Developing a Computational Framework To Advance Bioprocess Scale-Up

Wang, Guan; Haringa, Cees; Noorman, Henk; Chu, Ju; Zhuang, Yingping

DOI
10.1016/j.tibtech.2020.01.009

Publication date
2020

Document Version
Proof

Published in
Trends in Biotechnology

Citation (APA)
Wang, G., Haringa, C., Noorman, H., Chu, J., & Zhuang, Y. (2020). Developing a Computational Framework To Advance Bioprocess Scale-Up. *Trends in Biotechnology*, *38*(8), 846-856. https://doi.org/10.1016/j.tibtech.2020.01.009

Trends in Biotechnology

CellPress REVIEWS

Opinion

Developing a Computational Framework To Advance Bioprocess Scale-Up

Guan Wang,[1,*] Cees Haringa,[2] Henk Noorman,[2,3] Ju Chu,[1] and Yingping Zhuang[1,*]

Bioprocess scale-up is a critical step in process development. However, loss of production performance upon scaling-up, including reduced titer, yield, or productivity, has often been observed, hindering the commercialization of biotech innovations. Recent developments in scale-down studies assisted by computational fluid dynamics (CFD) and powerful stimulus–response metabolic models afford better process prediction and evaluation, enabling faster scale-up with minimal losses. In the future, an ideal bioprocess design would be guided by an *in silico* model that integrates cellular physiology (spatiotemporal multiscale cellular models) and fluid dynamics (CFD models). Nonetheless, there are challenges associated with both establishing predictive metabolic models and CFD coupling. By highlighting these and providing possible solutions here, we aim to advance the development of a computational framework to accelerate bioprocess scale-up.

Highlights

Metabolomics has been used alone or combined with other 'omics' tools to contribute to industrial systems biology and synthetic biology.

Mechanisms governing population heterogeneity in industrial bioprocessing need rigorous investigation, and thereafter should be incorporated into more powerful cellular models.

A predictive scale-down model should be rationally guided by model fluid studies in large-scale bioprocesses.

Coupling cellular kinetics with fluid dynamics accelerates science-based design of both microbial cell factories and industrial-scale bioreactors.

Potential and Challenges of Computer-Aided Bioprocess Design

Worldwide, the number of industrial bioprocesses is steadily growing because of larger market demand for the bulk production of chemicals, biofuels, materials, nutrition ingredients, and healthcare products. This demand accompanies the urgent need to transition from a fossil-based to a bio-based economy [1–3]. The growing markets come with an increasing need for industrial-scale bioprocesses, together with a demand for large and efficient bioreactors. During development, one big challenge during the transfer from the lab scale to the industrial scale has long been the so-called 'scale-up effect', which is often accompanied by reduced commercial indicators (by 10–30%), such as biomass and product–substrate yield, as well as productivity [4]. This reduction in performance is mostly associated with mass and heat transfer in addition to mixing issues in industrial-scale bioreactors, wherein a subpopulation of microbes with reduced production efficiency and capacity are formed, a phenomenon often referred to as 'population heterogeneity' [5]. In microbial bioprocesses, this omnipresent phenomenon originates from both intrinsic (the stochastic nature of gene expression) and extrinsic factors (variations in environmental conditions) [5,6]. Intrinsic factors dictate that, even under the same environmental conditions, cells may behave differently [7–11]. Extrinsic factors show that the emergence of (temporary) slower rates (e.g., μ and q_P) is mostly ascribed to a lack of substrate or oxygen, or the onset of overflow metabolism at high substrate availability in large-scale containers [4,11–13]. Consequently, elucidating these up-scaling related mechanisms (e.g., phenotypic heterogeneity) for reducing or circumventing performance loss is not only an academic goal, but also urgently needed in industrial practices [14].

Bioreactor scaling-up from laboratory-scale (milliliter to tens of liters) to large-scale (hundreds to thousands of cubic meters) requires proof of comparability of process and of product quality, which traditionally follows a time-consuming, lab-intensive, stepwise scale-up scheme. This mostly depends on expert empiricism and does not consider possible changes of strain

[1]State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology (ECUST), Shanghai, People's Republic of China
[2]DSM Biotechnology Center, Delft, The Netherlands
[3]Department of Biotechnology, Delft University of Technology, Delft, The Netherlands

*Correspondence:
guanwang@ecust.edu.cn (G. Wang) and ypzhuang@ecust.edu.cn (Y. Zhuang).

physiology in the gas–liquid flow field and substrate, and/or oxygen concentration field of the industrial reactor. Within the biopharmaceutical industry, to ensure cost-effective, reproducible, high-quality production, guiding principles, such as **process analytical technology** (PAT) (see Glossary) and **quality by design** (QbD), have been launched and advocated by the US Food and Drug Administration (FDA) to allow an improved understanding and control of the process and, thus, quality-guaranteed biomanufacturing [15]. Thanks to the development of in-line or at-line process biosensors and real-time data visualization, a plethora of studies have reported the identification of the rate-limiting factor through process data-association analysis [16–18]. In addition, cross-scale insights can be obtained if process investigation and reasoning go from macro- to microscale and from extra- to intracellular (Box 1).

However, to allow science-based design of industrial workhorses and bioreactors (i.e., in a more cost-constrained setting outside the biopharmaceutical space), challenges associated with bioreactor scale-up remain. In particular, to fully benefit from a down-scale optimization strategy, where industrial conditions are mimicked in the lab, lack of data sets from the industrial scale remains a significant hurdle. To bridge the gap from promise to practice, **CFD** models can be applied to generate detailed flow field information (e.g., flow pattern, mixing time, and concentration profiles), which provides suggested advice on process optimization and bioreactor configuration (Box 2) [19–22]. Furthermore, by tracking single cell lifelines and assessing the production performance at a large scale, pioneering work by Lapin and coworkers showed that the **Euler–Lagrange CFD method** can be applied to evaluate the dynamic behavior of yeast cells [23] and the effect of glucose gradients on the metabolic response in *Escherichia coli* with simple glucose uptake kinetics [24]. To capture more detailed information about the system, kinetic models should be set up with kinetic regulation information (e.g., a gene-regulation model for product formation) [25]. The principle governing the model establishment should be to keep models simple, but detailed enough to preserve the critical reaction kinetics [26]. In addition, models should be sufficiently computationally efficient to allow coupling with CFD models and reasonable simulation time.

In the future, science-based design of industrial processes and bioreactors (involving e.g., strain characteristics, reactor configuration, and operating parameters) as well as prior global assessment of process performance across the scales can be achieved by coupling cellular kinetics and fluid dynamics (Figure 1, Key Figure). The coupled full-scale predictive model can then generate more valuable information to guide decision-making for intelligent biomanufacturing, which fulfills the 'Industry 4.0' idea toward digitalization and automation [27]. Nonetheless, challenges associated with the set-up of kinetic models with high predictive ability and integration of those with CFD models still require general and robust solutions.

Metabolic Model Reduction by Metabolite Lumping/Pooling

Different from stoichiometric models (e.g., genome-scale metabolic models), kinetic models can provide dynamic behavior of cells in response to genetic and/or environmental perturbations. To allow a more detailed kinetic description of the cell, incorporating more biochemical mechanisms (e.g., substrate uptake, production formation, futile cycles, carbohydrate storage/mobilization, and induction of key enzymes) is obvious; nevertheless, there are challenges related with parameterization of large kinetic models, which are caused by: (i) the lack of representative metabolic data reflecting cellular regulation responding to dynamic environments present in large-scale bioreactors; (ii) the insensitivity of kinetic parameters within a nonlinear kinetic structure under tested conditions; (iii) lack of informed design-of-experiments studies that allow parameter identification and estimation; and (iv) too little use of data reconciliation methods to minimize uncertainty and maximize consistency of experimental data. As an alternative, kinetic models should be

Glossary

Compartment model: spatially resolved description of a reactor that is generally coarser than a CFD representation; in a compartment model, flows are using mapped from experimental/CFD data rather than calculated.
Computational fluid dynamics (CFD): aims to analyze and solve fluid flow-related problems using numerical analysis and data structures.
Critical process parameters (CPPs): key variables monitored in pharmaceutical manufacturing to show deviations in standard production operations as well as in product quality or changes in critical quality attributes.
Critical quality attributes (CQAs): key physiochemical or biological attributes that can be detected to ensure final product within acceptable quality limits.
Euler-Lagrange CFD method: CFD approach where dispersed entities (particles, droplets, and bubbles) are represented by virtual particles of which positions, velocities, and so on, are tracked, whereas the continuous phase is represented by a field. This in contrast to Euler–Euler methods, which keep track of the concentration of dispersed entities using a concentration and velocity fields. In the current context, biomass is the Lagrangian phase.
Key performance parameters: specify key biosystem capabilities that must be achieved to meet its operational goals in a biological process.
Macromixing: large-scale mixing, caused by mechanical stirring and/or gassing.
Mesomixing: turbulent dispersion of a feed stream shortly after it enters a reactor.
Metabolite lumping/pooling: metabolites inside the cell can be grouped as a series of simplified metabolite pools (e.g., amino acid pool and storage pool) based on both metabolite properties and metabolite turnover timescales.
Metabolite turnover time: can be approximated by dividing intracellular metabolite level Xi (μmol/gDW) by metabolic fluxes Vi (μmol/gDW/s).
Micromixing: mixing/transport of material at the smallest scales, for example in Kolmogorov eddies, but also transport of substrate from the liquid to the biomass.

Box 1. Cross-Scale Insights in Guanosine Fermentation

Guanosine, an essential ingredient of the bisodium ribonucleotide 'I+G', is an expensive food additive [16] and is part of the growth-related fermentation products. As shown in Figure I, in a typical fermentation process, guanosine does not accumulate in the fermentation broth during the growth stage (Phase I, 0–12 h). After ~12 h, the fermentation process enters the late logarithmic phase, at which time it begins to enter the high-speed guanosine production phase with the large accumulation of guanosine (Phase II, 12–40 h). During the late stage, a significant decrease in the rate of guanosine production is unavoidable, which is also the main reason why the guanosine yield cannot be further improved (Phase III, 40 h–end). From the parameter profile during guanosine fermentation (Figure I), we can also see that the oxygen uptake rate (OUR) and the carbon evolution rate (CER) start to decrease at ~40 h, which indicates that the oxidative capacity of the bacteria is decreasing, while the respiratory quotient (RQ) remains essentially unchanged. However, such an abrupt decrease of guanosine productivity is accompanied by an increase in both sugar consumption and ammonia feed rate. The carbon balance suggests that some unknown organic compounds (e.g., amino acids or organic acids) accumulate after 40-h fermentation. Thus, it can be hypothesized that the data correlation analysis on the reactor scale reveals a metabolic shift on the cellular scale.

To substantiate this cross-scale insight, the first step is to ascertain and identify these unknown metabolic intermediates in the fermentation broth. The results of high-performance liquid chromatography (HPLC) further confirmed that alanine and pyruvate accumulate after 40-h fermentation, where the final alanine concentration can be 12.6 times higher than the initial concentration. Consistent with this, the activities of alanine dehydrogenase and phosphofructokinase increased while the activity of glucose-6-phosphate dehydrogenase decreased with the culture age. Furthermore, by setting up a simple stochiometric model containing the central metabolism and the product biosynthesis, the flux ratio [hexose monophosphate pathway (HMP)/Embden–Meyerhof–Parnas pathway (EMP)] shows a sharp decrease over the fermentation process, which also directly points to the overflow metabolism toward the EMP and, thus, alanine production.

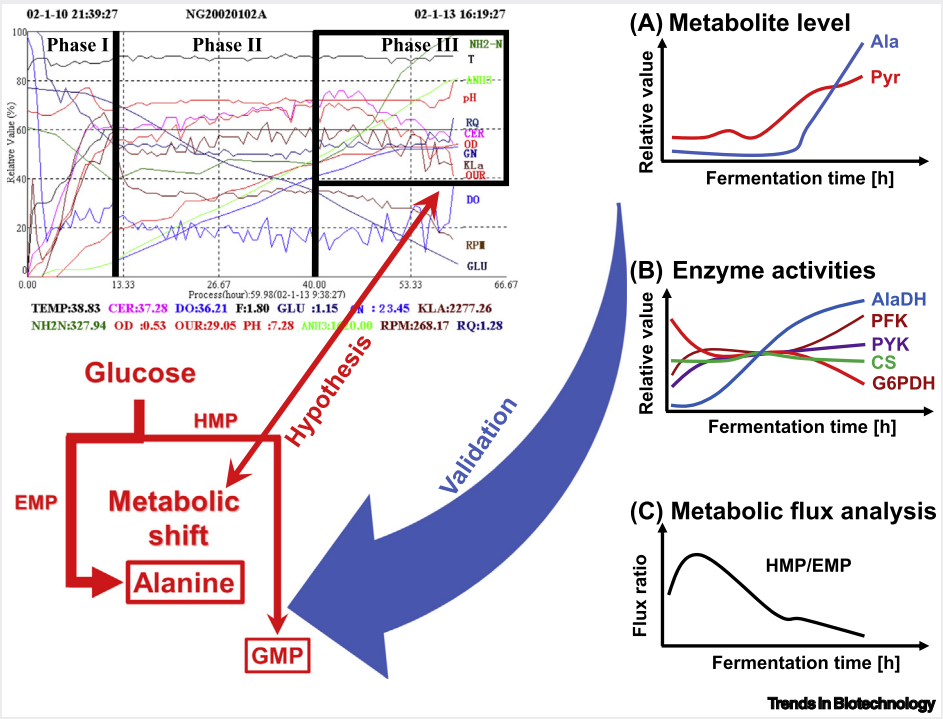

Figure I. Cross-Scale Insights in Guanosine Fermentation. The process macroscopic data sets suggested that there is an overflow metabolism towards alanine accumulation after 40 h guanosine fermentation. This hypothesis was then substantiated by measurements of metabolite, enzyme activities, and metabolic flux. Abbreviations: K_La, oxygen transfer coefficient; Ala, alanine; AlaDH, alanine dehydrogenase; ANH_3, accumulated consumption of ammonia; CER, carbon evolution rate; CS, citrate synthetase; DO, dissolved oxygen; EMP, Embden–Meyerhof–Parnas pathway; F, air flow rate; G6PDH, glucose-6-phosphate dehydrogenase; GLU, glucose; GMP, guanosine monophosphate; TEMP/T, temperature; GN, guanosine; HMP, hexose monophosphate pathway; NH_2-N, amino nitrogen; OD, optical density; OUR, oxygen uptake rate; PFK, phosphofructokinase; PYK, pyruvate kinase; Pyr, pyruvate; RPM, revolutions per minute; RQ, respiratory quotient.

Population balance approach: defines how populations of separate entities develop in specific properties over time.

Process analytical technology (PAT): important process initiative that allows the design, analysis, and control of biopharmaceutical manufacturing processes by measuring CPPs that affects CQA.

Quality by design (QbD): based on PAT and quality risk management, QbD approaches allow biopharmaceutical development to begin with predefined objectives and emphasize product and process understanding as well as process control.

Box 2. CFD-Assisted Process Optimization and Reactor Redesign

CFD has been used to evaluate detailed gas–liquid flow field, such as flow pattern, gas hold-up, mixing time, volumetric power input, and limiting-substrate and oxygen concentration profiles, which is important for prior assessment of mass transfer (k_La), shear stress, and mixing intensity (95% homogeneity) upon bioreactor scale-up. These critical up-scaling factors, such as mixing [69–71], shear environment [57,72,73], and mass transfer [58,74,75], have been dealt with in different bioreactors with different microorganisms or mammalian cells using CFD simulation approaches.

As an example, in a pilot-scale (12 m^3) fermentation of cephalosporin C (CPC) production by an industrial strain *Acremonium chrysogenum*, soybean oil was used as the main carbon source for CPC production; however, the process respiratory quotient (RQ), which is an ideal indicator of oxidation of the substrate, soybean oil, was 0.9–1.0 instead of 0.7. This indicated that utilization of soybean oil was limited because of its low-density property and nonhomogeneous broth when a conventional impeller configuration (four-layered Rushton turbines) was used. To address this, the CFD simulation result suggested that a novel impeller design (two upper axial flow impeller and two lower radial flow impellers) would greatly stimulate soybean oil utilization for CPC production because of more homogeneous mixing of this carbon source. In addition, the up-scaling fermentation results showed that, by using the novel impeller configuration, the RQ value was closer to the theoretical value of 0.7, suggesting a rapid utilization of soybean oil. Also, a 10% increase in CPC production was successfully achieved while the power consumption was decreased by 25% [76].

formulated with reduced structure but yet preserved with enough dynamic features [26]. For example, approximative kinetic formats (e.g., linear-logarithmic) have been recommended in metabolic network modeling for this purpose [28]. In addition, to get more information about metabolic regulation inside the cells, metabolite dynamics in response to extracellular stimuli should be considered because metabolome data provide dynamic and immediate recordings of the cellular response at timescales of seconds to tens of seconds, which is within the order of magnitude of the global liquid circulation time in large-scale bioreactors [29]. Regarding a specific product, most relevant are metabolite dynamics through the central metabolic network and metabolites associated with product biosynthetic pathways. To bypass setting up a kinetic model for each reaction, metabolites can be lumped into different pools (i.e., **metabolite lumping/pooling**) depending on their **metabolite turnover time** and properties. For instance, based on time hierarchy and metabolite properties, a recent study reported a nine-pool metabolic structured model describing both biomass growth and penicillin production [30]. This established model can provide predictions about metabolic pool dynamics across timescales from seconds to days (Figure 2). To achieve this, metabolic data for model parameterization should be representative of the actual metabolic response at a large scale. Otherwise, kinetic information or regulatory mechanisms can be missed.

CFD-Guided Representative Scale-Down Studies

A predictive scale-down model should represent the actual conditions experienced by the cells in the large-scale process. However, most previously published scale-down studies do not take all industrial-scale flow and reaction conditions into account, but rather only a limited set, and some are more worst-case scenarios [31]. Despite their limitations, these studies have provided valuable insights into cellular metabolic responses and also give suggestions for the improvement of the large-scale performance.

In industrial-scale bioreactors, single cell trajectories or 'lifelines' generated using computation describe individual metabolic dynamics and productivity (Figure 1). Population heterogeneity, which frequently arises from different cell lifelines forced by fluid dynamics, has been acknowledged as a key factor affecting overall production performance, but its impact on an industrial scale is seldom accounted for [32]. Therefore, it is valuable to mimic these cell lifelines closely in lab-scale scale-down simulators. CFD simulation is now available to provide details about the flow conditions that are required for representative scale-down designs. In these down-scaling experiments, the effect of industrial scale-sensitive factors (e.g., concentration gradients of substrate, shear stress, dissolved oxygen, and protons) on the cellular metabolic response can be

Key Figure

Overview of the Coupled Model Combining both Cellular Kinetics and Fluid Dynamics

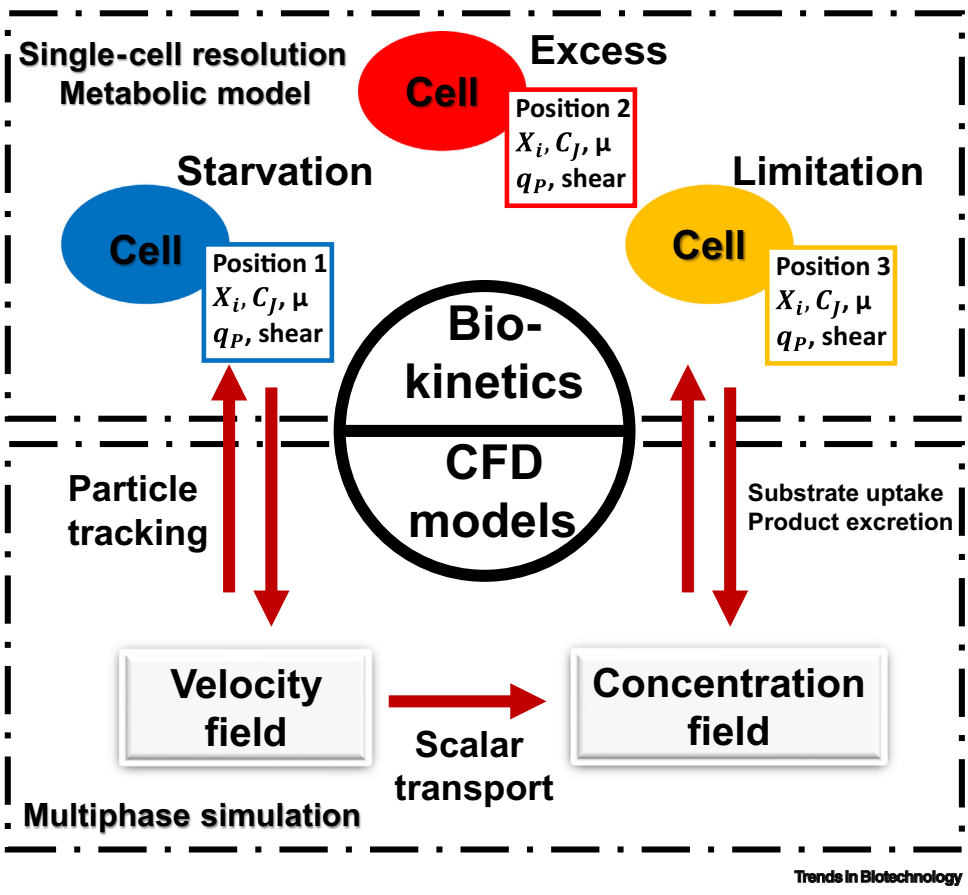

Figure 1. Cells with different colors represent different metabolic regimes (e.g., substrate excess, limitation, and starvation). These cells are tracked in different positions of the large-scale fermenter and each position features profiles of intracellular (X_i, i= glycolytic intermediates, amino acids, or storage carbohydrates) and extracellular (C_j, j=glucose, dissolved oxygen, or precursor) (lumped) metabolite pools as well as specific biomass growth rate (μ) and productivity (q_P). Abbreviation: CFD, computational fluid dynamics.

evaluated on a small scale [4]. For instance, the volumetric power inputs representative of both the impeller region and the bulk zone of a 54-m^3 industrial-scale penicillin fermentation were estimated by CFD simulations, which were then scaled-down to lab-scale chemostat cultures of a high-yielding *Penicillium chrysogenum* strain [33]. The results showed that the reduced penicillin production performance may be caused by glucosensing-induced metabolic rearrangement; an extremely low and slowly decreasing glucose concentration at higher power input might trigger ATP-consuming futile cycling through forming a cycling flux between the pentose phosphate (PP) pathway and the reversed Embden–Meyerhof–Parnas (EMP) pathway. Consequently, less ATP can be channeled towards penicillin formation [33]. However, in this lab-scale setting, the calculated glucose consumption time ($\tau_{r,\ s}$) is smaller than the broth circulation time (t_c), which

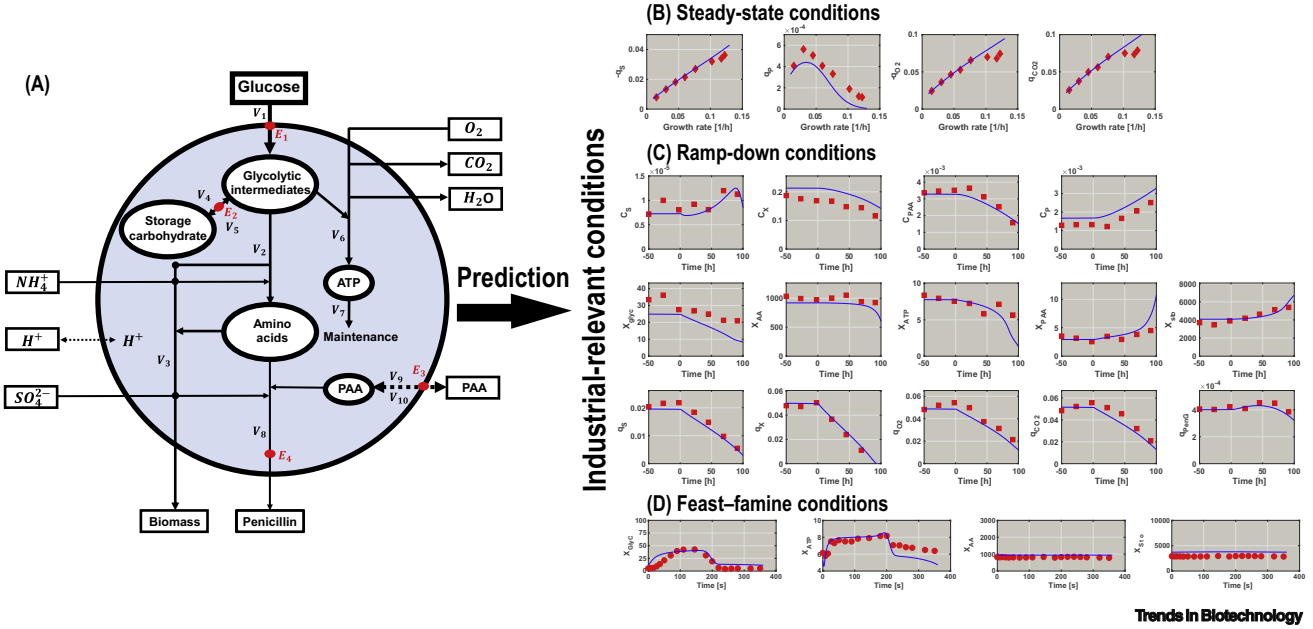

Figure 2. Overview of the Nine-Pool Model for *Penicillium chrysogenum*. (A) Lumped metabolic pools, such as glycolytic intermediates, storage carbohydrates, and amino acids, are defined by metabolite properties and turnover timescales. The final model contains five lumped intracellular metabolite pools (glycolytic intermediates, amino acids, ATP, PAA, and storage carbohydrates), four enzyme (capacity) pools (glucose uptake, PAA export, penicillin conversion, and storage conversion), and ten extracellular components, connected via ten intracellular reactions. Comparison of model predictions and experimental data: (B) as a function of the specific growth rate under chemostat conditions; (C) as a function of time during steady state (–50 to 0 h) and ramp phases (0 to 100 h); (D) under a complete feast–famine cycle of 360 s using block-wise feeding (36 s on, 324 s off). Extracellular C_i (mol/kg), intracellular X_i (mmol/gDW) and specific rates q_i (mol/CmolX/h). Experimental data (red symbols) and simulated result (blue lines) predicted by the nine-pool model. Reproduced, with permission, from [30]. Abbreviation: PAA, phenylacetic acid.

likely gives rise to the gradients of glucose concentration [33]. Therefore, the well-accepted assumption of ideal mixing in lab-scale bioreactors may not be true in certain cases because of a high strain-specific affinity for the substrate. In addition, radial flow patterns governed by radially pumping impellers may aggravate this nonideal mixing phenomenon at the lab scale [34]. In industrial practice, substrate gradients experienced by cells often lead to productivity drop and overflow metabolism [35]. Likewise, based on the CFD simulation results of the 54-m^3 fermenter [36], scale-down simulators in the form of both intermittent feeding with different cycle times (30 s, 3 min, and 6 min) imposed on a single bioreactor as well as a two-compartment system with a mean broth circulation time of 6 min were applied to mimic three representative metabolic regimes (i.e., glucose excess, limitation, and starvation) [12]. Penicillin productivity was reduced in all scale-down simulators and the results from the larger fermenter could be reproduced more accurately by using longer cycle times (3 min and 6 min) rather than short cycle times (30 s) within the same order of magnitude as the global circulation time (20–40 s) [12]. This suggests that the circulation time distribution and, thus, more extensive population effects have major roles in the final performance. However, the cells in these scale-down simulators did not undergo all three representative metabolic regimes identified for the industrial-scale bioreactor [12]. More elaborate downscaling design can be performed via a multivessel scale-down simulator using the five degrees of freedom proposed by Noorman [31]: (i) number of compartments (i.e., number of metabolic regimes); (ii) compartment volume distribution (i.e., volume fraction of each metabolic regime); (iii) recycle rate among the volumes; (iv) flow pattern in each of the volumes; and (v) smooth or 'noisy' feed supply. Furthermore, model-based parallel scaling-down designs in mini-bioreactors or droplet microfluidics can be implemented to accelerate phenotyping before bioprocess scale-up [37,38].

At the same time, in these extreme down-scaling studies, perturbed metabolome data can be collected to capture the cell dynamics and provide 'training' metabolite data for establishing metabolic models (e.g., the nine-pool metabolic structured model) [30]. After incorporating metabolite dynamics into the model, the kinetic properties and metabolic regulation relevant to actual industrial-scale conditions can be predicted (Figure 2).

Assessing Metabolic Response in Industrial Bioreactors via CFD Coupling

From experimental scale-down simulations [4,39,40], it has become evident that the heterogeneous environment in a large-scale reactor potentially impacts the metabolic response of the cell, and thereby exerts an impact on process performance. While CFD simulations with unstructured cellular kinetics [41–45] are valuable to gain insights into environmental gradients and scale-down strategies, these simulations cannot capture the response of the organism to such conditions, because these inherently assume the instantaneous adaptation of the cell to surrounding conditions. Structured cellular models can be combined with CFD simulations to asses this impact, with two philosophies being currently used: **population balance approaches** [46,47], and the Euler–Lagrange (or agent-based) methodology [23,48]. Table 1 lists applications of CFD–cellular reaction dynamics (CRD) models in the literature. Both approaches have their merits and drawbacks [49]. The most advantageous aspect of the population balance approach is scalability to large domains, but as a drawback, heterogeneity in multiple intracellular pools is challenging to include. Morchain and coworkers showed that the population balance approach can be used to study growth rate heterogeneity in a large bioreactor [47]. Pigou and coworkers [50] modeled subsequently introduced structured kinetics, modeling a 22-m^3 cultivation of *E. coli*, with the kinetic model successfully capturing production of acetate as a by-product. In this study, heterogeneity was still described by using the growth rate as a single state variable for the population. To reduce computational demand, a reactor **compartment model** rather than full CFD simulation was used. Advances in the use of moment methods should make population

Table 1. Overview of Research Describing Spatially Resolved Bioreactor Assessment from the Cellular Perspective and/or Including Structured Kinetic Models

Case	Aim	Approach (software)	Reaction model	Refs
Saccharomyces cerevisiae, 0.07/0.9 m^3, stirred tank reactor	NAD^+/NADH cycle synchronization	Agent-based (in-house)	Structured, full coupling	[23]
Escherichia coli, 0.9 m^3, stirred tank reactor	PEP/PYR uptake inhibition	Agent-based (in-house)	Structured, full coupling	[24]
Generic, 0.07/70 m^3, stirred tank reactor	Growth rate distribution	Population balance (ANSYS FLUENT)	Unstructured + growth rate distribution.	[47]
E. coli, 22 m^3 stirred tank reactor	Glucose gradient + acetate production	Population balance (in-house)	Structured + growth rate distribution.	[50]
Carthamus tinctorius L., 5–15 L, stirred tank reactor	Strain exposure quantification	Agent based (ANSYS FLUENT)	Unstructured death kinetics	[57]
S. cerevisiae, 0.24 m^3, bubble column	Glucose gradient	Agent based (ANSYS CFX)	Unstructured, uptake only	[77]
Penicillium chrysogenum, 54 m^3, stirred tank reactor	Glucose gradient	Agent based (ANSYS FLUENT)	Unstructured, uptake only	[36]
S. cerevisiae, 22 m^3, stirred tank reactor	Glucose gradient	Agent based (ANSYS FLUENT)	Unstructured, uptake only	[56]
P. chrysogenum, 54 m^3, stirred tank reactor	Glucose gradient + penicillin production	Agent based (ANSYS FLUENT)	Structured, simplified coupling	[48]
Pseudomonas putida, 54 m^3, stirred tank reactor	Glucose gradient + replication regime	Agent based (ANSYS FLUENT)	Unstructured, Pirt model	[55]
Clostridium ljungdahlii DSM 13528, 125 m^3, bubble column	CO gradient	Agent based (ANSYS FLUENT)	Unstructured, uptake + yield.	[58]

balances more computationally accessible [51,52], while 2D population balances enabled the study of population heterogeneity beyond only the growth rate [53].

The agent-based approach more easily accounts for the use of multiple intracellular pools to capture multidimensional heterogeneity. This has been done by coupling structured kinetic models, describing metabolic, transcriptional, and/or enzymatic responses to environmental conditions, to the individual agents. Currently, to avoid excessive computational demands, most agent-based bioreactor studies still apply unstructured kinetics for substrate uptake, evaluating the resulting lifelines for downscaling purposes as described earlier [36,54–56]. In some cases, the response of the organism was evaluated post processing. Haringa and colleagues observed using a structured kinetic model that extracellular glucose gradients led to a substantial reduction in penicillin production [30]. Liu and colleagues showed that shear-rate lifelines with a simple cell death kinetic model could describe cell death in cultivations of *Carthamus tinctorius L.* cells [57]. More recently, Siebler and colleagues studied the probability of *Clostridium ljungdahlii* cells undergoing transcriptional changes based on their residence time under stress conditions in a syngas fermentation [58].

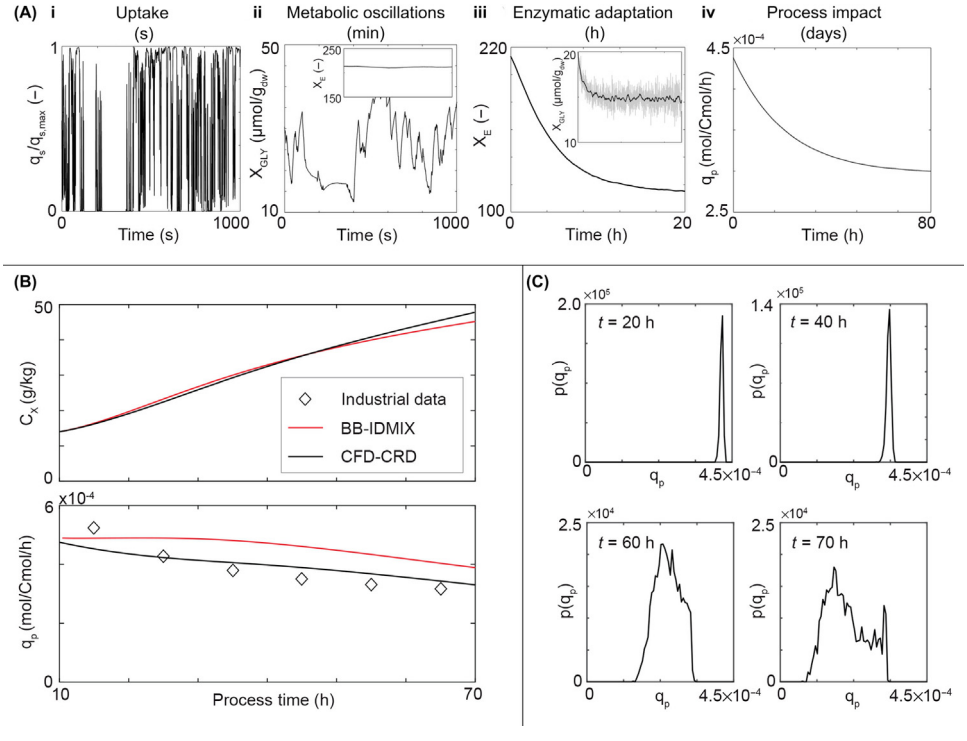

Figure 3. Results of a Coupled Computational Fluid Dynamics (CFD)–Cellular Reaction Dynamics (CRD) Simulation for a Penicillin Production Process. (A) Illustration of process response at different timescales. Extracellular fluctuations in substrate concentration cause fluctuations in substrate uptake rate (i) on second timescale, leading to metabolic variations on the minute timescale (ii); enzyme levels are constant at this timescale (ii, inset). On an hour timescale, metabolic fluctuations cause enzyme-level adaptation (iii); these in turn can affect metabolic fluctuations (inset; rolling average in black, fluctuations in gray). Finally, enzyme adaptation changes the penicillin production rate over the full process time (iv). (B) Performance of CFD–CRD in an industrial fed batch process compared with a black-box, ideal-mixed model (BB-IDMIX). (C) Due to differences observed early during the process, individual metabolic responses lead to divergence in the population as a function of time; after 20 h, all organisms have nearly the same production rate, while after 60 h a wider distribution is observed (sampled from 2500 tracked organisms). Adapted with permission from [48] (B).

Haringa and colleagues included the impact of intracellular variations on glucose uptake, although some simplifications of CFD–CRD coupling were possible because intracellular feedback occurred at the comparatively slow enzymatic level [48]. The resulting model could predict the decrease in penicillin production much better than an ideal-mixing unstructured model, which overpredicted production, or a CFD-coupled unstructured model, which strongly underpredicted it (Figure 3). In addition, spontaneous emergence of population heterogeneity in the glucose uptake capacity was observed, although this observation has not yet been tested experimentally [48]. Lapin and colleagues used an approach including rapid metabolic feedback. First, they modeled synchronization–desynchronization in intracellular NAD^+/NADH ratios in *Saccharomyces cerevisiae* in response to extracellular glucose fluctuations [23]. In follow-up work using *E. coli*, they captured the inhibitory impact of excreted pyruvate and phosphoenolpyruvate on glucose uptake; an effect where not only intracellular, but also extracellular concentration dynamics were determined by the cellular response [24]. The major challenge of the agent-based model is to simulate sufficient particles to avoid artificial gradients in biomass concentration [59], and the required number of particles will increase if local heterogeneity in the uptake rate needs to be included [49]. These limitations may be alleviated as graphical processing units (GPU)-driven agent-based models make their way to bioprocess engineering[i].

Concluding Remarks and Future Prospects

Computational frameworks via metabolic–hydrodynamic coupling can provide informative insights into cell lifelines (subpopulation distribution and evolution) and metabolic responses (metabolite dynamics, growth, and productivity) triggered by environmental fluctuations in an industrial bioreactor. In addition, this coupled framework can, in principle, serve to comprehend, predict, and evaluate the effects of adding, removing, or modifying molecular components and/or pathways (e.g., substrate uptake kinetics, production excretion capacity, futile cycles, carbohydrate storage/mobilization, and induction of key enzymes) of a cell factory for more robust performance, while making suggestions for the design of the bioreactor and fermentation process.

The set-up of highly predictive metabolic models is the first step to allow follow-up coupling to CFD models. However, there are significant challenges regarding the acquisition of the *in vivo* highly dynamic and representative metabolite data sets needed to establish a metabolic structured kinetic model that can describe the day-scale to second-scale dynamics of cell growth and product formation (see Outstanding Questions). Although the concept of lumping metabolite pools based on a time hierarchy and their properties reduces the complexity of model parameterization while preserving basic dynamic behavior, other key metabolic information, such as redox potential (NADH/NAD^+) and energy charge {[ATP+(ADP/2)]/(ATP+ADP+AMP)}, should be incorporated in the future. Also, parameters in the kinetic model should be repeatedly estimated by obtaining more industrial scale-relevant conditions. For example, despite the high predictive ability of the previously established nine-pool model, metabolic data sets were obtained mainly from glucose-perturbed conditions, not considering the oscillating oxygen scenario as well as the influence of combinations thereof. Therefore, there is still room for model extension and upgrading. In addition, to push bioprocessing towards digitalization and automation, process models should in the future become self-adapting using real-time data sets and making adaptive decisions for process control and development [60]. In the meantime, **critical process parameters** (CPPs), **critical quality attributes** (CQAs), and **key performance parameters** (KPPs) should be established to improve process quality and reduce lot-to-lot variability and even batch failures. However, there is difficulty in identifying the most appropriate CPPs, CQAs, and KPPs during the product lifecycle [61]. Further developments in biotechnology will largely depend on the extensive use of big data. To address this, big data analytics, artificial intelligence (AI), and machine learning will be critical to link CPPs to CQAs and identify KPPs in intelligent biomanufacturing [62].

Outstanding Questions

From an industrial perspective, what are the main influential factors in a fluid flow field inside an industrial fermenter?

Through the 'eyes' of the organism, what are the cellular responses to environmental perturbations (e.g., alternating substrate feast–famine conditions) in an industrial fermenter?

Towards self-regulated homeostasis, what mechanisms do organisms use to regulate their metabolic response in both steady state and under highly dynamic conditions?

How can we switch industrial production processes from complex to defined media, such that science-based approaches can be used to reduce lot-to-lot variability and ensure process stability and performance?

How can we generate cell dynamics in lab-scale experiments that are representative of large scales, such that a prior assessment of process performance under mimicked industrial conditions can be achieved?

How can we obtain true snapshots of both the average metabolome and the fluctuations around the mean such that a quantitative single-cell kinetic model can be established and the phenomenon of population heterogeneity can be evaluated?

In the field of bioreactor modeling, many challenges remain beyond the obvious aspect of computation time. Most CFD models are based on (air-)water experiments, and the impact of rheology and the presence of surfactants in fermentation broths are insufficiently accounted for; an effort is needed both in experimental hydrodynamics and simulation. In addition to **macromixing** because of impeller and gas–liquid interactions, **mesomixing** (initial feed mixing) and **micromixing** (e.g., boundary layer and intrapellet mass transfer) effects could be relevant in some situations [63,64]. Regarding the computational aspect, many exiting techniques are under development. Using GPUs may allow for greatly increased particle numbers in Lagrangian simulations. In addition, there are innovative developments of faster CFD solvers and options for high-resolution flow reconstruction, such as rCFD [65] and AI-based approaches, such as ANSYS reduced-order modeling. Combining these developments will help to evaluate complete bioprocesses rather than snapshots, or to screen wider ranges of possible cell designs and process configurations, with a higher level of realism. In the meantime, advances in analysis methods and microfluidic cultivations may unlock a new generation of scale-down simulations, where both the population and single-cell response can be studied in higher resolution [38,49,66–68].

Acknowledgments

The authors are grateful to Wouter van Winden, Wenjun Tang, and Jianye Xia for critical comments, and Siliang Zhang for providing data sets regarding guanosine fermentations. This work was financially supported by the National Natural Science Foundation of China (Grant no. 31900073), the Science and Technology Commission of Shanghai Municipality (Grant no. 19ZR1413600), the National Key Research and Development Program (Grant no. 2017ZX7402003), and the 111 Project (Grant no. B18022).

Resources

[i]www.flamegpu.com/

References

1. Aguilar, A. *et al.* (2019) Bioeconomy for sustainable development. *Biotechnol. J.* 14, e1800638
2. Lokko, Y. *et al.* (2018) Biotechnology and the bioeconomy-towards inclusive and sustainable industrial development. *Nat. Biotechnol.* 40, 5–10
3. Festel, G. (2010) Industrial biotechnology: market size, company types, business models, and growth strategies. *Ind. Biotechnol.* 6, 88–94
4. Lara, A.R. *et al.* (2006) Living with heterogeneities in bioreactors: understanding the effects of environmental gradients on cells. *Mol. Biotechnol.* 34, 355–381
5. Heins, A.L. and Weuster-Botz, D. (2018) Population heterogeneity in microbial bioprocesses: origin, analysis, mechanisms, and future perspectives. *Bioprocess Biosyst. Eng.* 41, 889–916
6. Takhaveev, V. and Heinemann, M. (2018) Metabolic heterogeneity in clonal microbial populations. *Curr. Opin. Microbiol.* 45, 30–38
7. Delvigne, F. *et al.* (2014) Metabolic variability in bioprocessing: implications of microbial phenotypic heterogeneity. *Trends Biotechnol.* 32, 608–616
8. Delvigne, F. and Goffin, P. (2014) Microbial heterogeneity affects bioprocess robustness: dynamic single-cell analysis contributes to understanding of microbial populations. *Biotechnol. J.* 9, 61–72
9. Heins, A.L. *et al.* (2015) Experimental and *in silico* investigation of population heterogeneity in continuous *Saccharomyces cerevisiae* scale-down fermentation in a two-compartment setup. *J. Chem. Technol. Biotechnol.* 90, 324–340
10. Heins, A.L. *et al.* (2019) The effect of acetate on population heterogeneity in different cellular characteristics of *Escherichia coli* in aerobic batch cultures. *Biotechnol. Prog.* 35, e2796
11. Heins, A.L. *et al.* (2019) Quantitative cytomics to understand population heterogeneity in response to changes in substrate availability in *Escherichia coli* and *Saccharomyces cerevisiae* chemostats. *Front. Bioeng. Biotechnol.* 7, 187
12. Wang, G. *et al.* (2018) Comparative performance of different scale-down simulators of substrate gradients in *Penicillium chrysogenum* cultures: the need of a biological systems response analysis. *Microb. Biotechnol.* 11, 486–497
13. Smith, A. *et al.* (2018) The culture environment influences both gene regulation and phenotypic heterogeneity in *Escherichia coli*. *Front. Microbiol.* 9, 1739
14. Takors, R. (2012) Scale-up of microbial processes: impacts, tools and open questions. *J. Biotechnol.* 160, 3–9
15. Glassey, J. *et al.* (2011) Process analytical technology (PAT) for biopharmaceuticals. *Biotechnol. J.* 6, 369–377
16. Zhang, S. *et al.* (2004) A multi-scale study of industrial fermentation processes and their optimization. *Adv. Biochem. Eng. Biotechnol.* 87, 97–150
17. Zhang, S.L. *et al.* (2006) From multi-scale methodology to systems biology: to integrate strain improvement and fermentation optimization. *J. Chem. Technol. Biotechnol.* 81, 734–745
18. Wang, Y. *et al.* (2009) Industrial bioprocess control and optimization in the context of systems biotechnology. *Biotechnol. Adv.* 27, 989–995
19. Ali, H. *et al.* (2019) CFD and kinetic-based modeling to optimize the Sparger design of a large-scale photobioreactor for scaling up of biofuel production. *Biotechnol. Bioeng.* 116, 2200–2211
20. Villiger, T.K. *et al.* (2018) Experimental and CFD physical characterization of animal cell bioreactors: from micro- to production scale. *Biochem. Eng. J.* 131, 84–94
21. Pan, A. *et al.* (2017) CFD simulation of average and local gas–liquid flow properties in stirred tank reactors with multiple rushton impellers. *J. Chem. Eng. Japan* 50, 878–891
22. Delafosse, A. *et al.* (2014) CFD-based compartment model for description of mixing in bioreactors. *Chem. Eng. Sci.* 106, 76–85
23. Lapin, A. *et al.* (2004) Dynamic behavior of microbial populations in stirred bioreactors simulated with Euler-Lagrange methods: traveling along the lifelines of single cells. *Ind. Eng. Chem. Res.* 43, 4647–4656
24. Lapin, A. *et al.* (2006) Modeling the dynamics of *E. coli* populations in the three-dimensional turbulent field of a stirred-tank bioreactor—a structured–segregated approach. *Chem. Eng. Sci.* 61, 4783–4797
25. Douma, R.D. *et al.* (2010) Dynamic gene expression regulation model for growth and penicillin production in *Penicillium chrysogenum*. *Biotechnol. Bioeng.* 106, 608–618

26. Dano, S. *et al.* (2006) Reduction of a biochemical model with preservation of its basic dynamic properties. *FEBS J.* 273, 4862–4877
27. Narayanan, H. *et al.* (2019) Bioprocessing in the Digital Age: the role of process models. *Biotechnol. J.* 15, e1900172
28. Heijnen, J.J. (2005) Approximative kinetic formats used in metabolic network modeling. *Biotechnol. Bioeng.* 91, 534–545
29. Wang, G. *et al.* (2015) Integration of microbial kinetics and fluid dynamics toward model-driven scale-up of industrial bioprocesses. *Eng. Life Sci.* 15, 20–29
30. Tang, W. *et al.* (2017) A 9-pool metabolic structured kinetic model describing days to seconds dynamics of growth and product formation by *Penicillium chrysogenum*. *Biotechnol. Bioeng.* 114, 1733–1743
31. Noorman, H. (2011) An industrial perspective on bioreactor scale-down: what we can learn from combined large-scale bioprocess and model fluid studies. *Biotechnol. J.* 6, 934–943
32. Fernandes, R.L. *et al.* (2011) Experimental methods and modeling techniques for description of cell population heterogeneity. *Biotechnol. Adv.* 29, 575–599
33. Wang, G. *et al.* (2018) Power input effects on degeneration in prolonged penicillin chemostat cultures: a systems analysis at flux, residual glucose, metabolite, and transcript levels. *Biotechnol. Bioeng.* 115, 114–125
34. Sinclair, C.G. and Brown, D.E. (1970) Effect of incomplete mixing on the analysis of the static behaviour of continuous cultures. *Biotechnol. Bioeng.* 12, 1001–1017
35. Enfors, S.O. *et al.* (2001) Physiological responses to mixing in large scale bioreactors. *J. Biotechnol.* 85, 175–185
36. Haringa, C. *et al.* (2016) Euler-Lagrange computational fluid dynamics for (bio)reactor scale down: an analysis of organism lifelines. *Eng. Life Sci.* 16, 652–663
37. Anane, E. *et al.* (2019) A model-based framework for parallel scale-down fed-batch cultivations in mini-bioreactors for accelerated phenotyping. *Biotechnol. Bioeng.* 116, 2906–2918
38. Ho, P. *et al.* (2019) Reproduction of large-scale bioreactor conditions on microfluidic chips. *Microorganisms* 7, 105
39. Neubauer, P. and Junne, S. (2010) Scale-down simulators for metabolic analysis of large-scale bioprocesses. *Curr. Opin. Biotechnol.* 21, 114–121
40. de Jonge, L.P. *et al.* (2011) Scale-down of penicillin production in *Penicillium chrysogenum*. *Biotechnol. J.* 6, 944–958
41. Larsson, G. *et al.* (1996) Substrate gradients in bioreactors: origin and consequences. *Bioprocess Eng.* 14, 281–289
42. Spann, R. *et al.* (2019) CFD predicted pH gradients in lactic acid bacteria cultivations. *Biotechnol. Bioeng.* 116, 769–780
43. Laakkonen, M. *et al.* (2006) Dynamic modeling of local reaction conditions in an agitated aerobic fermenter. *AIChE J.* 52, 1673–1689
44. Elqotbi, M. *et al.* (2013) CFD modelling of two-phase stirred bioreaction systems by segregated solution of the Euler-Euler model. *Comput. Chem. Eng.* 48, 113–120
45. Park, S. and Li, Y. (2015) Integration of biological kinetics and computational fluid dynamics to model the growth of *Nannochloropsis salina* in an open channel raceway. *Biotechnol. Bioeng.* 112, 923–933
46. Morchain, J. *et al.* (2013) Coupling of biokinetic and population balance models to account for biological heterogeneity in bioreactors. *AIChE J.* 59, 369–379
47. Morchain, J. *et al.* (2014) A coupled population balance model and CFD approach for the simulation of mixing issues in lab-scale and industrial bioreactors. *AIChE J.* 60, 27–40
48. Haringa, C. *et al.* (2018) Computational fluid dynamics simulation of an industrial *P. chrysogenum* fermentation with a coupled 9-pool metabolic model: towards rational scale-down and design optimization. *Chem. Eng. Sci.* 175, 12–24
49. Haringa, C. *et al.* (2018) From industrial fermentor to CFD-guided downscaling: what have we learned? *Biochem. Eng. J.* 140, 57–71
50. Pigou, M. and Morchain, J. (2015) Investigating the interactions between physical and biological heterogeneities in bioreactors using compartment, population balance and metabolic models. *Chem. Eng. Sci.* 126, 267–282
51. Pigou, M. *et al.* (2018) New developments of the extended quadrature method of moments to solve population balance equations. *J. Comput. Physics* 365, 243–268
52. Pigou, M. *et al.* (2017) An assessment of methods of moments for the simulation of population dynamics in large-scale bioreactors. *Chem. Eng. Sci.* 171, 218–232
53. Quedeville, V. *et al.* (2018) A two-dimensional population balance model for cell growth including multiple uptake systems. *Chem. Eng. Res. Design* 132, 966–981
54. Noorman, H.J. and Heijnen, J.J. (2017) Biochemical engineering's grand adventure. *Chem. Eng. Sci.* 170, 677–693
55. Kuschel, M. *et al.* (2017) Lagrangian trajectories to predict the formation of population heterogeneity in large-scale bioreactors. *Bioengineering* 4, 27
56. Haringa, C. *et al.* (2017) Euler-Lagrange analysis towards representative down-scaling of a 22 m^3 aerobic *S. cerevisiae* fermentation. *Chem. Eng. Sci.* 170, 653–669
57. Liu, Y. *et al.* (2016) Application of Euler–Lagrange CFD for quantitative evaluating the effect of shear force on *Carthamus tinctorius L.* cell in a stirred tank bioreactor. *Biochem. Eng. J.* 114, 209–217
58. Siebler, F. *et al.* (2019) The impact of CO gradients on *C. ljungdahlii* in a 125 m^3 bubble column: mass transfer, circulation time and lifeline analysis. *Chem. Eng. Sci.* 207, 410–423
59. Haringa, C. *et al.* (2017) Lagrangian modeling of hydrodynamic–kinetic interactions in (bio)chemical reactors: practical implementation and setup guidelines. *Chem. Eng. Sci.* 157, 159–168
60. Bournazou, M.N.C. *et al.* (2017) Online optimal experimental re-design in robotic parallel fed-batch cultivation facilities. *Biotechnol. Bioeng.* 114, 610–619
61. Oliveira, A.L. (2019) Biotechnology, big data and artificial intelligence. *Biotechnol. J.* 14, e1800613
62. Oyetunde, T. *et al.* (2018) Leveraging knowledge engineering and machine learning for microbial bio-manufacturing. *Biotechnol. Adv.* 36, 1308–1315
63. Linkès, M. *et al.* (2014) Numerical investigation of subgrid mixing effects on the calculation of biological reaction rates. *Chem. Eng. Sci.* 116, 473–485
64. Morchain, J. *et al.* (2017) A population balance model for bioreactors combining interdivision time distributions and micromixing concepts. *Biochem. Eng. J.* 126, 135–145
65. Pirker, S. and Lichtenegger, T. (2018) Efficient time-extrapolation of single- and multiphase simulations by transport based recurrence CFD (rCFD). *Chem. Eng. Sci.* 188, 65–83
66. Binder, D. *et al.* (2017) Homogenizing bacterial cell factories: analysis and engineering of phenotypic heterogeneity. *Metab. Eng.* 42, 145–156
67. Lemoine, A. *et al.* (2017) Tools for the determination of population heterogeneity caused by inhomogeneous cultivation conditions. *J. Biotechnol.* 251, 84–93
68. Rosenthal, K. *et al.* (2017) Beyond the bulk: disclosing the life of single microbial cells. *FEMS Microbiol. Rev.* 41, 751–780
69. Haringa, C. *et al.* (2018) Inter-compartment interaction in multi-impeller mixing: Part I. Experiments and multiple reference frame CFD. *Chem. Eng. Res. Des.* 136, 870–885
70. Haringa, C. *et al.* (2018) Inter-compartment interaction in multi-impeller mixing. Part II. Experiments, sliding mesh and large Eddy simulations. *Chem. Eng. Res. Des.* 136, 886–899
71. McClure, D.D. *et al.* (2015) Mixing in bubble column reactors: experimental study and CFD modeling. *Chem. Eng. J.* 264, 291–301
72. Hutmacher, D.W. and Singh, H. (2008) Computational fluid dynamics for improved bioreactor design and 3D culture. *Trends Biotechnol.* 26, 166–172
73. Verma, R. *et al.* (2019) Computational fluid dynamic analysis of hydrodynamic shear stress generated by different impeller combinations in stirred bioreactor. *Biochem. E. J.* 151, 107312
74. Kerdouss, F. *et al.* (2008) Two-phase mass transfer coefficient prediction in stirred vessel with a CFD model. *Comput. Chem. Eng.* 32, 1943–1955
75. McClure, D.D. *et al.* (2018) Oxygen transfer in pilot-scale contactors: an experimental and computational investigation into the effect of contactor design. *Chem. Eng. J.* 344, 173–183
76. Yang, Y. *et al.* (2012) A novel impeller configuration to improve fungal physiology performance and energy conservation for cephalosporin C production. *J. Biotechnol.* 161, 250–256
77. McClure, D.D. *et al.* (2016) Characterizing bubble column bioreactor performance using computational fluid dynamics. *Chem. Eng. Sci.* 144, 58–74