectroscopy is one of  
242 the most widely used techniques for online monitoring and control of cell cultures.  
243 It stands out for its spectra's sharpness and compatibility with aqueous systems  
244 (Li et al., 2018). Several studies have reported the use of Raman spectroscopy for  
245 online monitoring in CHO systems in the last few years (Li et al., 2018; Santos  
246 et al., 2018; Feidl et al., 2019; Yilmaz et al., 2020; W Eyster et al., 2021; Chen  
247 et al., 2021; A Gibbons et al., 2022; Domján et al., 2022; Schwarz et al., 2022;  
248 Romann et al., 2022; Yousefi-Darani et al., 2022; Yang et al., 2024). The online  
249 monitoring of key process variables facilitates the development of mathematical  
250 feedback models where operating conditions can be adjusted to obtain the maximum  
251 yield in protein production. Mechanistic modeling and Raman spectroscopy have  
252 been combined for monitoring antibody chromatographic purification by Feidl et al.  
253 (2019). This monitoring was achieved by combining data from a kinetic model and  
254 a Raman analyzer employing an extended Kalman filter. This technique proved  
255 robust, allowing accurate estimation of antibody concentrations with reduced noise.

256 The richness of the database that will be used later for model calibration is of crucial  
257 importance, namely regarding the robustness and prediction potential. To optimize  
258 the experimental work needed to build this database, it is strongly recommended to  
259 use a design of experiment (DOE) that considers the variability of the main system  
260 variables, within and outside the operational ranges used in industry. Both fed-  
261 batch and batch tests are recommended. Notably, one part of the tests should not

262 be used in the learning step but kept for validation.

### 263 **3. Metabolic pathway identification**

264 In bioprocess modeling, identifying metabolic pathways is crucial as it represents the  
265 core of mechanistic modeling and its formulation. To that purpose, metabolic flux  
266 analysis (MFA) is the most widely used technique. MFA uses *in vivo* isotopic mark-  
267 ers of metabolites ( $^{13}\text{C}$ ) and modeling to quantify fluxes through the major metabolic  
268 pathways at steady state (Sengupta et al., 2011). In recent years, MFA has been  
269 extended to assess metabolic transients during fermentation, considering cellular dy-  
270 namics, referred to as dynamic MFA (DMFA) (Antoniewicz, 2013; Martínez et al.,  
271 2015). Conventional and dynamic MFA approaches can be complemented by flux  
272 balance analysis (FBA), a mathematical modeling method frequently employed by  
273 metabolic engineers to quantitatively simulate steady-state genome-scale metabolic  
274 reconstructions (Kauffman et al., 2003; Gianchandani et al., 2010; Martínez et al.,  
275 2013; Ivarsson et al., 2015; Hefzi et al., 2016; Huang et al., 2017; Gutierrez et al.,  
276 2020; Schinn et al., 2021), in which, as for DMFA, dynamic conditions can also be ap-  
277 plied (DFBA) (Antoniewicz, 2013). FBA utilizes linear programming to optimize a  
278 flux distribution towards a defined objective, considering physicochemical and ther-  
279 modynamic constraints (Orth et al., 2010). While FBA demands significantly fewer  
280 experimental resources than the extensive requirements of  $^{13}\text{C}$ -MFA, the latter is  
281 the preferred method for identifying metabolic pathways within biological systems.  
282 It offers higher precision and the ability to generate a more comprehensive flux  
283 map (Antoniewicz, 2021). Consequently, this study will emphasize the  $^{13}\text{C}$ -MFA  
284 approach.

#### 285 *3.1. Metabolic flux analyses*

286 Several recent MFA approaches have been developed to understand CHO cell metabolism.  
287 In most cases, two distinct metabolic phases can be observed. The first is the ex-

288 potential phase, characterized by rapid growth and the secretion of inhibitory by-  
 289 products such as lactate and ammonium. During this phase, protein production is  
 290 limited. It is followed by the stationary or non-growth phase, during which protein  
 291 production increases significantly and cell growth decreases. During the exponential  
 292 phase, a high flux of glycolysis with considerable lactate production and a strong  
 293 association of anaplerotic processes with the TCA cycle is observed. In contrast,  
 294 the stationary or non-growth phase exhibits a reduced glycolysis flux, a net lactate  
 295 consumption, an oxidative flux of the pentose phosphate pathway, and a reduced  
 296 rate of anaplerosis (Ahn and Antoniewicz, 2011; Templeton et al., 2013).

297 Fig. 3 summarizes the metabolic pathways involving carbon-containing compounds  
 298 that play a crucial role in catabolic and anabolic processes in CHO cells (known as  
 299 central metabolism carbon) determined by using  $^{13}\text{C}$  as reported in the literature.

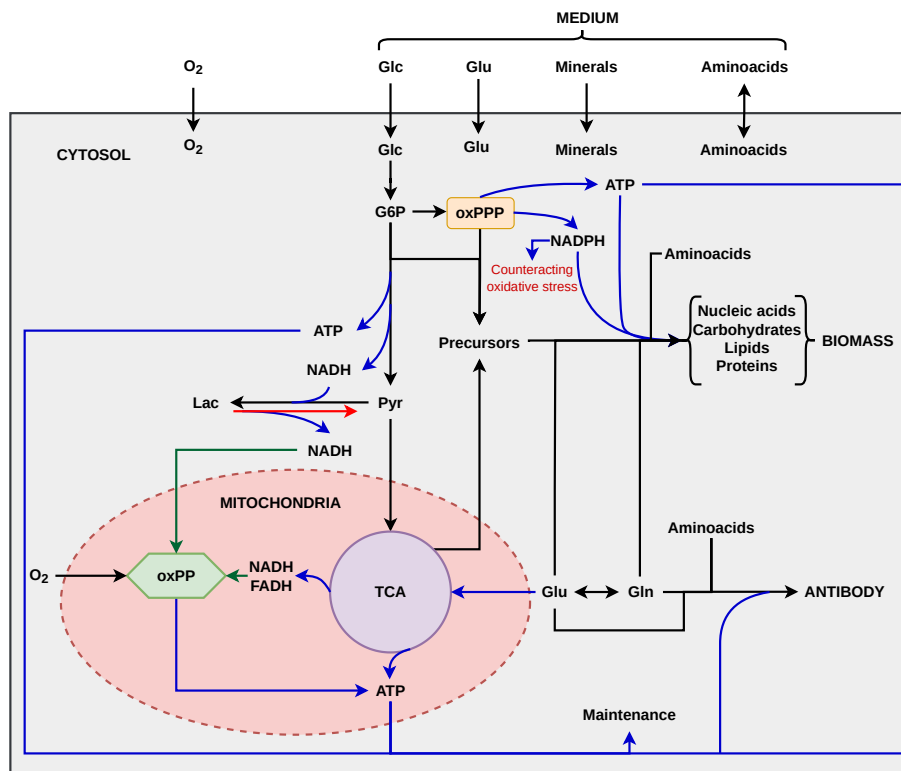


Fig. 3: Simplified schema of central metabolism carbon in CHO cells based on literature reports (Glc: Glucose, Glu: Glutamate, G6P: Glucose-6-phosphate, Pyr: Pyruvate, Lac: Lactate, Gln: Glutamine, oxPPP: Oxidative pentose phosphate pathway, TCA: Tricarboxylic acid cycle, and oxPP: Oxidative phosphorylation).

300 Anaplerosis is mainly observed through converting pyruvate to oxaloacetate and  
301 glutamate to  $\alpha$ -ketoglutarate. In both phases, pyruvate dehydrogenase, and TCA  
302 cycle fluxes are similar ([Ahn and Antoniewicz, 2011](#)). During the stationary phase,  
303 an elevated glucose flux diverted to oxPPP, providing high NADPH production,  
304 is observed ([Sengupta et al., 2011](#)). This elevated NADPH production via oxPPP  
305 could be associated with macromolecule biosynthesis or a defensive process of the  
306 cell to counteract oxidative stress (Fig. 3). [Sengupta et al. \(2011\)](#) also found that,  
307 during the stationary phase, most of the pyruvate produced in glycolysis was me-  
308 tabolized in the TCA cycle with little or no lactate production. This finding aligns  
309 with [Templeton et al. \(2013\)](#), who developed a more detailed MFA, observing a cor-  
310 relation between the peak of antibody production and the highest oxidative activity  
311 in the Krebs cycle. The authors confirmed the absence of lactate production dur-  
312 ing the stationary phase until the metabolite reached low levels, restarting lactate  
313 production. Interestingly, during the stationary phase, the energy efficiency of the  
314 cells using lactate (total ATP produced per total C-mol substrate consumed) six  
315 times higher than in lactate-producing cells was reported by [Martínez et al. \(2013\)](#)  
316 by flux balance analysis. These results underline the importance of the exponen-  
317 tial phase, during which rapid growth ensures a large number of cells or "protein  
318 micro-factories," simultaneously providing lactate, an efficient energy source for the  
319 stationary phase.

320 [Martínez et al. \(2015\)](#) conducted an interesting DMFA work for studying the dy-  
321 namics of metabolic shifts caused by temperature changes. Their study demon-  
322 strated that inducing mild hypothermia in the system significantly reduces growth  
323 and overall metabolic rates, potentially improving the stability of recombinant pro-  
324 tein production. These results suggest that protein production is antagonistic to the  
325 growth process. As the cell reduces its metabolic flux for cell growth due to temper-  
326 ature decrease, its biological activity is redirected towards protein production. This



327 behavior is especially relevant considering these cells have been engineered, selected,  
 328 and adapted to produce recombinant proteins, behaving like malignant cells to grow  
 329 indefinitely. Cell decay is also an important process that should not be neglected in  
 330 these systems. Being present throughout the process is more significant during the  
 331 decline phase. The specific mortality averaged over the whole process rate ranges  
 332 between 0.013 and 0.107 d<sup>-1</sup> (Templeton et al., 2013). In brief, most MFA stud-  
 333 ies report strong cell growth and weak protein production during the exponential  
 334 phase mainly based on glucose and glutamate consumption, which is followed by a  
 335 transition of metabolism to the stationary phase, where strong protein production  
 336 and weak cell growth are based mainly on the simultaneous utilization of glutamate,  
 337 lactate, and glucose, with low or no lactate production. In general, peak antibody  
 338 production is associated with increased oxidative metabolism activity. Cell decay  
 339 has been reported during the entire process. Fig. 4 shows a compilation of the  
 340 phenomena occurring during CHO cell metabolism identified from the analysis of  
 341 metabolic fluxes reported in the literature.

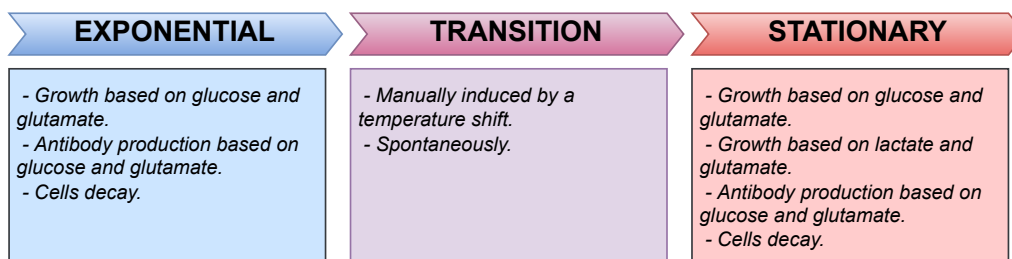


Fig. 4: Metabolic pathways involved in the different stages of CHO cell metabolism.

342 The exponential phase is accompanied by the production of by-products, such as lac-  
 343 tate and ammonia, which inhibit cell metabolism. However, the lactate production  
 344 during the exponential phase ensures the redox balance in the cytosol. It becomes  
 345 an essential carbon and energy source during the stationary phase (Fig. 3).

346 3.2. Metabolic pathways description

347 The metabolic pathways identified from the metabolic flux analysis works reported  
 348 in the previous section are summarized in figures 5 and 6, respectively, for the  
 349 exponential and stationary phases. The extent of each metabolic pathway identified  
 350 may depend on the CHO strain.

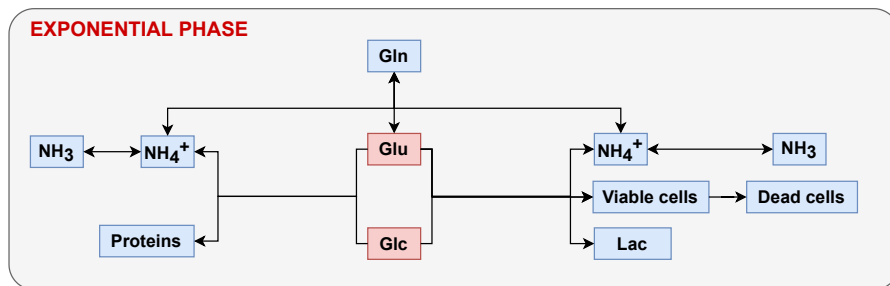


Fig. 5: Simplified schema of metabolic pathways taking place during the exponential phase (Glc: Glucose, Glu: Glutamate, Gln: Glutamine, and Lac: Lactate, with pink and blue colors denoting substrates and products, respectively).

351 The modeling of the stationary phase is more complex than the exponential phase,  
 352 as it involves a more significant number of metabolic pathways co-occurring (e.g.,  
 353 protein production based on lactate, cell growth based on lactate, protein produc-  
 354 tion based on glucose, and cell growth based on glucose). In this phase, lactate is  
 355 an energy source primarily for cell growth. While lactate plays a role in protein  
 356 production, its contribution is minimal compared to glucose. Consequently, some  
 357 authors have dismissed the significance of lactate in protein production ([Jimenez del](#)  
 358 [Val et al., 2016](#); [Kotidis et al., 2019](#)). However, its minor contribution is crucial when  
 359 the model distinguishes anabolism from catabolism by accounting for the ATP/ADP  
 360 and NADH/NAD<sup>+</sup> electron transport chains ([Nolan and Lee, 2011](#)).

361 This remark on the role of lactate illustrates how the choice of metabolic pathways  
 362 to be considered is a crucial question for modeling. The answer depends on the costs  
 363 and benefits between their influence on the main output variables and the complexity  
 364 of the parameter determination. Additionally, the number of stoichiometric and

365 dynamic parameters to be defined depends on the model complexity, which could  
 366 be problematic for direct experimental determination. As it will be explained below,  
 367 this question of parameter determination can be partly addressed by inverse analysis.  
 368 Note that a more comprehensive model does not necessarily mean greater precision.  
 369 A complexity that is too high and unbalanced with sufficient data can hinder its  
 370 practical application (numerical solution problems, over-parametrized model, error  
 371 propagation, large CPU time, etc.). Therefore, the challenge for modelers is to  
 372 obtain a suitable trade-off.

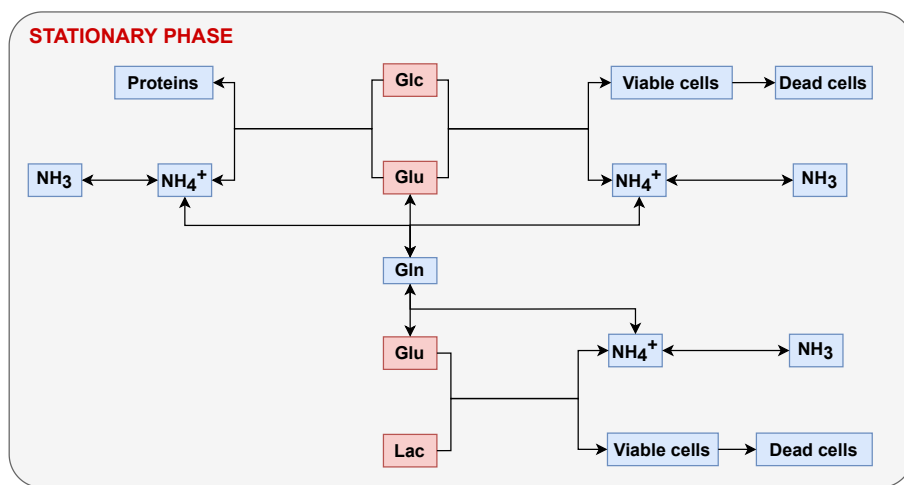


Fig. 6: Simplified schema of metabolic pathways occurring during the stationary phase (Glc: Glucose, Glu: Glutamate, Gln: Glutamine, and Lac: Lactate, with pink and blue colors denoting substrates and products, respectively).

373 Moreover, although both phases have common metabolic pathways, they differ con-  
 374 siderably in stoichiometry and kinetics. Consequently, both phases must be sep-  
 375 arately formulated using submodels. These two submodels can coexist, but the  
 376 pathways of the stationary phase submodel (Fig. 6) should be activated once the  
 377 metabolic pathways of the exponential phase submodel (Fig. 5) have been deacti-  
 378 vated.

### 379 3.3. Metabolic shift

380 The metabolic transition of CHO from the exponential to the stationary phase is  
 381 usually intentionally induced by a temperature shift (temperature drop) (López-

382 [Meza et al., 2016](#); [Jimenez del Val et al., 2016](#)). Nevertheless, factors that trigger a  
 383 spontaneous metabolic shift remain a challenge for modelers, as it is crucial to obtain  
 384 a predictive model that can address unexpected situations or propose innovative  
 385 protocols. The metabolic transition from lactate production to lactate consumption  
 386 in CHO cells is an essential phenomenon strongly linked to cell culture longevity  
 387 and protein production yield ([Brunner et al., 2018](#)). However, this metabolic shift is  
 388 difficult to control due to the unknown mechanisms involved, which are still under  
 389 investigation ([Zagari et al., 2013](#); [Hartley et al., 2018](#); [Hong et al., 2018](#)).

390 Several studies have been conducted to identify the factors or the combination of  
 391 factors that trigger this phenomenon. The literature has highlighted various major  
 392 triggering factors (Fig. 7): limiting substrates such as glucose ([Altamirano et al.,](#)  
 393 [2004](#); [Tsao et al., 2005](#); [Altamirano et al., 2006](#); [Martínez et al., 2013](#); [Zagari et al.,](#)  
 394 [2013](#); [Brunner et al., 2021](#)), glutamate, or glutamine ([Zagari et al., 2013](#); [Ghorba-](#)  
 395 [niaghdam et al., 2014](#); [Wahrheit et al., 2014](#)), the effect of pH ([Zalai et al., 2015](#);  
 396 [Liste-Calleja et al., 2015](#); [Ivarsson et al., 2015](#)), and inhibitory by-products such as  
 397 lactate ([Mulukutla et al., 2012](#); [Kyriakopoulos and Kontoravdi, 2014](#); [Pereira et al.,](#)  
 398 [2018](#)), and carbon dioxide ([Brunner et al., 2018](#); [Xu et al., 2018](#)).

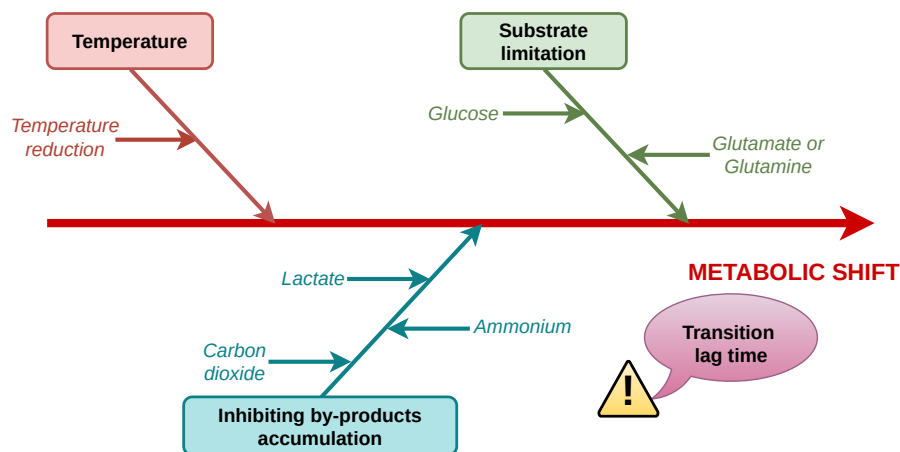


Fig. 7: Variables proposed in the literature as possible triggers for metabolic shift in the CHO cell system.

399 Substrate limitation (glucose, glutamate, or glutamine) can activate alternative  
400 metabolic pathways in which the carbon source is alternated from glucose to lac-  
401 tate. Many studies have reported that spontaneous metabolic shift is associated  
402 with pH (Zalai et al., 2015; Liste-Calleja et al., 2015; Ivarsson et al., 2015), which  
403 contradicts other studies in which metabolic shift has occurred spontaneously at  
404 constant pH (Altamirano et al., 2006; Mulukutla et al., 2012; Martínez et al., 2013;  
405 Ghorbaniaghdam et al., 2014). Martínez-Monge et al. (2019) tested pH-controlled  
406 and uncontrolled CHO cell metabolism. When pH was controlled in the bioreac-  
407 tor, only lactate consumption was observed once glucose was completely depleted.  
408 However, since these experiments were performed in batch mode, it can be assumed  
409 that the metabolic shift was triggered by substrate limitation, in this case, glucose  
410 limitation. It would be interesting to perform this experiment with sufficient glucose  
411 (batch feeding). In contrast, when pH was not controlled, it decreased due to the  
412 accumulation of secreted lactic acid, triggering its consumption when pH was below  
413 6.80. According to these authors, the metabolic shift could be induced by adding  
414 lactate to the initial medium and setting the pH below 6.80.

415 Finally, the antagonism between cell growth and protein production reported in sec-  
416 tion 3.1 could explain the spontaneous metabolic shift. For example, inhibiting cell  
417 growth due to the accumulation of by-products could lead the cell to redirect their  
418 production capacity towards protein synthesis (Mulukutla et al., 2012; Kyriakopou-  
419 los and Kontoravdi, 2014; Pereira et al., 2018). This redirection could be a way to  
420 maintain maximum biological activity of the cell, consistent with its malignant be-  
421 havior, while conveniently consuming an essential cell growth inhibitor by-product  
422 such as lactate.

423 In conclusion, according to the literature analysis, a combination of temperature and  
424 limiting substrates triggers the metabolic shift. On the other hand, the spontaneous  
425 metabolic shift occurs even with controlled and uncontrolled pH, implying that pH

426 could be an indirect macroscopic indicator of the metabolic shift. This parameter  
 427 is directly related to inhibiting by-products already identified as triggering factors,  
 428 like lactate and carbon dioxide. Another factor to consider is ammonium, which  
 429 is an inhibitory by-product of cell metabolism that, at the same time, contributes  
 430 to lowering the pH of the system. To our knowledge, ammonium has not yet been  
 431 reported in the literature as a triggering factor. For this reason, it would be in-  
 432 teresting to intentionally try to induce metabolic shifts by adding these inhibiting  
 433 by-products to discern their contribution to metabolic shifts.

#### 434 3.4. Main acid contributors to pH

435 To quantify the pH variation in the CHO system, it is necessary to consider the  
 436 main metabolites, the substrate feed, and the acid-base solutions added during  
 437 metabolism to regulate pH, where the main acid-base contributors must be identi-  
 438 fied. The reduction of pH during the degradation process is a direct consequence of  
 439 CHO cell metabolism. All metabolic pathways are accompanied by carbon dioxide  
 440 production and other inhibiting by-products such as lactate and ammonium that  
 441 shift equilibrium towards hydronium formation, thus reducing the pH value. The  
 442 chemical equilibria corresponding to the substrates and metabolites involved in CHO  
 443 cell metabolism are shown in Table 1.

Table 1: Main acid contributors to pH variation

No.	pH contributor	Chemical equilibrium	$K_a$ at 25 °C
1	Ammonium	$NH_4^+ + H_2O \rightleftharpoons NH_3 + H_3O^+$	$5.60 \times 10^{-10}$
2	Lactate	$HLac + H_2O \rightleftharpoons Lac^- + H_3O^+$	$1.38 \times 10^{-4}$
		$H_2CO_3 = CO_2 + H_2O$	
3	Carbon dioxide	$H_2CO_3 + H_2O \rightleftharpoons HCO_3^- + H_3O^+$	$4.20 \times 10^{-7}$
		$HCO_3^- + H_2O \rightleftharpoons CO_3^{2-} + H_3O^+$	$4.80 \times 10^{-11}$

444 Considering the equilibrium constant values at 25 °C and the same species concen-  
 445 tration, the order of contribution to the system acidity is  $HLac > NH_4^+ > CO_2$ .  
 446 However, the actual acid contribution depends on the extent to which each species

447 it is available in the system. One must remember that lactate is produced dur-  
448 ing the exponential phase and consumed during the stationary phase. Meanwhile,  
449 ammonium and CO<sub>2</sub> are produced throughout the process. Nevertheless, as pH is  
450 generally regulated, the influence of pH is discarded in most metabolic models. In  
451 contrast, it is important to note hyperosmolality, a non-physiological increase in  
452 osmolality that affects cell physiology, morphology, and proteome, resulting from  
453 the addition of concentrated feed and base solutions for pH regulation. While it  
454 enhances the specific antibody production rate, it inhibits cell growth, thereby af-  
455 fecting the final antibody titers (Kim et al., 2002; Min Lee and Koo, 2009; Zhang  
456 et al., 2010; Romanova et al., 2022). Incorporating hyperosmolality into mechanistic  
457 metabolic models could improve the accuracy of the impact of feeding on CHO cell  
458 metabolism.

#### 459 **4. Mathematical formulation**

460 The formulation is a key building block to derive a mechanistic model. Even though  
461 they are not mandatory, the following assumptions are considered in most published  
462 works:

- 463 1. Perfectly stirred bioreactor,
- 464 2. Constant stoichiometric and kinetics parameter values,
- 465 3. Metabolic shift can be induced by temperature change and the combination  
466 of ammonium and lactate inhibiting effect,
- 467 4. Different stoichiometry and kinetics for exponential and stationary phases,  
468 even when they can change over the metabolic process, as can be corroborated  
469 in Templeton et al. (2013),
- 470 5. Constant chemical composition of CHO cells, regardless of the metabolic pro-  
471 cess and the nature of the substrate. It is to be expected that the chemical

472 composition of CHO cells will be different when glucose or lactate is used as  
473 the primary carbon source,

474 6. No oxygen limitation is considered, as aeration is regulated to ensure adequate  
475 oxygen levels during the process to avoid the reactive oxygen species formation  
476 that affects the cell's metabolism, reducing the proteins' productivity ([Hand-  
477 logten et al., 2018](#)) considerably,

478 7. pH regulation.

479 Assumption 1 allows the 0D model to be derived. They are based on ODEs (Or-  
480 dinary Differential Equations), which allow the efforts to be concentrated on the  
481 metabolism. Constant stoichiometric and kinetics parameter values are quasi-mandatory  
482 for tackling complexity. Yet, due to the existence of two phases and co-current path-  
483 ways, the global stoichiometry over an entire process depends on culture conditions.  
484 Assumption 3 is essential as it connects the submodel describing the exponential  
485 phase with the submodel describing the stationary phase. Consequently, successful  
486 identification and mathematical description of the combination of effects of the fac-  
487 tors that trigger this metabolic shift represents a significant challenge for modelers.  
488 Although not mandatory, the two last assumptions are usually fulfilled in produc-  
489 tion and allow the model to be simplified without restriction. Finally, it is essential  
490 to note that all the above assumptions are based on the condition that the CHO  
491 cell system is fully acclimatized to the culture medium and ready for production.

#### 492 *4.1. Kinetics of biological processes*

493 Models based on physical, chemical, and biological principles offer a robust option in  
494 process engineering. However, the rates at which these principles occur is mandatory  
495 to obtain a predictive model. This must include the dynamics of all metabolic  
496 pathways, but also side phenomena such as the viable/non-viable cell densities and  
497 balances in the compartments considered by the formulation: at least the cells



498 and the bioreactor (nutrient/metabolite concentrations) (Tsopanoglou and del Val,  
 499 2021) and, depending on the model, the concentration obtained from balances in  
 500 subcompartments of the cells, such as mitochondria. For a bioprocess operated in  
 501 fed-batch mode, the following matter balances reads as:

$$V \frac{dS_{(j)}}{dt} = Q_{in}(S_{in(j)} - S_{(j)}) + V \sum_{i=1}^{i=n} r_{S_{(i)}}, \quad (1)$$

$$V \frac{dX}{dt} = -Q_{in}X + V \sum_{i=1}^{i=n} r_{X_{(i)}}, \quad (2)$$

$$V \frac{dP_{(j)}}{dt} = -Q_{in}P_{(j)} + V \sum_{i=1}^{i=n} r_{P_{(i)}}, \quad (3)$$

$$\frac{dV}{dt} = Q_{in}, \quad (4)$$

502 where  $Q_{in}$  is the volumetric flow rate of feed (l/h),  $S_{(j)}$  is the limiting substrate  $-j$   
 503 (g/l),  $S_{in(j)}$  is the  $S_{(j)}$  concentration in the feed (g/l),  $P_{(j)}$  is the product concen-  
 504 tration  $-j$  (g/l),  $V$  is the bioreactor volume (l),  $X$  is the biomass concentration  
 505 (cells/l),  $r_{X_{(i)}}$  is the volumetric growth rate for process  $-i$  (cells/l/h),  $r_{P_{(i)}}$  is the  
 506 volumetric metabolites production rate for process  $-i$  (cells/l/h), and  $r_{S_{(i)}}$  is the  
 507 volumetric substrates consumption rate for process  $-i$  (cells/l/h) .

508 In general, the most widely accepted mathematical expressions for considering the  
 509 limiting effect and the inhibitory effect on the specific growth rates are Monod's  
 510 law and its variant, respectively, which are represented as a single expression as the  
 511 product of all limiting and inhibitory substrates (Bree et al., 1988; Xing et al., 2010;  
 512 Jimenez del Val et al., 2016; López-Meza et al., 2016; Kotidis et al., 2019; Yahia  
 513 et al., 2021; Tsopanoglou and del Val, 2021). The Monod variant mathematical  
 514 expression is also commonly used as a turnover rate function for changing metabolic  
 515 pathways. The first term of Eq. 6 represents the substrate's limiting effect, whereas

516 the by-products' inhibitory effect is represented in the second term of Eq. 6:

$$r_{X(i)} = \mu_{(i)}X, \quad (5)$$

$$\mu_{(i)} = \mu_{max(i)} \prod_{j=1}^{j=m} \frac{S_{(j)}}{K_{S(j)} + S_{(j)}} \prod_{j=1}^{j=k} \frac{K_{S(j)}^{inh}}{K_{S(j)}^{inh} + S_{(j)}^{inh}}, \quad (6)$$

517 where  $\mu_{(i)}$  is the specific growth rate for process  $-i$  ( $\text{h}^{-1}$ ),  $\mu_{max(i)}$  is the maximum  
 518 growth rate for process  $-i$  ( $\text{h}^{-1}$ ),  $S_{(j)}$  is the limiting substrate  $-j$  ( $\text{g/l}$ ),  $K_{S(j)}$  is the  
 519 half-saturation coefficient for the limiting substrate  $S_{(j)}$  ( $\text{g/l}$ ),  $S_{(j)}^{inh}$  is the inhibiting  
 520 by-product  $-j$  ( $\text{g/l}$ ) and  $K_{S(j)}^{inh}$  is the half-saturation coefficient for the inhibiting  
 521 by-product  $S_{(j)}^{inh}$  ( $\text{g/l}$ ).

522 The expression rates of the remaining model variables are expressed as a function  
 523 of cell growth rate considering the stoichiometric relationships between the different  
 524 chemical species and the cells formed during the biological process in a manner  
 525 analogous to chemical reactions. Negative and positive sign values are added to the  
 526 expression rate for substrate consumption (Eq. 7) and metabolite production (Eq.  
 527 8), respectively:

$$r_{S(i)} = -\frac{1}{Y_{X/S(j)}}r_{X(i)}, \quad (7)$$

$$r_{P(i)} = \frac{1}{Y_{X/P(j)}}r_{X(i)}, \quad (8)$$

528 where  $Y_{X/S(j)}$  is the  $X$  yield from  $S_{(j)}$  ( $\text{cells/g}$ ) and  $Y_{X/P(j)}$  is the  $X$  yield from  
 529  $P_{(j)}$  ( $\text{cells/g}$ ).

530 However, even when in CHO cell metabolism modeling, Monod's law and its inhi-  
 531 bition variant are the most used, it will be advisable to consider another kinetics  
 532 expression already used in bioprocess modeling to find an adequate description of  
 533 most of the variables' effect of the phenomena described. Fig. 8 shows examples

534 of kinetics expressions used in bioprocess modeling. The Moser Equation adds the  
535 parameter  $\lambda$  to Monod's law that describes the growth rate to the limiting substrate  
536 concentration. The Contois Equation states the proportionality between the effec-  
537 tive saturation constant and the biomass concentration  $X$ . At high  $X$ ,  $\mu$  is inversely  
538 proportional to  $X$ . This is sometimes used to represent a diffusion limitation in floc-  
539 culating or immobilized biomass (Snape et al., 2008). Expressions that combine the  
540 limiting and inhibitory effect for the same substrate, as is the case of the Haldane,  
541 Edwards, Webb, and Luong models, could be very useful in describing metabolic  
542 processes from substrates that have an inhibiting effect on the cell but can be utilized  
543 as a carbon source, as is the case for lactate in CHO systems. According to Edwards  
544 (1970), some mathematical expressions proposed to describe product inhibition can  
545 be borrowed to correlate substrate inhibition. For example, the Luong (1987) model  
546 was obtained from the Levenspiel (1980) model, while the Edwards (1970) model  
547 was derived from the Aiba et al. (1968) model. The Levenspiel (1980) and Aiba  
548 et al. (1968) models describe the inhibitory effect of the individual products. The  
549 logistic equation was initially proposed by the UK sociologist Thomas Malthus to  
550 describe "the law of population growth" at the end of the 18th century (Malthus,  
551 1986), which was later used by the Belgian mathematician Pierre Franois Verhulst  
552 (Verhulst, 1838) to describe the biological population kinetics, in particular the self-  
553 limiting growth (Xu, 2020). According to this expression, the specific growth rate  
554 decreases linearly with an increasing cell population ( $X$ ) and reaches zero at the  
555 maximum population. Indeed, this behavior is observed in most CHO systems. The  
556 use of this mathematical expression for modeling CHO cell metabolism has already  
557 been evaluated by Shirsat et al. (2015), obtaining better results than when Monod's  
558 law is used. In addition, the logistic equation has also been used to describe product  
559 inhibition (Contois, 1959; Fujimoto, 1963) (Fig. 8).

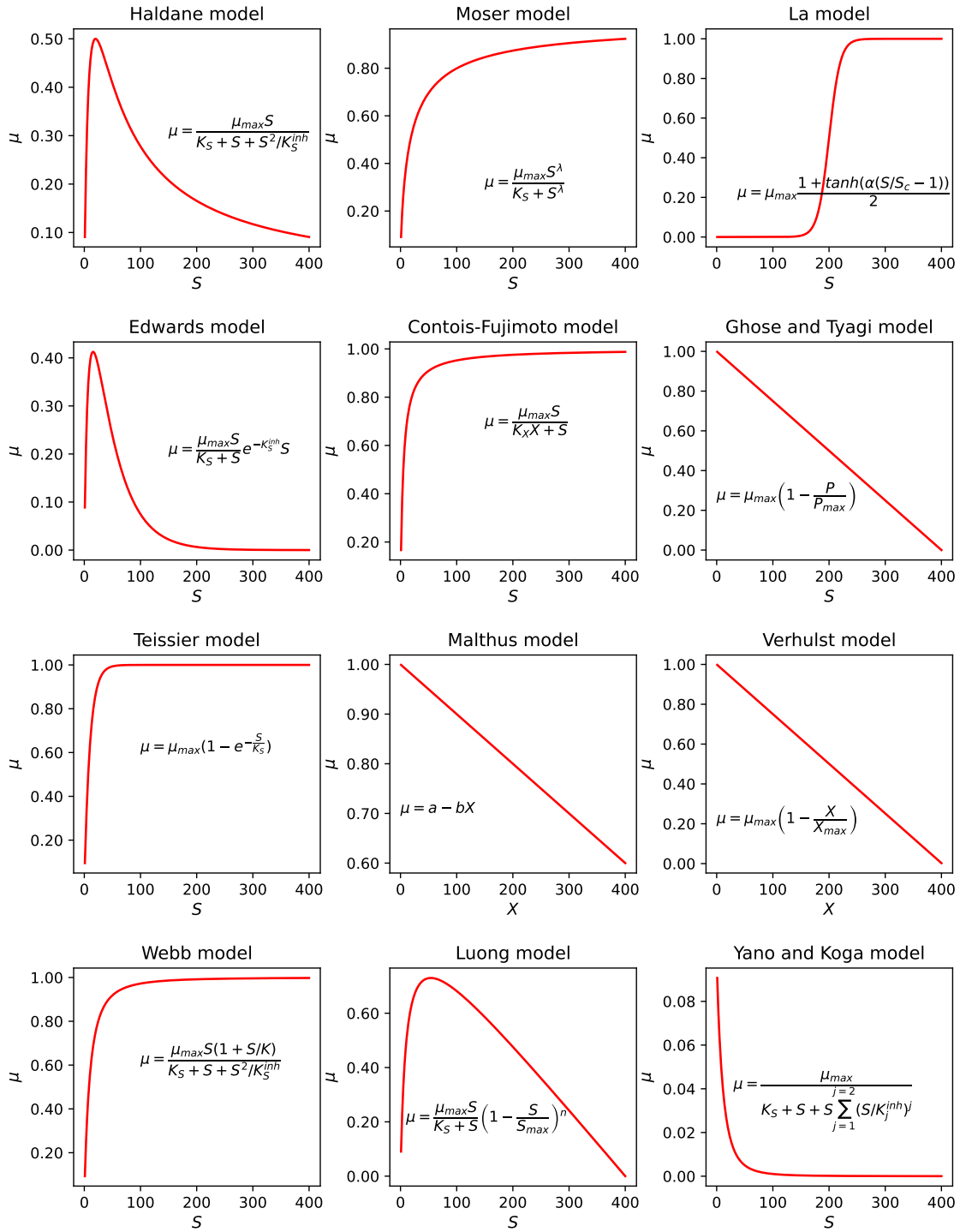


Fig. 8: Example of kinetics expressions used for modeling substrate limitation and by-products inhibition in bioprocesses (Verhulst, 1838; Teissier et al., 1936; Contois, 1959; Fujimoto, 1963; Webb, 1963; Yano and Koga, 1969; Edwards, 1970; Ghose and Tyagi, 1979; Malthus, 1986; Luong, 1987; Moser, 1988; Selișteanu et al., 2007; La et al., 2020).

560 However, even when this mathematical expression describes a similar behavior to  
561 that reported for CHO cell metabolism, it does not consider the metabolic shift. It  
562 is more related to the by-product inhibition accumulation. Nevertheless, the use of  
563 this mathematical expression is attractive because of the aspects mentioned above.  
564 The stepwise function proposed by [La et al. \(2020\)](#) is defined by two parameters:  
565 the shift value  $S_c$  defines the concentration value at which transition occurs, and the  
566  $\alpha$  parameter defines the sharpness of this transition. This powerful function can be  
567 used for describing switching between metabolic pathways, limiting substrate, and  
568 inhibiting by-product effects by manipulating the  $S_c$  and  $\alpha$  values. It allows the  
569 value at which switching occurs and the rate at which it occurs around that value  
570 to be set independently. In this article, we have selected a set of functions that  
571 can tackle most situations. To go further, we recommend the review proposed by  
572 [Mulchandani and Luong \(1989\)](#), where the applicability and limitations of several  
573 functions are analyzed.

#### 574 4.2. Kinetics of glutamine metabolism

575 The utilization and synthesis of glutamine can be represented as a chemical equilib-  
576 rium, where the glutamine substrate is in equilibrium with ammonia and glutamate  
577 (Eq. 9):



578 The equilibrium constant is then expressed as:

$$K_{Gln} = \frac{[\text{Gln}][\text{ADP}]}{[\text{NH}_4^+][\text{ATP}][\text{Glu}]} \quad (10)$$

579 As can be corroborated with Eq. 9 the glutamine formation from glutamate and  
580 ammonia requires energy in the form of ATP. Thus, the glutamine production rate

581 can be expressed as follows:

$$\frac{dGln}{dt} = k_{Gln}(K_{Gln}[NH_4^+][ATP][Glu] - [Gln][ADP]), \quad (11)$$

582 where  $k_{Gln}$  is the specific equilibrium displacement rate ( $s^{-1}$ ). [Batstone et al. \(2002\)](#)  
 583 suggested setting the specific equilibrium displacement rate one order of magnitude  
 584 higher than the highest biological rate constant to reduce model stiffness. The  
 585 reaction equilibrium markedly favors synthesis; the equilibrium constant at pH 7.0  
 586 and 37 °C was 1200 ([Levintow and Meister, 1954](#)).

### 587 4.3. Kinetics of physical and chemical processes

588 For considering pH prediction in the model, the hydronium concentration must be  
 589 estimated:

$$pH = -\log[H_3O^+] \quad (12)$$

590 To estimate hydronium concentration, the main acid-base compounds involved in  
 591 the CHO cell metabolism should be considered: ammonium, lactic acid, and carbon  
 592 dioxide. The chemical equilibrium system shown in Table 1 allows us to set the  
 593 following ordinary differential equation system:

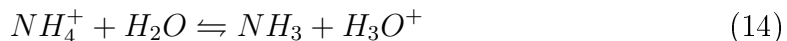
Table 2: Kinetics expressions corresponding to main pH contributors

No.	pH contributor	Kinetics expression
1	Ammonium	$\frac{d[H_3O^+]}{dt} = k_{H_3O^+}(K_{a,NH_4^+}[NH_4^+] - [NH_3][H_3O^+])$
2	Lactate	$\frac{d[H_3O^+]}{dt} = k_{H_3O^+}(K_{a,HLac}[HLac] - [Lac^-][H_3O^+])$
3	Carbon dioxide	$\frac{d[H_3O^+]}{dt} = k_{H_3O^+}(K_{a,H_2CO_3}[CO_2] - [HCO_3^-][H_3O^+])$ $\frac{d[H_3O^+]}{dt} = k_{H_3O^+}(K_{a,HCO_3^-}[HCO_3^-] - [CO_3^{2-}][H_3O^+])$

594 Where  $k_{H_3O^+}$  is the specific equilibrium displacement rate ( $s^{-1}$ ).

$$\frac{d[H_3O^+]}{dt} = \left(\frac{d[H_3O^+]}{dt}\right)_{NH_4^+} + \left(\frac{d[H_3O^+]}{dt}\right)_{HLac} + \left(\frac{d[H_3O^+]}{dt}\right)_{CO_2} \quad (13)$$

595 In biological systems, ammonia stripping takes place when ammonia is produced in  
 596 the presence of continuous aeration. For a better understanding, the ammonia equi-  
 597 librium should be first analyzed. According to the chemical equilibrium of Eq. 14,  
 598 the production of ammonium via metabolic activity shifts the equilibrium, favoring  
 599 free ammonia production.



600 The free ammonia is liberated from the bioreactor liquid medium phase by aeration.  
 601 Due to the low partial pressure of ammonia in the atmospheric air, the ammonia air  
 602 flow saturation occurs instantaneously when it enters the biological reactor through-  
 603 out the diffusion system. The free ammonia production rate can be expressed as  
 604 follows:

$$\frac{d[NH_3]}{dt} = k_{NH_3}(K_{a,NH_4^+}[NH_4^+] - [NH_3][H_3O^+]), \quad (15)$$

605 where  $k_{NH_3}$  is the specific equilibrium displacement rate ( $s^{-1}$ )(with the same rule  
 606 proposed by [Batstone et al. \(2002\)](#) to reduce the system stiffness).

607 In diffused air systems, saturation with ammonia occurs within the first few mil-  
 608 limeters of the ascent of the air bubbles through the liquid. Because this saturation  
 609 process is almost instantaneous, the measured duration of ammonia desorption in  
 610 the diffused air system is, in effect, a measure of the volume of air needed to achieve  
 611 the observed removal of ammonia ([Srinath and Loehr, 1974](#)). Therefore, the ammo-  
 612 nia removal process can be expressed using the following mathematical expression:

$$\frac{d[NH_3]}{dt} = k_{NH_3,st}[NH_3] \quad (16)$$

613 According to [Srinath and Loehr \(1974\)](#), the specific ammonia removal rate by aer-  
 614 ation can be calculated using the following equation:

$$k_{NH_3,st} = 0.021e^{[1.93 \times 10^{-4} \frac{q_{Air}}{V} + 0.062(T-5)]}, \quad (17)$$

615 where  $k_{NH_3,st}$  is the specific ammonia removal rate ( $h^{-1}$ ),  $q_{Air}$  is the air flow injected  
616 to the bioreactor ( $cm^3/min$ ),  $V$  is the bioreactor volume (l) and  $T$  is the tempera-  
617 ture inside the bioreactor ( $^{\circ}C$ ). The Eq. 17 considers the airflow intensity and the  
618 temperature influence on the ammonia removal process.

#### 619 4.4. Critical analysis of the published CHO mechanistic models

620 In the literature, two main model approaches can be found to describe CHO cell  
621 metabolism:

- 622 1. Considering the overall biological reactions (Xing et al., 2010; Jimenez del Val  
623 et al., 2016; López-Meza et al., 2016; Yahia et al., 2021),
- 624 2. Separating anabolism and catabolism reactions, considering ATP/ADP and  
625 NADH/NAD<sup>+</sup> electron transporters (Nolan and Lee, 2011).

626 Both approaches can describe the glycosylation process (Jimenez del Val et al., 2016;  
627 Kotidis et al., 2019). The protein quality depends on glycosylation, which occurs  
628 within the endoplasmatic reticulum and Golgi apparatus. This process takes place  
629 along the protein secretory pathway. It involves attaching an oligosaccharide chain  
630 to an amino acid residue, primarily asparagine (N-linked) or serine/threonine (O-  
631 linked glycosylation) (Galleguillos et al., 2017). The pros and cons of these models  
632 are thoroughly analyzed in Table 3. In most of these models, Monod's law and  
633 its variant for inhibition are used to describe the metabolic process rates, where  
634 glucose and glutamine or glutamate are the main limiting substrates (Xing et al.,  
635 2010; Jimenez del Val et al., 2016; López-Meza et al., 2016). In contrast, lactate  
636 and ammonia are the main inhibitory substrates (Xing et al., 2010; Kotidis et al.,  
637 2019). Lactate, an inhibitory product of CHO cell metabolism, becomes a limiting  
638 substrate during the stationary phase (Nolan and Lee, 2011; Kotidis et al., 2019).



Table 3: Pros and cons of mechanistic CHO models

Model description	Pros	Cons
<a href="#">Xing et al. (2010)</a> : Model based on Monod's law and its variant for inhibition.	<ul style="list-style-type: none"> <li>• Protein production independent of cell growth,</li> <li>• <math>\mu</math> is used as an emerging factor of metabolic shift,</li> <li>• Adequate prediction of the main model variables.</li> </ul>	<ul style="list-style-type: none"> <li>• No <i>Lac</i> utilization for cell growth during the stationary phase.</li> </ul>
<a href="#">Nolan and Lee (2011)</a> : Model based on empirical kinetic rates describing intracellular catabolic and anabolic reactions.	<ul style="list-style-type: none"> <li>• Protein production modeled independently of cell growth,</li> <li>• <i>Lac</i> utilization for cell growth during the stationary phase,</li> <li>• Adequate prediction of the main model variables.</li> </ul>	<ul style="list-style-type: none"> <li>• No decay process inclusion,</li> <li>• No <math>NH_4^+</math> inhibitory effect inclusion,</li> <li>• Complex and over-parameterized model.</li> </ul>
<a href="#">Jimenez del Val et al. (2016)</a> : Model linking <i>mAbs</i> glycosylation with cell secretory capacity based on Monod's law and its variant for inhibition.	<ul style="list-style-type: none"> <li>• Integration of protein glycosylation with cellular secretory capacity,</li> <li>• Protein production modeled independently of cell growth,</li> <li>• Adequate prediction of the main model variables.</li> </ul>	<ul style="list-style-type: none"> <li>• No <i>Lac</i> inhibitory effect inclusion,</li> <li>• No <i>Glu</i> limiting effect inclusion,</li> <li>• No <math>NH_4^+</math> production inclusion, hence its inhibiting effect.</li> </ul>
<a href="#">López-Meza et al. (2016)</a> : Model based on Monod's law for cell growth and Luedeking-Piret model for <i>mAbs</i> production.	<ul style="list-style-type: none"> <li>• Satisfactory prediction of <math>\mu</math>, <math>X_v</math>, and protein concentration over time.</li> </ul>	<ul style="list-style-type: none"> <li>• No consideration of <i>Lac</i> and <math>NH_4^+</math> production and their inhibiting effect,</li> <li>• No <i>Lac</i> utilization for cell growth,</li> <li>• Protein production associated with the cell growth process.</li> </ul>
<a href="#">Kotidis et al. (2019)</a> : Model describing the influence of glycosylation precursor feeding on CHO cell metabolism based on Monod's laws, its variant for inhibition, and empirical rates.	<ul style="list-style-type: none"> <li>• <i>Lac</i> production used as a trigger of metabolic shift.</li> <li>• Inclusion of <i>Lac</i> and <math>NH_4^+</math> inhibiting effects and glycosylation,</li> <li>• <i>Lac</i> utilization for cell growth during the stationary phase,</li> <li>• Adequate prediction of the main model variables.</li> </ul>	<ul style="list-style-type: none"> <li>• Protein production associated with the cell growth process.</li> </ul>
<a href="#">Yahia et al. (2021)</a> : Model based on mixed Monod's law and its variant for inhibition.	<ul style="list-style-type: none"> <li>• Adequate prediction of protein production over time.</li> </ul>	<ul style="list-style-type: none"> <li>• No consideration of <i>Lac</i> and <math>NH_4^+</math> production and their inhibiting effect,</li> <li>• No <i>Glc</i> and <i>Lac</i> utilization for cell growth and protein production,</li> <li>• Protein production associated with the cell growth process.</li> </ul>

639 On the other hand, some empirical expression rates are also used to describe metabolic  
640 reactions involving intermediate metabolites, as in modeling central carbon metabolism  
641 (Nolan and Lee, 2011). Different stoichiometry and kinetics are assumed when  
642 the temperature is changed to trigger the metabolic shift (Jimenez del Val et al.,  
643 2016; López-Meza et al., 2016). In contrast, they are assumed to be constant when  
644 the metabolic shift occurs spontaneously (Xing et al., 2010; Kotidis et al., 2019).  
645 The metabolic shift is crucial in CHO cell metabolism modeling as it connects the  
646 exponential growth and stationary phases. Therefore, accurately predicting the  
647 metabolic shift is a genuine necessity for modeling. Yet this prediction remains  
648 an open challenge as this phenomenon is poorly understood. The most common  
649 indicators used in modeling for connecting both metabolic phases are temperature  
650 (Jimenez del Val et al., 2016), lactate concentration (Kotidis et al., 2019), and spe-  
651 cific growth rate ( $\mu$ ) as an emerging indicator (Xing et al., 2010). Indeed,  $\mu$  depends  
652 on limiting substrates, inhibiting by-products, and the temperature, which includes  
653 the main factors associated with this phenomenon as reported in the literature (Fig.  
654 7). Several models describe the process of protein production associated with cell  
655 growth, where protein is considered a metabolite of cell growth, especially during  
656 the exponential phase (López-Meza et al., 2016; Kotidis et al., 2019; Yahia et al.,  
657 2021). The MFA (Ahn and Antoniewicz, 2011; Sengupta et al., 2011; Templeton  
658 et al., 2013) and modeling studies (Xing et al., 2010; Kotidis et al., 2019) suggest  
659 that there is always a significant increase in protein production following a sub-  
660 stantial decrease in specific growth rates (either by an increase in the concentration  
661 of inhibitory metabolites or by a reduction in temperature). This finding demon-  
662 strates the antagonism between the two metabolic processes. It should also be noted  
663 that protein production is directly proportional to the concentration of viable cells.  
664 Consequently, to accurately predict protein production, it is crucial to consider cell  
665 growth and death processes. Cell death, influenced mainly by the toxic impact of

666 ammonia, plays a key role during the stationary phase (Xing et al., 2010).

667 The exponential phase is characterized by rapid cell growth and low protein produc-  
668 tion. For this reason, some authors omit this metabolic pathway in the exponential  
669 phase submodel. However, the yield of protein production from glucose is relatively  
670 low, requiring a significant amount of substrate for its synthesis, which can consid-  
671 erably influence the prediction of glucose consumption over time, even at very low  
672 protein production rates. Therefore, despite the low protein production, including  
673 it in the model is crucial.

674 Another important challenge of CHO cell metabolism models is the large number  
675 of kinetics and stoichiometric parameters involved, many of which are impossible to  
676 determine experimentally. In the case of models based on central carbon metabolism,  
677 although stoichiometry is known a priori, the kinetics involved are complex due to  
678 the number of intermediate processes, which in most cases include parameters and  
679 variables that are difficult to quantify. To summarize this literature review, the  
680 following aspects can be highlighted:

- 681 • The models offer poor investigation and description of metabolic pathways,
- 682 • In most cases, lactate use for cell growth is not considered,
- 683 • Criteria used for metabolic shift prediction are still not adequate,
- 684 • Two main pitfalls are observed: complex over-parameterized models and con-  
685 siderably simple models, both of which have little application in industry,
- 686 • The direct experimental determination of stoichiometric parameters is very  
687 difficult due to the simultaneous processes such as cell growth and protein  
688 production.

689 **5. Numerical solution and calibration**

690 *5.1. Computational solutions of ODEs*

The models under consideration in this work assumes the bioreactor to be perfectly stirred. The set of coupled equations is then a set of ODEs (Ordinary Differential Equations) of the generic form:

$$\begin{cases} u'(t) = F(u(t)) & (18a) \\ u(t_0) = u_0 & (18b) \end{cases}$$

691 Where  $u(t)$  is the vector of unknowns and  $F(u)$  is a vector of functions, the same  
692 size as  $u$ , giving the time-derivative of vector  $u$ .

693 Even though the equations are coupled and nonlinear, their computational solution  
694 is generally quite simple, using classical iterative methods to solve ODEs. The  
695 solution is obtained at discrete times  $t_n$  ( $u_n = u(t_n)$ ). The linear  $k$ -step methods,  
696 an important family of solutions, can be defined as a general expression:

$$\sum_{j=0}^k \alpha_j u_{n+j} = h \sum_{i=0}^k \beta_i F_{n+i} \quad \text{with} \quad \alpha_k = 1 \quad (19)$$

697 Where  $h$  is the time step and  $F_{n+j} = F(u_{n+j})$ .

698 Expression (19) is a recurrent expression that allows the value  $t(t_n) = u(t_n + h)$  to  
699 be computed from all previous values up to time  $t_n$ .  $b_k = 0$  for explicit methods  
700 and  $b_k \neq 0$  for implicit methods. For example, the simple forward (explicit) Euler  
701 method is obtained with  $k = 1, \alpha_0 = -1, \beta_1 = 0$ , and  $\beta_0 = 1$ . Using  $\beta_1 = 1$  and  
702  $\beta_0 = 0$  instead gives the backward (implicit) Euler method. Expression (19) also  
703 includes the multistep explicit Adams–Bashforth and implicit Adams–Mouton meth-  
704 ods (Butcher, 2000). Multistep methods aim to increase the order of the method  
705 (rate of convergence when  $h$  decreases). The Runge–Kutta methods are one-step

706 methods that use another strategy to increase the convergence order: in a single  
707 step, the function  $F$  is evaluated several times to obtain a Taylor expansion of the  
708 desired order.

709 The reader should, however, be aware that the solution order is not the single  
710 criterion for choosing a method. Formulations resulting from biological assumptions  
711 might be challenging to solve. This is the case, for example, with formulations  
712 involving several compartments (bioreactor, cells, organelles) and with very stiff  
713 functions, such as the switching varying over very narrow concentration levels (for  
714 example, the [La et al. \(2020\)](#) function shown in Fig. 8). In such cases, the system  
715 has components that vary on very different time scales, which poses challenges in  
716 selecting a suitable time step for numerical integration methods, resulting in stiff  
717 ODEs. The concept of convergence of stiff ODEs was introduced by [Dahlquist](#)  
718 (1963). He introduced the concept of A-convergence and proved that:

- 719 • A-convergence is possible only with implicit methods,
- 720 • A-convergence is not possible for methods above order 2.

721 It is important to keep this in mind when using the algorithms implemented in  
722 generic tools, such as the package *ode* in *R*, the class *scipy.integrate.ode* in *Python*  
723 or the existing *Matlab* solvers, to solve equations (18a) and (18b). The reader could  
724 refer to published works ([Butcher, 2000](#); [Cash, 2003](#)) to pick up the suitable options  
725 for these solvers. All these solvers perform quite well for standard problems. As  
726 ODEs are much less demanding than PDEs (Partial Differential Equations), the  
727 CPU time usually remains very low (in the order of some seconds or even less). One  
728 point of attention, though, is that the system of equations may have components  
729 with very different orders of magnitude. Due to variations on very small quantities,  
730 the evolution of some variables may not be correctly taken into account in the  
731 convergence criterion, leading to a bad control of the time step and erroneous results.

732 For very severe configurations, more sophisticated algorithms, if possible coded in  
733 low-level languages (Fortran or C) for their efficiency once compiled, should therefore  
734 not be excluded a priori. Once compiled, these tools could be embedded in high-  
735 level tools such as the ones cited above. Even with classical problems, the efficiency  
736 of such solvers in terms of CPU is also likely to open new routes for the usage of  
737 mechanistic modeling: i) using these tools in an optimization loop requiring many  
738 solutions to be computed, ii) online parameter identification or iii) hybrid modeling  
739 for example.

740 Several methods were proposed to efficiently solve stiff ODEs (Butcher, 2000; Ab-  
741 dulle and Medovikov, 2001; Cash, 2003; Fatunla, 2014; Lebedev, 2017). The family  
742 of exponential integrators (exponential Rosenbrock-type integrators) are certainly  
743 among the most efficient ones (Cox and Matthews, 2002; Tokman, 2006; Caliar  
744 and Ostermann, 2009; Hochbruck et al., 2009; Carr et al., 2013). To derive these  
745 exponential methods, the Jacobian of  $F$  at point  $u_n$ , noted  $J_n$ , is first introduced in  
746 equation (18a):

$$u'(t) = F_n + J_n(u(t) - u_n) + R(u(t)) \quad (20)$$

747 Where  $R$  is the remainder.

748 Using the integration factor  $e^{-F_n t}$ , equation (20) becomes

$$u(t_n + h) = u_n + (e^{J_n h} - I)J_n^{-1}F_n + \int_{t_n}^{t_n+h} e^{J_n(t_n+h-t)} R(u(t))dt \quad (21)$$

749 All exponential methods rely on suitable methods to evaluate:

- 750 • the second term of the left-hand side, which needs to compute  $\varphi(z) = \frac{\exp(z)-1}{z}$   
751 where  $z$  is a matrix,
- 752 • the third terms of the left-hand side, where  $R$  can simply be neglected, evalu-

753 ated assuming  $J$  to vary linearly over the time step, or by several evaluations,  
754 such as with a Runge–Kutta method.

## 755 5.2. Calibration and validation

756 The calibration process of a mathematical model is essential for its validation and  
757 future application (Rajamanickam et al., 2021). It consists of deterministic cal-  
758 culation of model parameter values consistent with data (Dawkins et al., 2001).  
759 The most appropriate strategy for model calibration is to experimentally determine  
760 as many model parameters as possible (direct determination)(González-Hernández  
761 et al., 2022). The remaining parameters can be taken from the literature if their  
762 value does not change significantly from one system to another (usually parameters  
763 with little influence on the model) or can be estimated by inverse analysis through an  
764 optimization process that searches for the unique combination of model parameters  
765 that allows a minimum deviation between the model output and the experimental  
766 data (indirect determination)(González-Hernández et al., 2022). Sometimes, when  
767 the model is simple, depending on the experience of modelers, this process can be  
768 carried out manually (Sin et al., 2008). The quality of the calibration process will  
769 depend on the quantity and quality of the data (Boudreau and McMillan, 2007; Abt  
770 et al., 2018) (economic resources availability, user competency, software access...).  
771 Therefore, the data collection process must be carefully carried out with an optimal  
772 and well-defined experimental design considering the variation of as many param-  
773 eters as possible to obtain representative data of the phenomena described in the  
774 model (Rajamanickam et al., 2021). If historical data is used, much attention should  
775 be paid to the data curation process and selecting representative data. Finally, it  
776 is essential to remark that the model, once calibrated, must be tested against a  
777 validation database under different operating conditions.

778 5.2.1. Parameter estimation by inverse analysis method

779 The inverse analysis can be performed through an optimization procedure, looking  
780 for a minimum value using objective or multi-objective functions:

$$\begin{aligned} \text{Min} : & \{f_1(x), f_2(x), f_3(x), \dots, f_n(x)\} \\ & \text{subject to: } x \in S, \end{aligned} \tag{22}$$

781 where  $x$  is the solution corresponding to the minimization process of  $n$  objective  
782 functions in the subspace  $S$  of the calibration parameter ranges.

783 Multi-objective optimization is essential when faced with real-world optimization  
784 problems since, in most cases, they are conditioned by multiple conflicting objec-  
785 tives with different levels of importance (Deb, 2014). In CHO models, two main  
786 metabolic phases are described mathematically, which present significant stoichio-  
787 metric and kinetic differences (Jimenez del Val et al., 2016). When CHO cells are  
788 used as a host for protein production in the pharmaceutical industry, more impor-  
789 tance is given to the stationary growth phase, characterized by a strong production  
790 of antibodies. In this case, multi-objective functions are helpful, as they allow for  
791 reinforcing the significance of specific phenomena variables during the calibration  
792 process, if necessary. This reinforcement can be achieved when the factor or vari-  
793 able's importance is given as a weighting average of all subfunctions:

$$F(x) = \sum_{i=1}^n w_i f_i(x) \quad \text{with: } \sum_{i=1}^n w_i = 1, \tag{23}$$

794 where  $w_i$  is the weighting factor, whose value is chosen by the decision maker ac-  
795 cording to his constraints (Florez et al., 2023).

796 For example, the mean relative error (MRE) between the experimental data (vi-  
797 able cells, glucose, glutamate, glutamine, lactate, ammonium, etc.) and the model  
798 prediction. This objective function can be adapted to weigh the importance of the



799 experiments, variables, and data points considered during the calibration process:

$$F(x) = \frac{1}{lmn} \sum_{i=1}^{\ell} w_i \sum_{j=1}^m w_j \sum_{k=1}^n w_k \left| \frac{y_{e(k,j)}^{(i)} - y_{m(k,j)}^{(i)}}{y_{e(k,j)}^{(i)}} \right| = \begin{cases} MRE, & \text{if } w_i, w_j, w_k = 1 \\ \neq MRE, & \text{if } w_i, w_j, w_k \neq 1 \end{cases}, \quad (24)$$

800 where  $y_{e(j,k)}^{(i)}$  is the experimental data value  $-k$  of variable  $-j$  in experiment  $-i$ ,  
 801  $y_{m(j,k)}^{(i)}$  is the model output data value  $-k$  of variable  $-j$  in experiment  $-i$ ,  $w_k$  is the  
 802 weight assigned to data value  $-k$ ,  $w_j$  is the weight assigned to variable  $-j$  and  $w_i$   
 803 is the weight assigned to experiment  $-i$ .

804 An alternative approach involves using self-guided fitness functions that speed up  
 805 the optimization process while reducing the risk of getting stuck in a local minimum.  
 806 As an example, we can propose a fitness function that effectively amalgamates ac-  
 807 curacy, utilizing the mean relative error (MRE), with variable trends (via Pearson's  
 808 correlation coefficient ( $r$ )):

$$fitness = \frac{1}{3} \left( \underbrace{MRE}_{\text{Accuracy}} + \underbrace{(1-r)}_{\text{Trend}} + \underbrace{|MRE - (1-r)|}_{\text{Equilibrium}} \right) \quad (25)$$

809 where:

$$r = \frac{1}{lm} \sum_{i=1}^{\ell} \sum_{j=1}^m \frac{n(\sum_{k=1}^{k=n} y_e y_m) - (\sum_{k=1}^{k=n} y_e)(\sum_{k=1}^{k=n} y_m)}{\sqrt{[n \sum_{k=1}^{k=n} y_e^2 - (\sum_{k=1}^{k=n} y_e)^2][n \sum_{k=1}^{k=n} y_m^2 - (\sum_{k=1}^{k=n} y_m)^2]} \quad (26)$$

810 In equation 25, the equilibrium term ensures an equitable contribution between  
 811 accuracy and trends. This type of function significantly accelerates the convergence  
 812 of the optimization process while decreasing the probability of becoming trapped in  
 813 a local minimum.

### 814 5.2.2. Calibration challenges

815 During calibration, the model simulations are compared to experiments (Villaverde  
816 et al., 2022). A typical experimental data set consists of a number of batch or  
817 fed-batch trials, for which the information collected consists of growing conditions  
818 and time series of certain variables (in-line and/or offline measurements). This can  
819 represent a huge data set but with few contrasted conditions. The system is then  
820 likely to be over-determined, which is necessary to gain accuracy and counterbal-  
821 ance experimental noise/variability. This allows the system to be projected onto a  
822 smaller space, where the solution is optimized. However, these conditions are not  
823 sufficient: the series of experiments must also be able to test each parameter, ob-  
824 taining significant variations in the measured variables when these parameters are  
825 modified. Indeed, one major challenge in bioprocess modeling is addressing over-  
826 parameterization (Mowbray et al., 2023), often resulting from an excessive number  
827 of parameters and/or highly correlated, that cannot be directly determined by ex-  
828 periments (Barz et al., 2015; Abt et al., 2018). Modelers often use inverse analysis  
829 but encounter difficulties due to the impossibility of generating experimental data  
830 that adequately captures the effect of all model parameters. Parametric Sensitivity  
831 Analysis is a widely adopted technique for this purpose (Kyriakopoulos et al., 2018).  
832 It effectively reduces the parameter calibration subspace by excluding less influen-  
833 tial parameters and, in some cases, omitting less significant phenomena, to obtain  
834 a determinate system.

835 Another method to address over-parameterization consists of developing simplified  
836 models (Sha et al., 2018), called lumped models. This method models the set of  
837 metabolic pathways within a cellular compartment, such as glycolysis, the Krebs  
838 cycle, or oxidative phosphorylation, as a single global metabolic pathway (La et al.,  
839 2020). Alternatively, it could also be considered the integration of these closely re-  
840 lated compartments as a unified entity through global biochemical reactions involv-

841 ing extracellular metabolites, a commonly employed practice in bioprocess modeling  
842 (Xing et al., 2010; Jimenez del Val et al., 2016; Kotidis et al., 2019).

843 Modelers also face challenges when introducing new parameters without prior ref-  
844 erences in the literature. These parameters often pose significant difficulties for  
845 direct experimental determination, making it necessary to employ inverse analysis.  
846 Consequently, a new challenge arises: establishing calibration bounds for these pa-  
847 rameters. In such situations, using numerical derivatives can offer valuable insights,  
848 allowing the understanding of how these parameters affect the fitness function. This  
849 process may require a tedious iterative procedure: the user analyzes the results at  
850 each stage and decides which parameter to select for evaluation at the next iteration  
851 stage, until the desired results are achieved.

### 852 *5.3. Optimization methods for minimizing the objective function*

853 First, note that the objective of this study is not to compare optimization methods  
854 but to mention the most used methods and explain their advantages and disad-  
855 vantages for the calibration of mechanistic models. Model calibration by inverse  
856 analysis requires simple and robust mathematical optimization methods. Swarm in-  
857 telligence and evolutionary computation (SIEC) are now some of the most popular  
858 optimization methods employed in scientific research (Bansal et al., 2019; Kumar  
859 et al., 2019). These methods use a stochastic approach, allowing one to solve many  
860 complex problems without demanding many mathematical properties (e.g., convex-  
861 ity, continuity, or the explicit definition of the objective function) (Bansal et al.,  
862 2019). Particle swarm optimization (PSO) is among the most popular and success-  
863 ful swarm intelligence algorithms (Bansal et al., 2019). A great advantage of these  
864 algorithms is their parallelization capacity, considering that they are based on popu-  
865 lations (fitness function evaluation) that considerably reduce time consumption but  
866 require a powerful computational capacity depending on the complexity of the fit-  
867 ness function. Evolutionary Computation (EC), mainly used to solve optimization

868 problems, comprises a series of problem-solving techniques based on the principles  
869 of biological evolution (e.g., natural selection and genetic inheritance) that allow for  
870 finding optimal global solutions (Bansal et al., 2019).

871 Mechanistic models present a great complexity considering the considerable num-  
872 ber of stoichiometric and kinetic parameters that may be involved. Generally, the  
873 calibration process is carried out under non-stationary operating conditions, and  
874 multiple local minima may be encountered. For these reasons, heuristic optimiza-  
875 tion methods such as bio-inspired algorithms are mainly used for this task.

876 Several population-based, fitness-oriented, and variation-driven evolutionary algo-  
877 rithms have been proposed in the last century (Yu and Gen, 2010). These algorithms  
878 evolve using different strategies by employing common genetic operators such as se-  
879 lection, mutation, and reproduction, which depend on individual structures defined  
880 by an environment (Khaparde et al., 2022). In particular, in the area of evolutionary  
881 algorithms, the genetic algorithm (GA) and differential evolution (DE) algorithms  
882 are the most popular in the scientific community (Chaudhary et al., 2019). Cur-  
883 rently, bio-inspired methods such as particle swarm optimization (PSO) have seen  
884 a remarkable application in solving several engineering problems.

885

### 886 *Genetic Algorithm*

887 The genetic algorithm, inspired by natural selection, is one of the most popular  
888 and widely used algorithms in various research areas due to its ease of implementa-  
889 tion and convincing concepts (Omidinasab and Goodarzimehr, 2020; Goodarzimehr  
890 et al., 2022). Genetic algorithms are stochastic mathematical optimization methods  
891 based on the processes of natural selection and Darwinian survival of the fittest  
892 (Wang, 2003). New populations are produced by iteratively using genetic operators  
893 (e.g., chromosomal representation, selection, crossover, and mutation) on the indi-  
894 viduals of a population. GAs are highly parallel, based on individual populations

895 of optimal candidate solutions that can be evaluated simultaneously. However, one  
896 of the main limitations of this method is its premature convergence, as they are  
897 sometimes trapped in local minima. Some researchers have suggested increasing  
898 diversity through selection pressure to avoid this problem (Katoch et al., 2021).

899

### 900 *Differential Evolution*

901 The differential evolution algorithm is a combinatorial algorithm based on popu-  
902 lations of individuals. Like the GA, it allows the resolution of complex real-world  
903 optimization problems, which, in most cases, are reduced to the search for the  
904 global minimum of non-differentiable, discontinuous, and nonlinear objective func-  
905 tions (Lilla et al., 2013). Similar to other evolutionary algorithms, DE is a method  
906 that performs a stochastic search using a population of candidate solutions, apply-  
907 ing mutation, crossover, and selection operators that drive the population toward  
908 better solutions in the optimization space (Georgioudakis and Plevris, 2020). DE dif-  
909 fers from traditional evolutionary algorithms in generating new candidate solutions,  
910 employing a greedy generation scheme (Vesterstrom and Thomsen, 2004). The DE  
911 algorithm is simple but robust, governed by few control parameters, and its structure  
912 facilitates parallel computation with high convergence speed. Even when DE has  
913 few control parameters, their adjustment remains difficult, and their inappropriate  
914 manipulation can lead to premature convergence or stagnation, being key strategies  
915 chosen for the mutation operator. Therefore, to improve the performance of this  
916 algorithm a self-adaptive parameter setting technique is required (Khaparde et al.,  
917 2022). In this sense, new advanced DE variants with adaptive and self-adaptive con-  
918 trol parameters have been developed (Eiben et al., 1999; Georgioudakis and Plevris,  
919 2020).

920

### 921 *Particle Swarm Optimization*

922 The particle swarm optimization method was introduced in the mid-1990s by [Kennedy](#)  
923 [and Eberhart \(1995\)](#). This method is among the most widely used bio-inspired algo-  
924 rithms for solving optimization problems that are loosely inspired by foraging flocks  
925 of birds ([Couceiro and Ghamisi, 2016](#)). According to this analogy, each bird (consid-  
926 ered a particle in the algorithm) uses its memory and the knowledge acquired by the  
927 swarm while searching for the best available food source ([Venter and Sobieszczanski-](#)  
928 [Sobieski, 2003](#)). This algorithm is governed by three fundamental operators: mem-  
929 ory, inertia, and socialization. Subsequently, PSO has been extensively used to solve  
930 real-world problems in various biological and medical applications, computer graph-  
931 ics, and music composition ([Sedighizadeh and Masehian, 2009](#)). However, the main  
932 weakness of this optimization method is its propensity to converge to local optima  
933 prematurely ([Banks et al., 2007](#); [Houssein et al., 2021](#)). One way to mitigate these  
934 problems is to work with optimization parameter ranges that are as narrow as pos-  
935 sible.

936

### 937 *Hybrid Optimization Methods*

938 Hybrid optimization methods have become increasingly popular in recent years, es-  
939 pecially in artificial intelligence, as they combine desirable properties of two or more  
940 optimization methods to increase their performance by mitigating their individual  
941 weaknesses ([Thangaraj et al., 2011](#)). It is important to remark that most of these  
942 algorithms have been modified by several authors to reduce the time consuming  
943 and facilitate the convergence process ([Trivedi et al., 2015](#); [Abidin, 2018](#); [Aguitoni](#)  
944 [et al., 2018](#); [Chaudhary et al., 2019](#)). These algorithms have been integrated, tak-  
945 ing advantage of the best of each one. Such connections may be implemented in  
946 many ways: (1) The use of a less accurate algorithm to initialize a more accurate  
947 algorithm (2) The parallel operation of two or more independent algorithms, with  
948 timely communication between them, to exchange the best individual solutions, (3)

949 The sequential execution of several algorithms in a single algorithm, (4) The com-  
 950 bination of genetic operators from different optimization methods in each iteration,  
 951 and others (Dziwiński and Bartczuk, 2019).

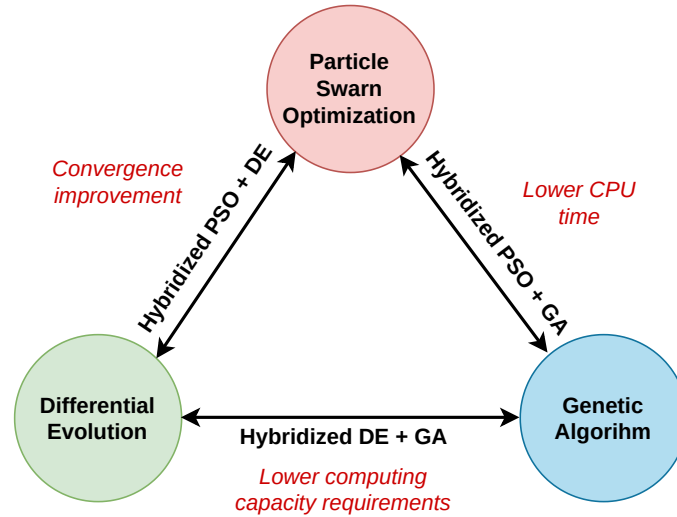


Fig. 9: Optimization methods hybridization

952 To achieve better performance (convergence speed, accuracy, and global optimiza-  
 953 tion ability) compared to individual methods, several hybrid approaches appear in  
 954 the literature, including DE and PSO (Hendtlass, 2001; Zhang and Xie, 2003; Talbi  
 955 and Batouche, 2004; Hao et al., 2007; Das et al., 2008; Vaisakh et al., 2009; Garcia-  
 956 Guarin et al., 2019; Dash et al., 2020; Li et al., 2021), PSO and GA (Mousa et al.,  
 957 2012; Esmine et al., 2013; Samuel and Rajan, 2015; Ali and Tawhid, 2017; Dziwiński  
 958 and Bartczuk, 2019; Pu et al., 2019; Omidinasab and Goodarzimehr, 2020; Fu et al.,  
 959 2021; Goodarzimehr et al., 2022), and GA and DE (Trivedi et al., 2015, 2016; Abidin,  
 960 2018; Aguitoni et al., 2018; Chaudhary et al., 2019; Fathy et al., 2020).

961 Fortunately, most of these algorithms are already included in optimized and veri-  
 962 fied packages in popular programming tools used by the scientific community, such  
 963 as MATLAB, Python, R, and others, which considerably facilitates the work of  
 964 modelers. However, the selection of the optimization method depends on the mod-  
 965 elers' experience, the fitness function's complexity, and computational capacity. We

966 strongly recommend that readers use the bio-inspired optimization method they  
967 know best, which is easy to handle with few control parameters, allowing them to  
968 obtain fast, efficient, and accurate results.

## 969 **6. Challenges and prospects of mechanistic modeling of CHO cells**

### 970 *6.1. Needs, concerns and improvement gaps*

971 Mechanistic modeling in the industry is rapidly expanding, finding applications in  
972 experimental design, scale-up, data analysis, product development, quality control,  
973 optimization, and decision-making, among other key areas (Hallow et al., 2010;  
974 Kuepfer et al., 2012; Helmlinger et al., 2017; Xing et al., 2023). In the pharma-  
975 ceutical industry, mechanistic modeling holds its prestige for its ability to convert  
976 process data into enhanced information to understand the process better, guide  
977 decision-making, and facilitate the development of digital and automated technolo-  
978 gies (Kuepfer et al., 2012; Sha et al., 2018; Narayanan et al., 2020). Examples of  
979 mechanistic modeling applications in CHO cell systems can be mentioned: e.g., Paul  
980 et al. (2019) improved the volumetric productivity of the CHO cell system operated  
981 in fed-batch mode by optimizing temperature and pH shifts by applying a mechanis-  
982 tic model, achieving an increase of 20% in the final product concentration; Kotidis  
983 et al. (2019) developed a mechanistic model (discussed earlier in section 4.4) that was  
984 successfully utilized to design and optimize the feeding strategy, achieving antibody  
985 concentrations exceeding 90% when compared to the control, without compromising  
986 the integral of viable cell density or the final antibody titer; Craven et al. (2014)  
987 used a nonlinear predictive model demonstrating its ability to achieve fixed-setpoint  
988 closed-loop glucose concentration control in a CHO cell system. As mentioned in  
989 this review article, the modeling of CHO cell metabolism has evolved considerably  
990 in recent decades, leading to a better understanding of the phenomena and a math-  
991 ematical description. This paper focused on operational models as candidates for



992 optimization and improved control/command of the production process. Simple  
993 and complex models can be found in the literature. In most cases, these models  
994 are over-parameterized, which makes their practical application difficult, as the sto-  
995 ichiometry and kinetics change with the CHO strain, requiring recalibration of the  
996 model parameters. These over-parameterized models require much more informa-  
997 tion about the system for their application, which is hindered by the widespread use  
998 of optimized commercial basal mediums for CHO culture, of which, in most cases,  
999 the exact composition is still unknown. Moreover, in CHO systems, the direct ex-  
1000 perimental determination of certain stoichiometric parameters is very complex due  
1001 to the simultaneous activation of several metabolic pathways involving the same  
1002 substrate, such as cell growth and protein production from glucose or lactate (Figs.  
1003 4–6). The inverse analysis is a powerful tool for estimating these parameters during  
1004 calibration. On the other hand, simple models are insufficient to predict CHO cell  
1005 metabolism due to the omission of essential processes or variables in the model.  
1006 Metabolic reconstructions at the genomic scale are complex and often challenging  
1007 to study in great detail. Recently, an interesting approach proposed by [Martínez](#)  
1008 [et al. \(2022\)](#) simplifies the model by choosing subsystems based on the topology of  
1009 the metabolic model, thus preserving the essential biochemical reactions.  
1010 Finding the right balance between the complexity and operability of CHO mod-  
1011 eling is part of the research gap that must be bridged. This determination requires  
1012 that the metabolic pathways included in the model should be fed by relevant pa-  
1013 rameters. For example, this strategy allowed the metabolic shifts to be predicted in  
1014 the case of yeast ([González-Hernández et al., 2022](#)). In our opinion, one of the main  
1015 challenges for future improvements of the CHO model is the relevant prediction of  
1016 spontaneous metabolic shifts from the exponential to the stationary phase.  
1017 Including organelles in the cell metabolism, leading to a compartmented model, and  
1018 including balances of energy (ATP/ADP) and electron carriers (NADH/NAD<sup>+</sup>)

1019 are two promising possibilities to get a model able to spontaneously predict the  
1020 metabolic shifts as a function of time or growth conditions (La et al., 2020). Al-  
1021 though several studies suggest the importance of these redox metabolites in con-  
1022 trolling variables, their inclusion in the models is complex by the fact that these  
1023 metabolites are found at such low levels (on the order of  $\mu\text{M}$ ), which makes it dif-  
1024 ficult to quantify them, and thus to determine the stoichiometry and associated  
1025 kinetics (Nolan and Lee, 2011).

1026 Glycosylation is another process of great interest to the pharmaceutical industry  
1027 as it is crucial in biological activity and stability, increases the half-life, and re-  
1028 duces the immunogenicity of protein therapeutics (Kuriakose et al., 2016). Although  
1029 this phenomenon has already been successfully coupled with a mechanistic model  
1030 (Jimenez del Val et al., 2016; Kotidis et al., 2019), it involves many enzymatic reac-  
1031 tions, which could lead to the over-parametrization of the model. Machine learning  
1032 and molecular modeling could be an appropriate combination to obtain a better  
1033 description of the structure, quality, and function of the protein obtained by glyco-  
1034 sylation.

1035 Over longer time horizons, one can imagine that a comprehensive description of the  
1036 metabolic pathways could be included in operational tools. This description would  
1037 bridge the significant gap that remains between the knowledge gained by metabolic  
1038 flux analyses and operational simulation. Eventually, genome-scale models, together  
1039 with the incorporation of enzyme constraint and enzyme kinetics, could produce pre-  
1040 dictive models from the genomes of specific strains (Price et al., 2004; Davidi and  
1041 Milo, 2017). This development would be a major step forward in bridging the gap  
1042 between metabolic engineering and the control and command of high-performance  
1043 microbial strains. The GECKO project, for example, proves that this will certainly  
1044 be possible in a not-to-far future (Sánchez et al., 2017; Domenzain et al., 2021).

1045 *6.2. The digital twin at the crossroads of mechanistic modeling and data science*

1046 Control command represents one of the most sought-after modeling applications, to  
1047 improve the final product's concentration and quality. In our opinion, it is essential  
1048 to maintain the mechanistic model as the core of the control system due to its  
1049 prediction potential. To this end, there is a notable trend towards using hybrid  
1050 models for control commands, particularly with new online sensors that immediately  
1051 collect sufficient and relevant data to machine learning tools. Our research team is  
1052 currently immersed in an ambitious project called CALIPSO, with the primary goal  
1053 of halving development time and doubling productivity.

1054 At short time horizons, the fast-growing field of data science is about to change  
1055 how to build and tune mechanistic models and will address many of the above-  
1056 mentioned limitations. In particular, the various building blocks presented in detail  
1057 will be designed and improved concomitantly rather than sequentially. Over the  
1058 past decade, machine learning has spread across many areas of engineering sci-  
1059 ence. Autonomous cars face recognition, and weather forecasting without solving  
1060 the equations of physics are probably the most popular examples of the success of  
1061 machine learning. Machine learning has also spread to the field of bio-modeling  
1062 ([Pozzobon et al., 2021](#)). However, in most of the studies in these review papers,  
1063 machine learning is seen as an alternative to mechanistic modeling ([Baker et al.,](#)  
1064 [2018](#)). In this sense, machine learning can cope with complex situations, provided  
1065 the training database is large enough. The predictive capability is restricted to the  
1066 domain paved by the database.

1067 Rather than using machine learning instead of mechanistic modeling, we believe  
1068 both areas are now mature enough to benefit from the best of both worlds: a pre-  
1069 dictive model capable of adapting to different products. Such a digital twin approach  
1070 would work offline or online ([Zhang et al., 2020](#); [Yang and Chen, 2021](#)). In this hy-  
1071 brid approach, the mechanistic model would remain at the heart of the interactions

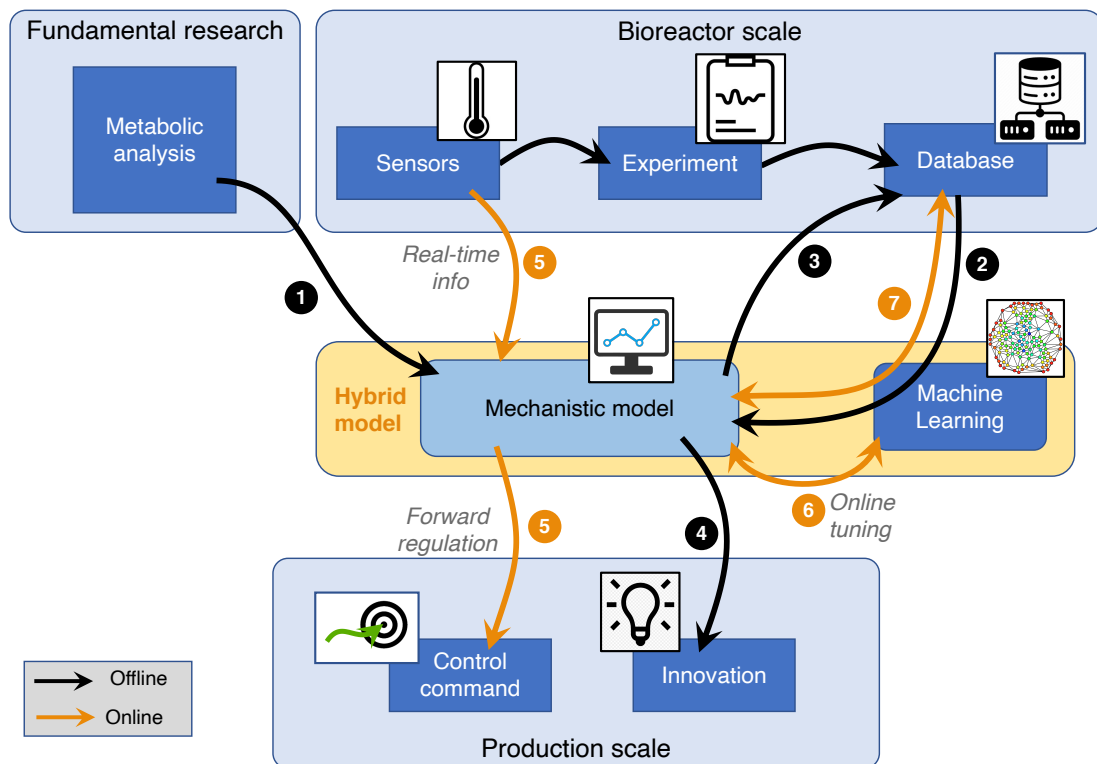


Fig. 10: The concept of hybrid modeling: using the synergies between experiment, mechanistic modeling, data science, and machine learning to fill the gap between modeling and process optimization (control command and innovation) by the concept of digital twin.

1072 (Fig. 10). Hybrid models not only deliver more precise predictive outcomes but also  
 1073 exhibit superior resilience and extrapolation abilities (Narayanan et al., 2019, 2021).

1074

1075 The offline work (black arrows in Fig. 10) starts with constructing the mechanistic  
 1076 model, assembling all building blocks detailed in this paper. The metabolic descrip-  
 1077 tion to be included in the model comes from fundamental studies (arrow 1). This is  
 1078 the cornerstone of mechanistic modeling. Besides, an experimental database must  
 1079 be generated from fermentation tests performed at the bioreactor level for various  
 1080 conditions using refined instrumentation (in-line and offline measurements). This  
 1081 database is primarily used for identifying the model parameters by inverse analy-  
 1082 sis, as explained in section 5. However, machine learning can also be used to test

1083 model assumptions, either to reduce the formulation in case of overestimation or  
1084 to recommend improvements (additional metabolic pathways, additional activat-  
1085 ing/inhibiting factors...) in particular, to represent non-standard data (this is why  
1086 arrow 2 passes through the machine learning box). Once calibrated, the mechanistic  
1087 model can be used offline to extend the database (arrow 4) or to imagine innovations  
1088 (arrow 5: new production protocols, new strategies to overcome problems...).

1089 The online use of the model (orange arrows of Fig. 10) includes at first the clas-  
1090 sical approach: using real-time information gathered by sensors together with the  
1091 mechanistic model for an efficient control/command of the bio-process (arrows 5).

1092 As the model is predictive, it is capable of forward regulation, which is a significant  
1093 advantage in reducing the culture time and making an early decision, for instance,  
1094 to stop a batch when its trajectory can no longer be recovered. To achieve better  
1095 performance in process control, machine learning can be used in two ways:

- 1096 • By using the complete set of information collected up to the present time,  
1097 machine learning can be used to tune online the model parameters. This is a  
1098 crucial advantage for bioprocesses (arrow 6),
- 1099 • by continuous queries to the database during cultivation to confirm whether  
1100 the current batch is atypical or inconsistent with the information provided  
1101 to the model. Identification of an atypical batch can reveal, for example,  
1102 contamination (arrow 7).

1103 In this last step of complexity, the mechanistic model and the database work in  
1104 synergy as a hybrid approach in which information provided by the mechanistic  
1105 model can benefit from the database, with queries triggered and analyzed by ma-  
1106 chine learning to complement the model. This process would allow, for example,  
1107 the use of simulations that are impossible to carry out in real-time, or for compar-  
1108 ing previous situations to detect specific issues (technical problems, sensor failure,

1109 product anomaly, etc.).

## 1110 **7. Conclusions**

1111 This review article details the various building blocks that must be assembled to  
1112 produce a mechanistic model of CHO cells for protein production. We have in-  
1113 tentively focused on operational models that can be used at a production scale.  
1114 The starting point is the set of metabolic pathways that must be provided to the  
1115 model. Once this input is completed, the activation and inhibition parameters and  
1116 the reaction rates must be defined to provide a self-contained model capable of re-  
1117 producing the dynamics of a bioreactor. One main concern is finding the right  
1118 balance between complexity and predictive capability. This balance includes the  
1119 choice of the metabolic description, the database quality able to fit the parameters,  
1120 and the prediction capabilities of the model in forward regulation.

1121 The main facts, recommendations, and prospects in the papers include:

- 1122 • To obtain a robust tool, over-parameterized models must be absolutely avoided,
- 1123 • We recommend using lactate for cell growth during the stationary phase as it  
1124 is not considered enough in the published works,
- 1125 • Inverse analysis stands out as a powerful tool for model calibration and val-  
1126 idation to overcome the complex direct experimental determination of stoi-  
1127 chiometric parameters due to concurrent processes, such as cell growth and  
1128 protein production,
- 1129 • Indicators used for metabolic shift prediction still need to be improved, which  
1130 motivates further investigations,
- 1131 • Compartmental models, able to account for balances in organelles and the  
1132 balance of energy and electron carriers, are two promising ways to predict the  
1133 metabolic shift better,

- 1134 • The intense use of machine learning and hybrid models, both offline and online,  
1135 will shape the future of mechanistic modeling,
- 1136 • In the long term, the considerable gap that remains between metabolic engi-  
1137 neering and command and control should be bridged.

## 1138 **8. Acknowledgments**

1139 Communauté urbaine du Grand Reims, Département de la Marne, Région Grand  
1140 Est and European Union (FEDER Champagne-Ardenne 2014-2020, FEDER Grand  
1141 Est 2021-2027) are acknowledged for their financial support to the Chair of Biotech-  
1142 nology of CentraleSupélec and the Centre Européen de Biotechnologie et de Bioé-  
1143 conomie (CEBB). This work was also co-funded by BPI France as part of the Calipso  
1144 project.

## 1145 **CRedit author statement**

1146 Y. González-Hernández: Conceptualization, Methodology, Formal analysis, Mod-  
1147 eling, Visualization, Writing-Original draft preparation, Writing-Review and Edit-  
1148 ing. P. Perré: Conceptualization, Methodology, Formal analysis, Modeling, Writ-  
1149 ing-Review and Editing, Supervision, Funding acquisition.

## 1150 **References**

- 1151 A Gibbons, L., Rafferty, C., Robinson, K., Abad, M., Maslanka, F., Le, N., Mo,  
1152 J., Clark, K., Madden, F., Hayes, R., et al., 2022. Raman based chemomet-  
1153 ric model development for glycation and glycosylation real time monitoring in a  
1154 manufacturing scale cho cell bioreactor process. *Biotechnol. Progr.* 38, e3223.
- 1155 Abdulle, A., Medovikov, A.A., 2001. Second order chebyshev methods based on  
1156 orthogonal polynomials. *Numer. Math.* 90, 1–18.

1157 Abidin, D., 2018. A hybrid genetic-differential evolution algorithm (hybgade) for a  
1158 constrained sequencing problem, in: 2018 International Conference on Artificial  
1159 Intelligence and Data Processing (IDAP), IEEE. pp. 1–6.

1160 Abt, V., Barz, T., Cruz-Bournazou, M.N., Herwig, C., Kroll, P., Möller, J., Pört-  
1161 ner, R., Schenkendorf, R., 2018. Model-based tools for optimal experiments in  
1162 bioprocess engineering. *Current opinion in chemical engineering* 22, 244–252.

1163 Aguitoni, M.C., Pavão, L.V., Siqueira, P.H., Jiménez, L., Ravagnani, M.A.d.S.S.,  
1164 2018. Heat exchanger network synthesis using genetic algorithm and differential  
1165 evolution. *Comput. Chem. Eng.* 117, 82–96.

1166 Ahn, W.S., Antoniewicz, M.R., 2011. Metabolic flux analysis of cho cells at growth  
1167 and non-growth phases using isotopic tracers and mass spectrometry. *Metab. Eng.*  
1168 13, 598–609.

1169 Aiba, S., Shoda, M., Nagatani, M., 1968. Kinetics of product inhibition in alcohol  
1170 fermentation. *Biotechnol. Bioeng.* 10, 845–864.

1171 Ali, A.F., Tawhid, M.A., 2017. A hybrid particle swarm optimization and genetic  
1172 algorithm with population partitioning for large scale optimization problems. *Ain*  
1173 *Shams Eng. J.* 8, 191–206.

1174 Altamirano, C., Illanes, A., Becerra, S., Cairó, J.J., Gòdia, F., 2006. Considerations  
1175 on the lactate consumption by cho cells in the presence of galactose. *J. Biotechnol.*  
1176 125, 547–556.

1177 Altamirano, C., Paredes, C., Illanes, A., Cairo, J., Godia, F., 2004. Strategies  
1178 for fed-batch cultivation of t-pa producing cho cells: substitution of glucose and  
1179 glutamine and rational design of culture medium. *J. Biotechnol.* 110, 171–179.

1180 Antoniewicz, M.R., 2013. Dynamic metabolic flux analysis—tools for probing tran-  
1181 sient states of metabolic networks. *Current opinion in biotechnology* 24, 973–978.

1182 Antoniewicz, M.R., 2021. A guide to metabolic flux analysis in metabolic engineer-  
1183 ing: Methods, tools and applications. *Metab. Eng.* 63, 2–12.



1184 Baker, R.E., Pena, J.M., Jayamohan, J., Jérusalem, A., 2018. Mechanistic models  
1185 versus machine learning, a fight worth fighting for the biological community? *Biol.*  
1186 *lett.* 14, 20170660.

1187 Banks, A., Vincent, J., Anyakoha, C., 2007. A review of particle swarm optimization.  
1188 part i: background and development. *Nat. Comput.* 6, 467–484.

1189 Bansal, J.C., Singh, P.K., Pal, N.R., 2019. Evolutionary and swarm intelligence  
1190 algorithms. volume 779. Springer.

1191 Barz, T., Körkel, S., Wozny, G., et al., 2015. Nonlinear ill-posed problem analysis  
1192 in model-based parameter estimation and experimental design. *Computers &*  
1193 *Chemical Engineering* 77, 24–42.

1194 Batstone, D.J., Keller, J., Angelidaki, I., Kalyuzhnyi, S., Pavlostathis, S., Rozzi, A.,  
1195 Sanders, W., Siegrist, H.a., Vavilin, V., 2002. The iwa anaerobic digestion model  
1196 no 1 (adm1). *Water Sci. Technol.* 45, 65–73.

1197 Boudreau, M.A., McMillan, G.K., 2007. New directions in bioprocess modeling and  
1198 control: maximizing process analytical technology benefits. *ISA*.

1199 Bree, M.A., Dhurjati, P., Geoghegan Jr, R.F., Robnett, B., 1988. Kinetic mod-  
1200 elling of hybridoma cell growth and immunoglobulin production in a large-scale  
1201 suspension culture. *Biotechnol. Bioeng.* 32, 1067–1072.

1202 Brunner, M., Doppler, P., Klein, T., Herwig, C., Fricke, J., 2018. Elevated pco<sub>2</sub>  
1203 affects the lactate metabolic shift in cho cell culture processes. *Eng. Life Sci.* 18,  
1204 204–214.

1205 Brunner, M., Kolb, K., Keitel, A., Stiefel, F., Wucherpfennig, T., Bechmann, J.,  
1206 Unsoeld, A., Schaub, J., 2021. Application of metabolic modeling for targeted  
1207 optimization of high seeding density processes. *Biotechnol. Bioeng.* 118, 1793–  
1208 1804.

1209 Budge, J.D., Roobol, J., Singh, G., Mozzanino, T., Knight, T.J., Povey, J., Dean,  
1210 A., Turner, S.J., Jaques, C.M., Young, R.J., et al., 2021. A proline metabolism

1211 selection system and its application to the engineering of lipid biosynthesis in  
1212 chinese hamster ovary cells. *Metab. Eng. Commun.* 13, e00179.

1213 Butcher, J.C., 2000. Numerical methods for ordinary differential equations in the  
1214 20th century. *J. Comput. Appl. Math.* 125, 1–29.

1215 Cacciatore, J.J., Chasin, L.A., Leonard, E.F., 2010. Gene amplification and vector  
1216 engineering to achieve rapid and high-level therapeutic protein production using  
1217 the dhfr-based cho cell selection system. *Biotechnol. Adv.* 28, 673–681.

1218 Caliari, M., Ostermann, A., 2009. Implementation of exponential rosenbrock-type  
1219 integrators. *Appl. Numer. Math.* 59, 568–581.

1220 Calmels, C., McCann, A., Malphettes, L., Andersen, M.R., 2019. Application of  
1221 a curated genome-scale metabolic model of cho dg44 to an industrial fed-batch  
1222 process. *Metab. Eng.* 51, 9–19.

1223 Carr, E.J., Turner, I.W., Perré, P., 2013. A variable-stepsize jacobian-free exponen-  
1224 tial integrator for simulating transport in heterogeneous porous media: Applica-  
1225 tion to wood drying. *J. Comput. Phys.* 233, 66–82.

1226 Carrillo-Cocom, L., Genel-Rey, T., Araíz-Hernández, D., López-Pacheco, F., López-  
1227 Meza, J., Rocha-Pizaña, M., Ramírez-Medrano, A., Alvarez, M., 2015. Amino  
1228 acid consumption in naive and recombinant cho cell cultures: producers of a  
1229 monoclonal antibody. *Cytotechnology* 67, 809–820.

1230 Cash, J., 2003. Efficient numerical methods for the solution of stiff initial-value  
1231 problems and differential algebraic equations. *Proceedings of the Royal Society of*  
1232 *London. Series A: Mathematical, Physical and Engineering Sciences* 459, 797–815.

1233 Chaudhary, D., Tailor, A.K., Sharma, V.P., Chaturvedi, S., 2019. Hygade: hybrid  
1234 of genetic algorithm and differential evolution algorithm, in: *2019 10th Interna-*  
1235 *tional Conference on Computing, Communication and Networking Technologies*  
1236 *(ICCCNT), IEEE.* pp. 1–4.

1237 Chen, G., Hu, J., Qin, Y., Zhou, W., 2021. Viable cell density on-line auto-control

1238 in perfusion cell culture aided by in-situ raman spectroscopy. *Biochem. Eng. J.*  
1239 172, 108063.

1240 Contois, D., 1959. Kinetics of bacterial growth: relationship between population  
1241 density and specific growth rate of continuous cultures. *Microbiology* 21, 40–50.

1242 Costa, A.R., Rodrigues, M.E., Henriques, M., Azeredo, J., Oliveira, R., 2010. Guide-  
1243 lines to cell engineering for monoclonal antibody production. *Eur. J. Pharm.*  
1244 *Biopharm.* 74, 127–138.

1245 Couceiro, M., Ghamisi, P., 2016. Particle swarm optimization, in: *Fractional order*  
1246 *darwinian particle swarm optimization*. Springer, pp. 1–10.

1247 Cox, S.M., Matthews, P.C., 2002. Exponential time differencing for stiff systems. *J.*  
1248 *Comput. Phys.* 176, 430–455.

1249 Craven, S., Whelan, J., Glennon, B., 2014. Glucose concentration control of a fed-  
1250 batch mammalian cell bioprocess using a nonlinear model predictive controller.  
1251 *Journal of Process Control* 24, 344–357.

1252 Dahlquist, G.G., 1963. A special stability problem for linear multistep methods.  
1253 *BIT Numer. Math.* 3, 27–43.

1254 Dang, C.V., 2010. Glutaminolysis: supplying carbon or nitrogen or both for cancer  
1255 cells? *Cell cycle* 9, 3884–3886.

1256 Das, S., Abraham, A., Konar, A., 2008. Particle swarm optimization and differential  
1257 evolution algorithms: technical analysis, applications and hybridization perspec-  
1258 tives, in: *Advances of computational intelligence in industrial systems*. Springer,  
1259 pp. 1–38.

1260 Dash, J., Dam, B., Swain, R., 2020. Design and implementation of sharp edge fir  
1261 filters using hybrid differential evolution particle swarm optimization. *AEU - Int.*  
1262 *J. Electron. Commun.* 114, 153019.

1263 Davidi, D., Milo, R., 2017. Lessons on enzyme kinetics from quantitative proteomics.  
1264 *Curr. Opin. Biotechnol.* 46, 81–89.

1265 Dawkins, C., Srinivasan, T.N., Whalley, J., 2001. Calibration, in: Handbook of  
1266 econometrics. Elsevier. volume 5, pp. 3653–3703.

1267 Dean, J., Reddy, P., 2013. Metabolic analysis of antibody producing cho cells in  
1268 fed-batch production. *Biotechnol. Bioeng.* 110, 1735–1747.

1269 Deb, K., 2014. Multi-objective optimization, in: Search methodologies. Springer,  
1270 pp. 403–449.

1271 Domenzain, I., Sánchez, B., Anton, M., Kerkhoven, E.J., Millán-Oropeza, A., Henry,  
1272 C., Siewers, V., Morrissey, J.P., Sonnenschein, N., Nielsen, J., 2021. Reconstruc-  
1273 tion of a catalogue of genome-scale metabolic models with enzymatic constraints  
1274 using gecko 2.0. *BioRxiv* .

1275 Domján, J., Pantea, E., Gyürkés, M., Madarász, L., Kozák, D., Farkas, A., Horváth,  
1276 B., Benkő, Z., Nagy, Z.K., Marosi, G., et al., 2022. Real-time amino acid and  
1277 glucose monitoring system for the automatic control of nutrient feeding in cho cell  
1278 culture using raman spectroscopy. *Biotechnol. J.* , 2100395.

1279 Dziwiński, P., Bartczuk, Ł., 2019. A new hybrid particle swarm optimization and  
1280 genetic algorithm method controlled by fuzzy logic. *IEEE Trans. Fuzzy Syst.* 28,  
1281 1140–1154.

1282 Edwards, V.H., 1970. The influence of high substrate concentrations on microbial  
1283 kinetics. *Biotechnol. Bioeng.* 12, 679–712.

1284 Eiben, Á.E., Hinterding, R., Michalewicz, Z., 1999. Parameter control in evolution-  
1285 ary algorithms. *IEEE Trans. Evol. Comput.* 3, 124–141.

1286 Esmin, A.A., Matwin, S., et al., 2013. Hpsom: a hybrid particle swarm optimization  
1287 algorithm with genetic mutation. *Int. J. Innov. Comput. Inf. Control.* 9, 1919–  
1288 1934.

1289 Fan, L., Kadura, I., Krebs, L.E., Hatfield, C.C., Shaw, M.M., Frye, C.C., 2012.  
1290 Improving the efficiency of cho cell line generation using glutamine synthetase  
1291 gene knockout cells. *Biotechnol. Bioeng.* 109, 1007–1015.

1292 Fan, L., Kadura, I., Krebs, L.E., Larson, J.L., Bowden, D.M., Frye, C.C., 2013.  
1293 Development of a highly-efficient cho cell line generation system with engineered  
1294 sv40e promoter. *J. Biotechnol.* 168, 652–658.

1295 Fathy, A., Abd Elaziz, M., Alharbi, A.G., 2020. A novel approach based on hy-  
1296 brid vortex search algorithm and differential evolution for identifying the optimal  
1297 parameters of pem fuel cell. *Renew. Energy* 146, 1833–1845.

1298 Fatunla, S.O., 2014. Numerical methods for initial value problems in ordinary dif-  
1299 ferential equations. Academic press.

1300 Feidl, F., Garbellini, S., Luna, M.F., Vogg, S., Souquet, J., Broly, H., Morbidelli,  
1301 M., Butté, A., 2019. Combining mechanistic modeling and raman spectroscopy  
1302 for monitoring antibody chromatographic purification. *Processes* 7, 683.

1303 Florez, D., Stéphan, A., Perré, P., Rémond, R., 2023. Probabilistic multi-objective  
1304 optimization of wood torrefaction conditions using a validated mechanistic model.  
1305 *Fuel* 335, 126932.

1306 Fu, X., Sun, Y., Wang, H., Li, H., 2021. Task scheduling of cloud computing based  
1307 on hybrid particle swarm algorithm and genetic algorithm. *Clust. Comput.* , 1–10.

1308 Fujimoto, Y., 1963. Kinetics of microbial growth and substrate consumption. *J.*  
1309 *Theor. Biol.* 5, 171–191.

1310 Galleguillos, S.N., Ruckerbauer, D., Gerstl, M.P., Borth, N., Hanscho, M.,  
1311 Zanghellini, J., 2017. What can mathematical modelling say about cho  
1312 metabolism and protein glycosylation? *Comput. Struct. Biotechnol. J.* 15, 212–  
1313 221.

1314 Garcia-Guarin, J., Rodriguez, D., Alvarez, D., Rivera, S., Cortes, C., Guzman, A.,  
1315 Bretas, A., Aguero, J.R., Bretas, N., 2019. Smart microgrids operation consider-  
1316 ing a variable neighborhood search: The differential evolutionary particle swarm  
1317 optimization algorithm. *Energies* 12, 3149.

1318 Georgioudakis, M., Plevris, V., 2020. A comparative study of differential evolution

1319 variants in constrained structural optimization. *Front. Built Environ.* 6, 102.

1320 Ghorbaniaghdam, A., Chen, J., Henry, O., Jolicoeur, M., 2014. Analyzing clonal  
1321 variation of monoclonal antibody-producing cho cell lines using an in silico  
1322 metabolomic platform. *PloS one* 9, e90832.

1323 Ghose, T., Tyagi, R., 1979. Rapid ethanol fermentation of cellulose hydrolysate. ii.  
1324 product and substrate inhibition and optimization of fermentor design. *Biotech-*  
1325 *nol. Bioeng.* 21, 1401–1420.

1326 Gianchandani, E.P., Chavali, A.K., Papin, J.A., 2010. The application of flux bal-  
1327 ance analysis in systems biology. *Wiley Interdiscip Rev Syst Biol Med* . 2, 372–382.

1328 González-Hernández, Y., Michiels, E., Perré, P., 2022. A comprehensive mechanistic  
1329 yeast model able to switch metabolism according to growth conditions. *Fermen-*  
1330 *tation* 8, 710.

1331 Goodarzimehr, V., Omidinasab, F., Taghizadieh, N., 2022. Optimum design of space  
1332 structures using hybrid particle swarm optimization and genetic algorithm. *World*  
1333 *J. Eng.* .

1334 Gutierrez, J.M., Feizi, A., Li, S., Kallehauge, T.B., Hefzi, H., Grav, L.M., Ley, D.,  
1335 Baycin Hizal, D., Betenbaugh, M.J., Voldborg, B., et al., 2020. Genome-scale  
1336 reconstructions of the mammalian secretory pathway predict metabolic costs and  
1337 limitations of protein secretion. *Nat. Commun.* 11, 68.

1338 Hallow, D.M., Mudryk, B.M., Braem, A.D., Tabora, J.E., Lyngberg, O.K., Bergum,  
1339 J.S., Rossano, L.T., Tummala, S., 2010. An example of utilizing mechanistic and  
1340 empirical modeling in quality by design. *Journal of Pharmaceutical Innovation* 5,  
1341 193–203.

1342 Handlogten, M.W., Zhu, M., Ahuja, S., 2018. Intracellular response of cho cells  
1343 to oxidative stress and its influence on metabolism and antibody production.  
1344 *Biochem. Eng. J.* 133, 12–20.

1345 Hao, Z.F., Guo, G.H., Huang, H., 2007. A particle swarm optimization algorithm

1346 with differential evolution, in: 2007 international conference on machine learning  
1347 and cybernetics, IEEE. pp. 1031–1035.

1348 Hartley, F., Walker, T., Chung, V., Morten, K., 2018. Mechanisms driving the  
1349 lactate switch in chinese hamster ovary cells. *Biotechnol. Bioeng.* 115, 1890–1903.

1350 Hefzi, H., Ang, K.S., Hanscho, M., Bordbar, A., Ruckerbauer, D., Lakshmanan, M.,  
1351 Orellana, C.A., Baycin-Hizal, D., Huang, Y., Ley, D., et al., 2016. A consen-  
1352 sus genome-scale reconstruction of chinese hamster ovary cell metabolism. *Cell*  
1353 *systems* 3, 434–443.

1354 Helmlinger, G., Al-Huniti, N., Aksenov, S., Peskov, K., Hallow, K.M., Chu, L.,  
1355 Boulton, D., Eriksson, U., Hamren, B., Lambert, C., et al., 2017. Drug-disease  
1356 modeling in the pharmaceutical industry-where mechanistic systems pharmacol-  
1357 ogy and statistical pharmacometrics meet. *European Journal of Pharmaceutical*  
1358 *Sciences* 109, S39–S46.

1359 Hendtlass, T., 2001. A combined swarm differential evolution algorithm for opti-  
1360 mization problems, in: *International conference on industrial, engineering and*  
1361 *other applications of applied intelligent systems*, Springer. pp. 11–18.

1362 Hochbruck, M., Ostermann, A., Schweitzer, J., 2009. Exponential rosenbrock-type  
1363 methods. *SIAM J. Numer. Anal.* 47, 786–803.

1364 Hong, J.K., Nargund, S., Lakshmanan, M., Kyriakopoulos, S., Kim, D.Y., Ang,  
1365 K.S., Leong, D., Yang, Y., Lee, D.Y., 2018. Comparative phenotypic analysis of  
1366 cho clones and culture media for lactate shift. *J. Biotechnol.* 283, 97–104.

1367 Houssein, E.H., Gad, A.G., Hussain, K., Suganthan, P.N., 2021. Major advances  
1368 in particle swarm optimization: theory, analysis, and application. *Swarm Evol.*  
1369 *Comput.* 63, 100868.

1370 Huang, Z., Lee, D.Y., Yoon, S., 2017. Quantitative intracellular flux modeling and  
1371 applications in biotherapeutic development and production using cho cell cultures.  
1372 *Biotechnol. Bioeng.* 114, 2717–2728.

1373 Ivarsson, M., Noh, H., Morbidelli, M., Soos, M., 2015. Insights into ph-induced  
1374 metabolic switch by flux balance analysis. *Biotechnol. Progr.* 31, 347–357.

1375 Katoch, S., Chauhan, S.S., Kumar, V., 2021. A review on genetic algorithm: past,  
1376 present, and future. *Multimed. Tools Appl.* 80, 8091–8126.

1377 Kauffman, K.J., Prakash, P., Edwards, J.S., 2003. Advances in flux balance analysis.  
1378 *Curr. Opin. Biotechnol.* 14, 491–496.

1379 Kennedy, J., Eberhart, R., 1995. Particle swarm optimization, in: *Proceedings of*  
1380 *ICNN'95-international conference on neural networks*, IEEE. pp. 1942–1948.

1381 Khaparde, A., Alassery, F., Kumar, A., Alotaibi, Y., Khalaf, O., Pillai, S., Al-  
1382 ghamdi, S., 2022. Differential evolution algorithm with hierarchical fair competi-  
1383 tion model. *Intell. Autom. Soft Comput.* 33, 1045–1062.

1384 Kim, M.S., Kim, N.S., Sung, Y.H., Lee, G.M., 2002. Biphasic culture strategy based  
1385 on hyperosmotic pressure for improved humanized antibody production in chinese  
1386 hamster ovary cell culture. *In Vitro Cell. Dev. Biol.* 38, 314–319.

1387 Kingston, R.E., Kaufman, R.J., Bebbington, C., Rolfe, M., 2002. Amplification  
1388 using cho cell expression vectors. *Curr. Protoc. Mol. Biol.* 60, 16–23.

1389 Kotidis, P., Jedrzejewski, P., Sou, S.N., Sellick, C., Polizzi, K., Del Val, I.J., Kon-  
1390 toravdi, C., 2019. Model-based optimization of antibody galactosylation in cho  
1391 cell culture. *Biotechnol. Bioeng.* 116, 1612–1626.

1392 Kuepfer, L., Lippert, J., Eissing, T., 2012. Multiscale mechanistic modeling in  
1393 pharmaceutical research and development. *Advances in Systems Biology* , 543–  
1394 561.

1395 Kumar, S., Nayyar, A., Paul, A., 2019. *Swarm intelligence and evolutionary algo-*  
1396 *rithms in healthcare and drug development.* CRC Press.

1397 Kuriakose, A., Chirmule, N., Nair, P., 2016. Immunogenicity of biotherapeutics:  
1398 causes and association with posttranslational modifications. *J. Immunol. Res.*  
1399 2016.



- 1400 Kyriakopoulos, S., Ang, K.S., Lakshmanan, M., Huang, Z., Yoon, S., Gunawan, R.,  
1401 Lee, D.Y., 2018. Kinetic modeling of mammalian cell culture bioprocessing: the  
1402 quest to advance biomanufacturing. *Biotechnol. J.* 13, 1700229.
- 1403 Kyriakopoulos, S., Kontoravdi, C., 2014. A framework for the systematic design of  
1404 fed-batch strategies in mammalian cell culture. *Biotechnol. Bioeng.* 111, 2466–  
1405 2476.
- 1406 La, A., Du, H., Taidi, B., Perre, P., 2020. A predictive dynamic yeast model based  
1407 on component, energy, and electron carrier balances. *Biotechnol. Bioeng.* 117,  
1408 2728–2740.
- 1409 Lebedev, V., 2017. How to solve stiff systems of differential equations by explicit  
1410 methods, in: *Numerical methods and applications*. CRC Press, pp. 45–80.
- 1411 Levenspiel, O., 1980. The monod equation: a revisit and a generalization to product  
1412 inhibition situations. *Biotechnol. Bioeng.* 22, 1671–1687.
- 1413 Levintow, L., Meister, A., 1954. Reversibility of the enzymatic synthesis of glu-  
1414 tamine. *J. Biol. Chem.* 209, 265–280.
- 1415 Li, H., Xiang, S., Yang, Y., Liu, C., 2021. Differential evolution particle swarm  
1416 optimization algorithm based on good point set for computing nash equilibrium  
1417 of finite noncooperative game. *AIMS Math.* 6, 1309–1323.
- 1418 Li, M.Y., Ebel, B., Paris, C., Chauchard, F., Guedon, E., Marc, A., 2018. Real-time  
1419 monitoring of antibody glycosylation site occupancy by in situ raman spectroscopy  
1420 during bioreactor cho cell cultures. *Biotechnol. Progr.* 34, 486–493.
- 1421 Lilla, A., Khan, M., Barendse, P., 2013. Comparison of differential evolution and  
1422 genetic algorithm in the design of permanent magnet generators, in: *2013 IEEE*  
1423 *International Conference on Industrial Technology (ICIT)*, pp. 266–271.
- 1424 Liste-Calleja, L., Lecina, M., Lopez-Repullo, J., Albiol, J., Solà, C., Cairó, J.J., 2015.  
1425 Lactate and glucose concomitant consumption as a self-regulated ph detoxification  
1426 mechanism in hek293 cell cultures. *Appl. Microbiol. Biotechnol.* 99, 9951–9960.

1427 López-Meza, J., Araíz-Hernández, D., Carrillo-Cocom, L.M., López-Pacheco, F., del  
1428 Refugio Rocha-Pizaña, M., Alvarez, M.M., 2016. Using simple models to describe  
1429 the kinetics of growth, glucose consumption, and monoclonal antibody formation  
1430 in naive and infliximab producer cho cells. *Cytotechnology* 68, 1287–1300.

1431 Luo, Y., Kurian, V., Ogunnaike, B.A., 2021. Bioprocess systems analysis, modeling,  
1432 estimation, and control. *Curr. Opin. Chem. Eng.* 33, 100705.

1433 Luong, J., 1987. Generalization of monod kinetics for analysis of growth data with  
1434 substrate inhibition. *Biotechnol. Bioeng.* 29, 242–248.

1435 MacDonald, M.A., Nöbel, M., Roche Recinos, D., Martínez, V.S., Schulz, B.L.,  
1436 Howard, C.B., Baker, K., Shave, E., Lee, Y.Y., Marcellin, E., et al., 2022. Perfu-  
1437 sion culture of chinese hamster ovary cells for bioprocessing applications. *Critical*  
1438 *reviews in biotechnology* 42, 1099–1115.

1439 Malthus, T.R., 1986. An essay on the principle of population (1798). *The Works of*  
1440 *Thomas Robert Malthus*, London, Pickering & Chatto Publishers 1, 1–139.

1441 Martínez, V.S., Buchsteiner, M., Gray, P., Nielsen, L.K., Quek, L.E., 2015. Dynamic  
1442 metabolic flux analysis using b-splines to study the effects of temperature shift  
1443 on cho cell metabolism. *Metabolic Engineering Communications* 2, 46–57.

1444 Martínez, V.S., Dietmair, S., Quek, L.E., Hodson, M.P., Gray, P., Nielsen, L.K.,  
1445 2013. Flux balance analysis of cho cells before and after a metabolic switch from  
1446 lactate production to consumption. *Biotechnol. Bioeng.* 110, 660–666.

1447 Martínez, V.S., Saa, P.A., Jooste, J., Tiwari, K., Quek, L.E., Nielsen, L.K., 2022.  
1448 The topology of genome-scale metabolic reconstructions unravels independent  
1449 modules and high network flexibility. *PLOS Computational Biology* 18, e1010203.

1450 Martínez-Monge, I., Comas, P., Triquell, J., Casablanco, A., Lecina, M., Paredes,  
1451 C., Cairó, J., 2019. Concomitant consumption of glucose and lactate: A novel  
1452 batch production process for cho cells. *Biochem. Eng. J.* 151, 107358.

1453 Marx, N., Eisenhut, P., Weinguny, M., Klanert, G., Borth, N., 2022. How to train

1454 your cell-towards controlling phenotypes by harnessing the epigenome of chinese  
1455 hamster ovary production cell lines. *Biotechnol. Adv.* , 107924.

1456 Matasci, M., Hacker, D.L., Baldi, L., Wurm, F.M., 2008. Recombinant therapeu-  
1457 tic protein production in cultivated mammalian cells: current status and future  
1458 prospects. *Drug Discov. Today Technol.* 5, e37–e42.

1459 McHugh, K.P., Xu, J., Aron, K.L., Borys, M.C., Li, Z.J., 2020. Effective temperature  
1460 shift strategy development and scale confirmation for simultaneous optimization  
1461 of protein productivity and quality in chinese hamster ovary cells. *Biotechnol.*  
1462 *Prog.* 36, e2959.

1463 Min Lee, G., Koo, J., 2009. Osmolarity effects, chinese hamster ovary cell cul-  
1464 ture. *Encyclopedia of industrial biotechnology: bioprocess, bioseparation, and*  
1465 *cell technology* , 1–8.

1466 Moser, A., 1988. The principles of bioprocess technol., in: *Bioprocess Technol.*  
1467 Springer, pp. 13–65.

1468 Mousa, A., El-Shorbagy, M.A., Abd-El-Wahed, W.F., 2012. Local search based  
1469 hybrid particle swarm optimization algorithm for multiobjective optimization.  
1470 *Swarm Evol. Comput.* 3, 1–14.

1471 Mowbray, M.R., Wu, C., Rogers, A.W., Rio-Chanona, E.A.D., Zhang, D., 2023. A  
1472 reinforcement learning-based hybrid modeling framework for bioprocess kinetics  
1473 identification. *Biotechnology and Bioengineering* 120, 154–168.

1474 Mulchandani, A., Luong, J., 1989. Microbial inhibition kinetics revisited. *Enzyme*  
1475 *Microb. Technol.* 11, 66–73.

1476 Mulukutla, B.C., Gramer, M., Hu, W.S., 2012. On metabolic shift to lactate con-  
1477 sumption in fed-batch culture of mammalian cells. *Metab. Eng.* 14, 138–149.

1478 Narayanan, H., Luna, M.F., von Stosch, M., Cruz Bournazou, M.N., Polotti, G.,  
1479 Morbidelli, M., Butté, A., Sokolov, M., 2020. Bioprocessing in the digital age:  
1480 the role of process models. *Biotechnology journal* 15, 1900172.

1481 Narayanan, H., Seidler, T., Luna, M.F., Sokolov, M., Morbidelli, M., Butté, A., 2021.  
1482 Hybrid models for the simulation and prediction of chromatographic processes for  
1483 protein capture. *Journal of Chromatography A* 1650, 462248.

1484 Narayanan, H., Sokolov, M., Morbidelli, M., Butté, A., 2019. A new generation of  
1485 predictive models: The added value of hybrid models for manufacturing processes  
1486 of therapeutic proteins. *Biotechnology and bioengineering* 116, 2540–2549.

1487 Noh, S.M., Sathyamurthy, M., Lee, G.M., 2013. Development of recombinant chinese  
1488 hamster ovary cell lines for therapeutic protein production. *Curr. Opin. Chem.* 2,  
1489 391–397.

1490 Noh, S.M., Shin, S., Lee, G.M., 2018. Comprehensive characterization of glutamine  
1491 synthetase-mediated selection for the establishment of recombinant cho cells pro-  
1492 ducing monoclonal antibodies. *Sci. Rep.* 8, 1–11.

1493 Nolan, R.P., Lee, K., 2011. Dynamic model of cho cell metabolism. *Metab. Eng.*  
1494 13, 108–124.

1495 Nolan, R.P., Lee, K., 2012. Dynamic model for cho cell engineering. *J. Biotechnol.*  
1496 158, 24–33.

1497 Omidinasab, F., Goodarzimehr, V., 2020. A hybrid particle swarm optimization and  
1498 genetic algorithm for truss structures with discrete variables. *J. Appl. Comput.*  
1499 *Mech.* 6, 593–604.

1500 Orth, J.D., Thiele, I., Palsson, B.Ø., 2010. What is flux balance analysis? *Nat.*  
1501 *Biotechnol.* 28, 245–248.

1502 Pan, X., Dalm, C., Wijffels, R.H., Martens, D.E., 2017a. Metabolic characterization  
1503 of a cho cell size increase phase in fed-batch cultures. *Applied microbiology and*  
1504 *biotechnology* 101, 8101–8113.

1505 Pan, X., Streefland, M., Dalm, C., Wijffels, R.H., Martens, D.E., 2017b. Selection of  
1506 chemically defined media for cho cell fed-batch culture processes. *Cytotechnology*  
1507 69, 39–56.

1508 Paul, K., Rajamanickam, V., Herwig, C., 2019. Model-based optimization of tem-  
1509 perature and ph shift to increase volumetric productivity of a chinese hamster  
1510 ovary fed-batch process. *Journal of bioscience and bioengineering* 128, 710–715.

1511 Pereira, S., Kildegaard, H.F., Andersen, M.R., 2018. Impact of cho metabolism on  
1512 cell growth and protein production: an overview of toxic and inhibiting metabo-  
1513 lites and nutrients. *Biotechnol. J.* 13, 1700499.

1514 Pozzobon, V., Levasseur, W., Guerin, C., Perré, P., 2021. Nitrate and nitrite as  
1515 mixed source of nitrogen for *Chlorella vulgaris*: fast nitrogen quantification using  
1516 spectrophotometer and machine learning. *J. Appl. Phycol.* 33, 1389–1397.

1517 Price, N.D., Reed, J.L., Palsson, B.Ø., 2004. Genome-scale models of microbial  
1518 cells: evaluating the consequences of constraints. *Nature Reviews Microbiology*  
1519 2, 886–897.

1520 Pu, H., Song, T., Schonfeld, P., Li, W., Zhang, H., Hu, J., Peng, X., Wang, J., 2019.  
1521 Mountain railway alignment optimization using stepwise & hybrid particle swarm  
1522 optimization incorporating genetic operators. *Appl. Soft Comput.* 78, 41–57.

1523 Rajamanickam, V., Babel, H., Montano-Herrera, L., Ehsani, A., Stiefel, F., Haider,  
1524 S., Presser, B., Knapp, B., 2021. About model validation in bioprocessing. *Pro-  
1525 cesses* 9, 961.

1526 Reinhart, D., Damjanovic, L., Kaisermayer, C., Kunert, R., 2015. Benchmarking  
1527 of commercially available cho cell culture media for antibody production. *Appl.  
1528 Microbiol. Biotechnol.* 99, 4645–4657.

1529 Ritacco, F.V., Wu, Y., Khetan, A., 2018. Cell culture media for recombinant protein  
1530 expression in chinese hamster ovary (cho) cells: history, key components, and  
1531 optimization strategies. *Biotechnol. Progr.* 34, 1407–1426.

1532 Rodrigues, M.E., Costa, A.R., Henriques, M., Azeredo, J., Oliveira, R., 2012. Com-  
1533 parison of commercial serum-free media for cho-k1 cell growth and monoclonal  
1534 antibody production. *Int. J. Pharm.* 437, 303–305.

1535 Romann, P., Kolar, J., Tobler, D., Herwig, C., Bielser, J.M., Villiger, T.K., 2022.  
1536 Advancing raman model calibration for perfusion bioprocesses using spiked har-  
1537 vest libraries. *Biotechnol. J.* , 2200184.

1538 Romanova, N., Schmitz, J., Strakeljahn, M., Grünberger, A., Bahnemann, J.,  
1539 Noll, T., 2022. Single-cell analysis of cho cells reveals clonal heterogeneity in  
1540 hyperosmolality-induced stress response. *Cells* 11, 1763.

1541 Samuel, G.G., Rajan, C.C.A., 2015. Hybrid: particle swarm optimization–genetic  
1542 algorithm and particle swarm optimization–shuffled frog leaping algorithm for  
1543 long-term generator maintenance scheduling. *Int. J. Electr. Power Energy Syst.*  
1544 65, 432–442.

1545 Sánchez, B.J., Zhang, C., Nilsson, A., Lahtvee, P.J., Kerkhoven, E.J., Nielsen, J.,  
1546 2017. Improving the phenotype predictions of a yeast genome-scale metabolic  
1547 model by incorporating enzymatic constraints. *Mol. Syst. Biol.* 13, 935.

1548 Santos, R.M., Kessler, J.M., Salou, P., Menezes, J.C., Peinado, A., 2018. Monitoring  
1549 mab cultivations with in-situ raman spectroscopy: The influence of spectral se-  
1550 lectivity on calibration models and industrial use as reliable pat tool. *Biotechnol.*  
1551 *Progr.* 34, 659–670.

1552 Schinn, S.M., Morrison, C., Wei, W., Zhang, L., Lewis, N.E., 2021. A genome-scale  
1553 metabolic network model and machine learning predict amino acid concentrations  
1554 in chinese hamster ovary cell cultures. *Biotechnol. Bioeng.* 118, 2118–2123.

1555 Schwarz, H., Mäkinen, M.E., Castan, A., Chotteau, V., 2022. Monitoring of amino  
1556 acids and antibody n-glycosylation in high cell density perfusion culture based on  
1557 raman spectroscopy. *Biochem. Eng. J.* 182, 108426.

1558 Sedighizadeh, D., Masehian, E., 2009. Particle swarm optimization methods, tax-  
1559 onomy and applications. *Int. J. Comput. Sci. Eng.* 1, 486.

1560 Selişteanu, D., Petre, E., Răsvan, V.B., 2007. Sliding mode and adaptive sliding-  
1561 mode control of a class of nonlinear bioprocesses. *Int. J. Adapt. Control Signal*

1562 Process. 21, 795–822.

1563 Sengupta, N., Rose, S.T., Morgan, J.A., 2011. Metabolic flux analysis of cho cell  
1564 metabolism in the late non-growth phase. *Biotechnol. Bioeng.* 108, 82–92.

1565 Sha, S., Huang, Z., Wang, Z., Yoon, S., 2018. Mechanistic modeling and applications  
1566 for cho cell culture development and production. *Curr. Opin. Chem. Eng.* 22, 54–  
1567 61.

1568 Shirsat, N., Mohd, A., Whelan, J., English, N.J., Glennon, B., Al-Rubeai, M., 2015.  
1569 Revisiting verhulst and monod models: analysis of batch and fed-batch cultures.  
1570 *Cytotechnology* 67, 515–530.

1571 Sin, G., De Pauw, D.J., Weijers, S., Vanrolleghem, P.A., 2008. An efficient approach  
1572 to automate the manual trial and error calibration of activated sludge models.  
1573 *Biotechnology and bioengineering* 100, 516–528.

1574 Snape, J.B., Dunn, I.J., Ingham, J., Přenosil, J.E., 2008. Dynamics of environmental  
1575 bioProcesses: modelling and simulation. John Wiley & Sons.

1576 Srinath, E.G., Loehr, R.C., 1974. Ammonia desorption by diffused aeration. *J.*  
1577 *Water Pollut. Control Fed.* , 1939–1957.

1578 Talbi, H., Batouche, M., 2004. Hybrid particle swarm with differential evolution  
1579 for multimodal image registration, in: 2004 IEEE International Conference on  
1580 Industrial Technology, 2004. IEEE ICIT'04., IEEE. pp. 1567–1572.

1581 Teissier, G., et al., 1936. Quantitative laws of growth. *Ann. Physiol. Physicochim.*  
1582 *Biol.* 12, 527–586.

1583 Templeton, N., Dean, J., Reddy, P., Young, J.D., 2013. Peak antibody production  
1584 is associated with increased oxidative metabolism in an industrially relevant fed-  
1585 batch cho cell culture. *Biotechnol. Bioeng.* 110, 2013–2024.

1586 Thangaraj, R., Pant, M., Abraham, A., Bouvry, P., 2011. Particle swarm opti-  
1587 mization: hybridization perspectives and experimental illustrations. *Appl. Math.*  
1588 *Comput.* 217, 5208–5226.

1589 Tokman, M., 2006. Efficient integration of large stiff systems of odes with exponential  
1590 propagation iterative (epi) methods. *J. Comput. Phys.* 213, 748–776.

1591 Torres, M., Zúñiga, R., Gutierrez, M., Vergara, M., Collazo, N., Reyes, J., Berrios,  
1592 J., Aguilon, J.C., Molina, M.C., Altamirano, C., 2018. Mild hypothermia upreg-  
1593 ulates myc and xbp1s expression and improves anti-tnf $\alpha$  production in cho cells.  
1594 *PloS one* 13, e0194510.

1595 Trivedi, A., Srinivasan, D., Biswas, S., Reindl, T., 2015. Hybridizing genetic al-  
1596 gorithm with differential evolution for solving the unit commitment scheduling  
1597 problem. *Swarm Evol. Comput.* 23, 50–64.

1598 Trivedi, A., Srinivasan, D., Biswas, S., Reindl, T., 2016. A genetic algorithm–  
1599 differential evolution based hybrid framework: case study on unit commitment  
1600 scheduling problem. *Inf. Sci.* 354, 275–300.

1601 Tsao, Y.S., Cardoso, A., Condon, R., Voloch, M., Lio, P., Lagos, J., Kearns, B., Liu,  
1602 Z., 2005. Monitoring chinese hamster ovary cell culture by the analysis of glucose  
1603 and lactate metabolism. *J. Biotechnol.* 118, 316–327.

1604 Tsopanoglou, A., del Val, I.J., 2021. Moving towards an era of hybrid modelling:  
1605 advantages and challenges of coupling mechanistic and data-driven models for  
1606 upstream pharmaceutical bioprocesses. *Curr. Opin. Chem. Eng.* 32, 100691.

1607 Vaisakh, K., Sridhar, M., Murthy, K.L., 2009. Differential evolution particle swarm  
1608 optimization algorithm for reduction of network loss and voltage instability, in:  
1609 2009 World Congress on Nature & Biologically Inspired Computing (NaBIC),  
1610 IEEE. pp. 391–396.

1611 Jimenez del Val, I., Fan, Y., Weilguny, D., 2016. Dynamics of immature mab  
1612 glycoform secretion during cho cell culture: An integrated modelling framework.  
1613 *Biotechnol. J.* 11, 610–623.

1614 Venter, G., Sobieszczanski-Sobieski, J., 2003. Particle swarm optimization. *AIAA*  
1615 *J.* 41, 1583–1589.



1616 Verhulst, P.F., 1838. Notice sur la loi que la population suit dans son accroissement.  
1617 Corresp. Math. Phys. 10, 113–129.

1618 Vesterstrom, J., Thomsen, R., 2004. A comparative study of differential evolution,  
1619 particle swarm optimization, and evolutionary algorithms on numerical bench-  
1620 mark problems, in: Proceedings of the 2004 congress on evolutionary computation  
1621 (IEEE Cat. No. 04TH8753), IEEE. pp. 1980–1987.

1622 Villaverde, A.F., Pathirana, D., Fröhlich, F., Hasenauer, J., Banga, J.R., 2022. A  
1623 protocol for dynamic model calibration. Briefings in Bioinformatics 23, bbab387.

1624 W Eyster, T., Talwar, S., Fernandez, J., Foster, S., Hayes, J., Allen, R., Reidinger,  
1625 S., Wan, B., Ji, X., Aon, J., et al., 2021. Tuning monoclonal antibody galactosy-  
1626 lation using raman spectroscopy-controlled lactic acid feeding. Biotechnol. Progr.  
1627 37, e3085.

1628 Wahrheit, J., Niklas, J., Heinzle, E., 2014. Metabolic control at the cytosol-  
1629 mitochondria interface in different growth phases of cho cells. Metab. Eng. 23,  
1630 9–21.

1631 Wang, S.C., 2003. Genetic algorithm, in: Interdisciplinary computing in java pro-  
1632 gramming. Springer, pp. 101–116.

1633 Webb, J.L., 1963. Enzyme and metabolic inhibitors .

1634 Xing, Z., Bishop, N., Leister, K., Li, Z.J., 2010. Modeling kinetics of a large-scale  
1635 fed-batch cho cell culture by markov chain monte carlo method. Biotechnol. Progr.  
1636 26, 208–219.

1637 Xing, Z., Nguyen, T.B., Kanai-Bai, G., Yamano-Adachi, N., Omasa, T., 2023. Con-  
1638 struction of a novel kinetic model for the production process of a cva6 vlp vaccine  
1639 in cho cells. Cytotechnology , 1–15.

1640 Xu, P., 2020. Analytical solution for a hybrid logistic-monod cell growth model  
1641 in batch and continuous stirred tank reactor culture. Biotechnol. Bioeng. 117,  
1642 873–878.

1643 Xu, S., Jiang, R., Mueller, R., Hoesli, N., Kretz, T., Bowers, J., Chen, H., 2018.  
1644 Probing lactate metabolism variations in large-scale bioreactors. *Biotechnol.*  
1645 *Progr.* 34, 756–766.

1646 Yahia, B.B., Malphettes, L., Heinzle, E., 2021. Predictive macroscopic modeling of  
1647 cell growth, metabolism and monoclonal antibody production: Case study of a  
1648 cho fed-batch production. *Metab. Eng.* 66, 204–216.

1649 Yang, N., Guerin, C., Kokanyan, N., Perré, P., 2024. Raman spectroscopy applied  
1650 to online monitoring of a bioreactor: Tackling the limit of detection. *Spectrochim.*  
1651 *Acta A Mol. Biomol. Spectrosc.* 304, 123343.

1652 Yang, R., Chen, J., 2021. Mechanistic and machine learning modeling of microwave  
1653 heating process in domestic ovens: A review. *Foods* 10, 2029.

1654 Yang, W., Zhang, J., Xiao, Y., Li, W., Wang, T., 2022. Screening strategies for  
1655 high-yield chinese hamster ovary cell clones. *Front. Bioeng. Biotechnol.* 10.

1656 Yano, T., Koga, S., 1969. Dynamic behavior of the chemostat subject to substrate  
1657 inhibition. *Biotechnol. Bioeng.* 11, 139–153.

1658 Yilmaz, D., Mehdizadeh, H., Navarro, D., Shehzad, A., O'Connor, M., McCormick,  
1659 P., 2020. Application of raman spectroscopy in monoclonal antibody producing  
1660 continuous systems for downstream process intensification. *Biotechnol. Progr.* 36,  
1661 e2947.

1662 Yousefi-Darani, A., Paquet-Durand, O., Von Wrochem, A., Classen, J., Tränkle, J.,  
1663 Mertens, M., Snelders, J., Chotteau, V., Mäkinen, M., Handl, A., et al., 2022.  
1664 Generic chemometric models for metabolite concentration prediction based on  
1665 raman spectra. *Sensors* 22, 5581.

1666 Yu, X., Gen, M., 2010. Introduction to evolutionary algorithms. Springer Science  
1667 & Business Media.

1668 Zagari, F., Jordan, M., Stettler, M., Broly, H., Wurm, F.M., 2013. Lactate  
1669 metabolism shift in cho cell culture: the role of mitochondrial oxidative activ-

1670 ity. *New Biotechnol.* 30, 238–245.

1671 Zalai, D., Koczka, K., Párta, L., Wechselberger, P., Klein, T., Herwig, C., 2015.

1672 Combining mechanistic and data-driven approaches to gain process knowledge on

1673 the control of the metabolic shift to lactate uptake in a fed-batch cho process.

1674 *Biotechnol. Progr.* 31, 1657–1668.

1675 Zhang, J., Petersen, S.D., Radivojevic, T., Ramirez, A., Pérez-Manríquez, A.,

1676 Abeliuk, E., Sánchez, B.J., Costello, Z., Chen, Y., Fero, M.J., et al., 2020. Com-

1677 bining mechanistic and machine learning models for predictive engineering and

1678 optimization of tryptophan metabolism. *Nat. commun.* 11, 1–13.

1679 Zhang, J.H., Shan, L.L., Liang, F., Du, C.Y., Li, J.J., 2022. Strategies and consid-

1680 erations for improving recombinant antibody production and quality in chinese

1681 hamster ovary cells. *Front. Bioeng. Biotechnol.* 10.

1682 Zhang, W.J., Xie, X.F., 2003. Depso: hybrid particle swarm with differential evo-

1683 lution operator, in: *SMC'03 Conference Proceedings. 2003 IEEE International*

1684 *Conference on Systems, Man and Cybernetics. Conference Theme-System Secu-*

1685 *rity and Assurance (Cat. No. 03CH37483), IEEE.* pp. 3816–3821.

1686 Zhang, X., Garcia, I.F., Baldi, L., Hacker, D.L., Wurm, F.M., 2010. Hyperosmolar-

1687 ity enhances transient recombinant protein yield in chinese hamster ovary cells.

1688 *Biotechnol. Lett.* 32, 1587–1592.