



Mini Review

In Silico Prediction of Large-Scale Microbial Production Performance: Constraints for Getting Proper Data-Driven Models

Julia Zieringer, Ralf Takors *

Institute of Biochemical Engineering, University of Stuttgart, Germany

ARTICLE INFO

Article history:

Received 26 March 2018

Received in revised form 11 June 2018

Accepted 12 June 2018

Available online 6 July 2018

Keywords:

Gene regulatory networks

Scale-down devices

CFD

Compartment models

CFD-based compartment models

ABSTRACT

Industrial bioreactors range from 10.000 to 700.000 L and characteristically show different zones of substrate availabilities, dissolved gas concentrations and pH values reflecting physical, technical and economic constraints of scale-up. Microbial producers are fluctuating inside the bioreactors thereby experiencing frequently changing micro-environmental conditions. The external stimuli induce responses on microbial metabolism and on transcriptional regulation programs. Both may deteriorate the expected microbial production performance in large scale compared to expectations deduced from ideal, well-mixed lab-scale conditions. Accordingly, predictive tools are needed to quantify large-scale impacts considering bioreactor heterogeneities. The review shows that the time is right to combine simulations of microbial kinetics with calculations of large-scale environmental conditions to predict the bioreactor performance. Accordingly, basic experimental procedures and computational tools are presented to derive proper microbial models and hydrodynamic conditions, and to link both for bioreactor modeling. Particular emphasis is laid on the identification of gene regulatory networks as the implementation of such models will surely gain momentum in future studies.

© 2018 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Contents

1. Introduction	246
2. Data-driven Approach	247
2.1. Experimental Set-Ups Mimicking Large-Scale Heterogeneities.	247
2.2. Experimental Access to Metabolic and Transcriptional Responses	247
2.3. Experimental Access to Single Cell Analysis	247
3. Modeling Microbial Growth With Different Granularity	248
3.1. Identifying Structured and Non-Structured Microbial Models	249
3.2. Identifying Gene Regulatory Networks (GRNs)	249
4. Simulating the Cellular Environment With Embedded Growing Cells	250
4.1. Modeling of Hydrodynamics and Mass Transfer	251
4.2. Hydrodynamic Modeling Linked to GRN Models	251
5. Conclusion and Perspectives	253
Conflict of Interest	253
References.	253

1. Introduction

With the advent of metabolic engineering in the 1990s [1], the engineers' view on microbes changed. Process optimization no longer

considered the extracellular environment (i.e. cultivation conditions) alone, but started to investigate intracellular mechanisms in addition [1, 2]. Since then, intracellular reaction rates have been quantified and models of regulatory processes finally aiming at identifying targets for further strain and process improvement have been derived. To some extent driven by the observations that cellular engineering always results in multiple and complex systemic responses [1], furthermore catalyzed by the avalanche of omics data that were accessible, systems biology and systems metabolic engineering emerged in 2000. In essence,

RRH: Preprint Submitted for review. Submission based on invitation of Gianni Panagiotou

* Corresponding author.

E-mail address: takors@ibt.uni-stuttgart.de (R. Takors).

holistic models have been developed that aim to provide as sound and comprehensive a cellular view as possible.

The development clearly reflects the general engineering mindset of investigating the whole system by modularization, quantitative analysis, reassembling and studying the interaction of the networked modules. The earliest, simple examples may be given by the Monod growth model [3], followed by more sophisticated approaches like the lactose operon considering feedback regulation in *Escherichia coli*, finally leading to complex models comprising multiple levels of cellular regulation [4]. While such movements led to the birth of systems biology [5] and systems metabolic engineering [6–9] core engineering activities such as scale-up were a matter of steady development, too.

Scale-up is the procedure to transfer lab-bioprocesses in production (large) conditions, often covering 7 to 8 orders of magnitude of volume. Unfortunately, loss or even failure of large-scale performance may occur. Detailed knowhow is necessary to prevent unwanted production losses. Accordingly, Oosterhuis and Kossen were the first who presented a scale-up simulator (1983) for investigating the impact of oxygen gradients on *Gluconobacter oxydans* [10]. They further introduced bioreactor compartment models to achieve the coarse spatial resolution of local oxygen transfer rates to identify micro- and anaerobic zones [11]. This line of thinking was followed by a series of similar studies [12–16] and reached a new level of complexity by linking simulations of hydrodynamics and mass transports with simple metabolic models of *Saccharomyces cerevisiae* and *E. coli* [17–21]. Notably, cellular dynamics were modeled by focusing on metabolism dynamics only. This is remarkable as systems biology has already shown that holistic models are able to cover a far broader range of complexity. Scale-up engineers have already pointed out [22] that profound knowhow is necessary to enable the best knowledge-based scale-up using in silico predictions.

This review addresses the current need for knowledge-based process scale-up by elucidating the putative contributions of modeling. The existing plethora of modeling approaches will be structured with respect to granularity and usefulness to (i) identify and (ii) model key regulatory phenomena and (iii) to link cellular models with predictions of large-scale hydrodynamics. It will be shown that the time is right to approach the challenging goal of in silico predicted large-scale performance of microbial producers.

2. Data-driven Approach

Comprehensive data sets are necessary to develop gene regulatory models, generated to answer the biological question of interest. This also holds true for elucidating complex metabolic and regulatory responses of producer cells that are exposed to industrial production conditions. One approach to collect representative data is to mimic large-scale conditions and to capture time series of regulatory dynamics as a basis for unraveling dynamic regulatory models. Such approaches usually require rapid sampling experiments that ‘freeze’ metabolic states monitored in scale-down experiments. Examples of experimental procedures are given in the following.

2.1. Experimental Set-Ups Mimicking Large-Scale Heterogeneities

In large-scale production processes micro-environmental inhomogeneities often occur. Insufficient mixing leads to severe axial and horizontal concentration gradients. Producer cells frequently cross these poorly mixed zones which triggers metabolic and transcriptional responses accordingly [23]. Because large-scale experimental data are rarely accessible, experimental scale-up simulators are typically applied, reflecting large-scale conditions [24]. Pioneering studies were performed by Oosterhuis et al. [10] using a two compartment system comprising two stirred tank reactors (STRs) to investigate the effect of different oxygen levels upon the gluconic acid fermentation of *Gluconobacter oxydans*. Since then, variations of the two compartment set up considered the combination of an STR and a plug flow reactor

(PFR). Reviews have been given by Delvigne et al. and Neubauer and Junne [22, 25, 26]. Fig. 1 depicts selected examples for several STR-STR and STR-PFR applications.

Experimental scale-up simulators do not merely consist of two compartments. Three compartment approaches have been studied as well. Examples are the STR-STR-STR cascade of Buchholz et al. [13] and the PFR-STR-PFR set-up of Lemoine et al. [28]. Accordingly, more complex scale-up scenarios could be analyzed.

Notably, two and three compartment scale-up simulators mirror the cellular responses on repeated, frequent stimuli. In contrast, investigations of single perturbations may be a proper tool for deriving distinct stimulus/response correlations, see Fig. 1 for examples. On this basis, explicit metabolic and transcriptional dynamics can be deduced that, when properly superimposed, result in the complex cellular response observed. However, signal transduction is highly networked in the cells which may cause the cross-interference of multiple stimuli. The co-incidence of multiple stimuli in large-scale fermentation is the rule rather than the exception [29, 30]. Accordingly, multiple stimulus/response studies are likely to gain importance in the future.

2.2. Experimental Access to Metabolic and Transcriptional Responses

Samples taken from the scale-up simulators need to be processed so that metabolic and transcriptional states are ‘frozen’ immediately. Metabolic inactivation and purification can be achieved via several approaches [31–34] and requires individual optimization for the given problem. Blocking intracellular transcription is achieved by sampling into RNA protect kits [14]. Correctly prepared, samples can be treated further to identify metabolic compositions via metabolic profiling or fingerprinting techniques [35–37], protein contents via affinity tags [38] or mass spectrometry [39] and transcript levels, either applying microarrays or, more preferred, next generation sequencing technologies analyzing mRNAs [40–42]. To reduce the overall sequencing expenses, library preparation usually is done via a rRNA depletion or poly-A enrichment step to remove non-coding rRNA.

Various methods for RNA Seq analysis are available and have been reviewed recently by Conesa et al. [43]. Regarding modeling, time series of transcripts are particularly important which requires methods of differential gene expression analysis. Fig. 2 provides an overview of a typical workflow making use of public R packages.

Once time series of transcripts are available, modelers may be interested in unraveling gene clusters showing similar transcription dynamics and data integration in dynamic models. Applicants may be guided via evaluating reports of Rapaport et al. [49], Hecker et al. and Banf et al. [50, 51]. Currently, algorithms such as DeSeq2 [52] and MaSigPro [53] are often applied.

Application examples are given by transcript time series and monitoring of metabolic changes reflecting the stimuli of glucose [14, 15, 54], nitrogen [55], oxygen [16, 56] or temperature stress [57] of *E. coli*. Data like this, derived from transcriptome analysis as it is described in Fig. 2, are the basis of proper validated mathematical models. Transcript analysis even enabled the engineering of robust *E. coli* strains [58] by attenuating the level of the alarmone ppGpp, the inducer of the stringent response regulation program. The new host is able to maintain high glucose uptake rates even under non or slow growing conditions.

2.3. Experimental Access to Single Cell Analysis

It is a well known fact that microbial populations in bioreactors are rather heterogeneous than homogeneous. A combination of stimuli such as local substrate availabilities, temperature and pH conditions may induce differences in cell cycle status, cell division, growth rates, etc. [59].

Such subpopulations can be experimentally analyzed via studies, for instance using fully-automated real-time, flow injection flow cytometry

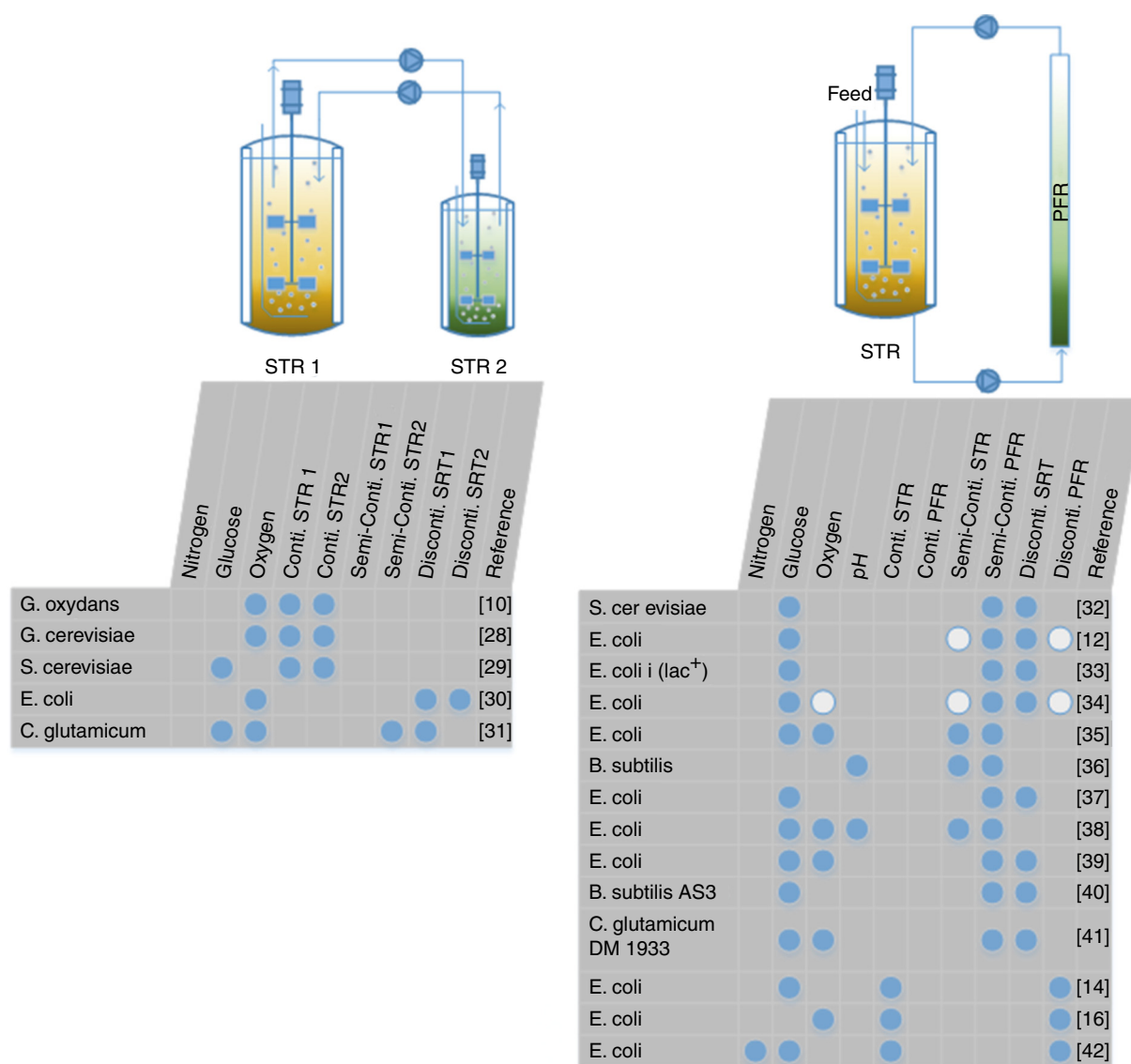


Fig. 1. Matrix of STR-STR and STR-PFR applications with different fluctuating conditions and operation modes (blue dots). The *E. coli* strain is the standard strain W3110. Alternative approaches or different operation modes within the same publication are displayed as blue circles with a white filling. The experimental setups are arranged by the year of publication. Investigations with redundant application information are mentioned once according to the most recent paper.

(FI-FCM) [60,61] or real-time imaging in combination with microfluidic cultivation devices [62–64].

Bennett and Hasty [65] reviewed several microfluidic devices which can be used to examine intracellular signaling pathways and the dynamics of gene regulation in bacteria, yeast and higher eukaryotes on a single cell basis. Often, on-line monitoring is combined with microfluidic studies to achieve full resolution of complex interactions. These technologies are expected to yield novel insights and allow the construction of mathematical models that more accurately describe the complex dynamics of gene regulation [65]. Lemoine et al. provided a review about tools for monitoring population heterogeneities on a single cell basis [66]. Today, even single cell transcription analysis using novel sequencing technologies is becoming achievable [67] which may further increase the quality of mathematical models.

3. Modeling Microbial Growth With Different Granularity

Based on properly analyzed experimental data sets, mathematical models can be derived to simulate the microbial behaviour under

different conditions with a varying level of detail. Following the well-known classification of Bailey [68] microbial models can be divided into non-structured/structured and non-segregated/segregated approaches. Non-structured/Non-segregated approaches represent the simplest growth models assuming average cells without subcellular detail. Such models are typically applied for bioprocess design. For the sake of simplicity, they are also applied in agent-based modeling for tracking individual cells. The consideration of subpopulations or individual cell properties leads to segregated approaches which, thanks to the improving availability of experimental data, is gaining more and more attraction. Structured, non-segregated models are commonly used for implementing the subcellular details of metabolic and transcriptional regulation, compartmentation or signal transduction [69,70]. They are computationally intensive but represent a powerful tool for predicting detailed cellular responses to extracellular stimuli. The most accurate approach are structured/segregated models [71], which for example describe the whole glycolysis process with reactions for each enzyme, depending on enzyme affinities and turn over rates. These parameters are more difficult to identify but transferable to other conditions. However,

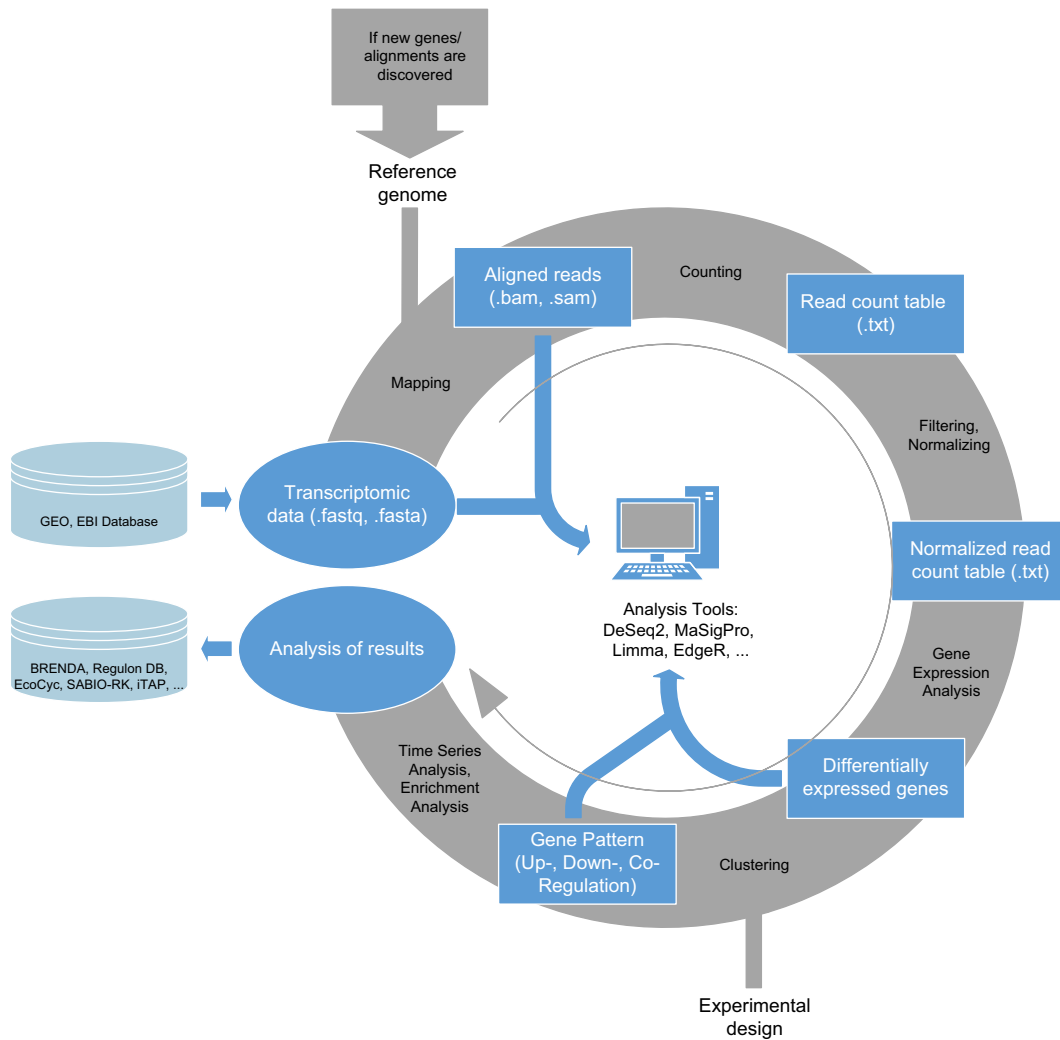


Fig. 2. Workflow illustrating the general procedure when analyzing gene expression data. RNA Seq analysis creates FASTA and FASTQ as raw data formats compiled in gene expression repositories (GEO, EBI), followed by SAM or BAM files for aligned reads. The analysis steps are, in general: (1) Mapping the reference genome onto the transcriptomic data, (2) Counting reads, (3) Filtering low read counts and normalizing counts, (4) Gene Expression Analysis, (5) Clustering, (6) Time Series Analysis and Enrichment Analysis. DeSeq2, MaSigPro, limma and edgeR are often applied packages within the language R to analyze transcriptomic data, in case of differential gene expression as well as gene pattern analysis. The resulting dynamic expressions and parameters are stored in databases like BRENDA [44], Regulon DB [45], EcoCyc [46], SABIO-RK [47] and iTAP [48].

models like this are limited in scale, due to the complexity of the cellular mechanisms and the single cell consideration, which results in a quadratic scaling problem.

3.1. Identifying Structured and Non-Structured Microbial Models

Non-segregated, structured models typically consist of a rigid network structure and a set of rate expressions including sensitive parameters. Knowledge of the network structure, the kinetic equations and the parameters is key to identifying a proper model. Often, such structures are determined following the bottom-up approach, i.e. the statistically profound identification of correlations between the structuring elements based on experimental data. The bottom-up concept can be applied to merge already existing small-scale models into large models [72–74]. Alternatively, top-down approaches aim for the identification of model parameters for a given structure. Accordingly, the top-down approach is a powerful tool for deciphering details of pathway interaction with the network, provided that the given structure is correct [75, 76]. Table 1 depicts a comparison of the two approaches, including prediction goals and limitations. Statements hold true irrespective of whether model complexity is limited to metabolic interactions or

whether superior regulation levels such as transcriptional or post-translational feedbacks are included.

3.2. Identifying Gene Regulatory Networks (GRNs)

When cells are exposed to dynamic stimuli, such as the fluctuating micro-environmental changes in large-scale bioreactors, they show short- and long-term physiological responses. Whereas the first are

Table 1
Comparison of bottom-up and top-down approach.

	Bottom-up	Top-down
Design steps	From single molecule to pattern	From single elements to relations
Model size	Small-scale	Coarse-grained large-scale Global
Model complexity	Detailed	Global cellular dynamics
Prediction goals	Detailed time-scale resolution	Neglect of single reaction steps
Limitations	Lack of kinetic parameters and in-depth knowledge	

dominated by metabolic interactions, the second include strategies for microbial adaptation usually comprising changes of transcriptome and proteome. However, recent findings [14–16, 54] have shown that transcriptional response occurs massively, even during short-term, sub-minute periods. Accordingly, GRN models gain importance even for modeling short-term responses which explicitly motivates their use.

A gene regulatory network links transcription factors to their target genes thereby creating a dynamic interaction map connecting external stimuli with internal transcriptional and even metabolic responses. Accordingly, GRN models may comprise the signal stimulus, its transduction to the receptor, the transcriptional response and downstream processes such as translation, post-translational modifications of protein and protein degradation. Altogether, these interactions form a very complex regulatory network. Roughly, the plentitude of GRNs may be divided into three representative approaches: continuous models (in this case based on ordinary differential equations) [71, 86–91], Boolean models [92–95] and probabilistic models [96–99]. These and other methods, such as Petri nets, Bayesian networks or neural networks, have been extensively reviewed by Karlebach and Shamir [100] and Machado et al. [101].

Stochastic models start from the assumption that gene expression should be described by random events e.g. caused by the shortage of mRNA molecules and factors of transcription. Similarly, initiation and elongation factors are scarce and may cause stochastic translation processes. Accordingly, the continuum paradigm, i.e. the assumption of a sufficient, homogeneous availability of each model entity, may be questionable and could be checked using the chemical master equation. In case the number of molecules per cell is too low, stochastic models should be considered [102]. Pragmatic guidelines have been published by Kremling et al. [103] and Turner et al. [97], identifying molecule numbers of about 100 per cell as the threshold value.

Alternatively, systems of ordinary differential equations (ODEs) can be applied ignoring stochastic transcription events and assuming cellular continuum instead. Cellular entities are simulated as continuous time courses. Such models require knowledge about gene regulatory mechanisms to select appropriate rate laws and to identify corresponding rate parameters for parameter estimation. Standardized formats simplify the development process and encourage the automatic construction of kinetic models. However, the prediction quality of the model may deteriorate for the prediction of cellular states that are not reflected by the experimental data. Checking thermodynamic feasibility for estimated fluxes is strongly encouraged to prevent misleading findings [104]. If only a small amount of data is available, Boolean approaches may be useful to model regulatory networks. As a key feature, Boolean models consider *on/off* activation and the inhibition of transcription factors and genes. Accordingly, such models are helpful to predict *on/off*-like gene switching but fail to simulate distinct time series of transcriptional dynamics that occur after frequent stimulations in large-scale bioreactors.

Moreover, Mochizuki [105] showed that high prediction qualities and the easy handling of Boolean models is limited to small-scale models. Consequently, large-scale GRNs should preferably be composed of ODEs [100]. A short summary of the different class of models with their specific advantages and disadvantages is given in Table 2.

Inferring gene regulatory networks from gene expression data remains a challenging task due to the large number of potential interactions, the relatively small number of available measurements and the intrinsic noise often caused by the biological variance which reflects the heterogeneity of the cell population. Despite success with automated model set-up and identification, manual curation of the inferred network interactions can become time intensive and cumbersome due to the amount of data investigated [106].

To achieve high prediction quality, kinetic parameters need to be accurate and sensitive, i.e. parameter variance should be low and parameter sensitivity should be high to enable highly accurate model prediction with the lowest amount of parameters necessary. Parameter values may be extracted from experimental data, taken from public databases or already existing dynamic models and kinetics used.

Parameters of the GRN model can be deduced from experimental data usually applying least-square error estimation. In essence, parameter estimation is an optimization problem which minimizes the weighted squared distance between simulation and experimental observation to achieve a parameter set for the least squares. Such approaches can be combined with model discrimination methods [107]. In case the parameter estimation problem does not have a unique solution, the solution space can be further constrained using thermodynamic laws or by expanding the experimental basis taking into account other experimental conditions to challenge the applicability of the model [108]. Parameter estimation methods have been reviewed by Lillacci and Khammash [109].

For example, small-scale regulatory networks of *E. coli* [71, 87, 110] and also large or genome-scale regulatory networks [98, 111–113] have already been published. In general, applicants should pay attention to the transferability of the models because reference conditions may be different compared to the current case which is likely to cause the improper extrapolation of experimental findings.

Once a model has been identified, its validity needs to be checked against new data sets that were not used for parameter identification. When models can successfully simulate such new data, it is a strong indication that the mechanistic principles and assumptions behind the model are sound. If a model fails to pass the validation step, the modeler needs to revise the previous steps of their modeling process.

Recent examples of data-driven GRN models are given by Erickson et al. [76], Palsson and Nielson [114–117] and others [98, 118–122]. For example, Klipp et al. [72, 123] describes bacterial growth transitions considering the proteome level, the complex interactions of the yeast cell cycle and the prediction of complex regulatory patterns following the mindset of optimized resource allocation in yeasts, respectively. The current developments of metabolism and gene expression (ME) models are in line with the pioneering approach of Chassagnole et al. [71] who published a comprehensive dynamic model of the central metabolism in *E. coli*.

4. Simulating the Cellular Environment With Embedded Growing Cells

Large-scale bioreactor conditions need to be calculated, aiming at a spatial resolution of the mass, momentum and energy balances via

Table 2
Summary of advantages, disadvantages and application of the above mentioned methods.

Model class	Application	Advantages	Disadvantages
Structured	Systems in transient state	Cellular compartmentation	Biological knowledge
Non-structured	Steady-state systems	Easy to build	Only phenomenological cell description
Segregated	Heterogeneous, individual cell systems	More representative and informative	Difficult to handle mathematically
Non-segregated	Systems with average cell description	Easy to build, for a large number of cells	Average cell description
Probabilistic	Randomly distributed events in time	Realistic behaviour	Only for low number of molecules in cell
Continuous	Evenly distributed events in time (cellular continuum)	Large-scale possible	Detailed knowledge about mechanisms and biological parameters
Boolean	Discrete dynamical system	Small amount of experimental data (On/Off conditions)	Fails to simulate distinct time series

numerical simulation. In particular, the Navier-Stokes equations (NSE) representing the conservation of momentum, the continuity equation representing the conservation of mass and the energy equation predicting the temperature in the fluid of a multiphase system have to be considered. The Navier-Stokes equations basically describe the motion of viscous fluid flows where the fluids are considered as a continuum rather than a number of colliding particles. Under the typical mixing conditions given, the occurrence of turbulent zones is likely. Turbulence is defined as a state consisting of structures such as eddies which affect molecular diffusion, heat transfer and the mixing behavior.

4.1. Modeling of Hydrodynamics and Mass Transfer

To describe hydrodynamic turbulence, multiple suggestions have been published, often applying the modified Reynolds-averaged Navier-Stokes equation (RANS) for multiphase systems [124]. Other alternative approaches, such as Large Eddy Simulation (LES) and Direct Numerical Simulation (DNS), offer increased accuracy, but require immense computational capacity [125, 126] as displayed in Fig. 3.

The Reynolds-averaged Navier-Stokes equations are time-averaged equations of motion for turbulent flows approximating different turbulent scales through fluctuating quantities, an idea first proposed by Reynolds [127]. RANS models offer the most economic approach for simulating complex turbulent flows, because turbulences are considered with different levels of complexity. The most common RANS turbulence models are classified with respect to the number of additional transport equations that need to be solved along with the RANS flow equation. Besides, the often used two-equation models, such as the standard $k-\epsilon$, $k-\omega$ or Renormalization group $k-\epsilon$ models, one-equation models (low-cost RANS models, e.g. the Spalart-Allmaras approach) or even zero-equation models which estimate the turbulence viscosity via the mean velocity and the length scale using an empirical formula are available [128]. Details are given in the review of Rodi [129].

In addition to the simulation of turbulence, the proper modeling of interactions between different phases (e.g. aqueous media, air bubbles and solid cells) is a challenge. Table 3 provides an overview of three common approaches.

Another way to predict hydrodynamics is the use of compartment models (CM) [130]. Characteristically, the reactor is divided into a subset of spatial parts, each assumed to be ideally mixed, see Fig. 4.

Compartment models are much less computationally demanding than CFD simulations and moreover, allow easy implementation of complex reaction schemes. Fluxes between the compartments are

often defined by considering global quantities which are not representative of the flow complexity. Moreover, incoming concentrations are instantly ideally mixed in the whole compartment and erratic changes occur, which are not observed in reality [24].

Recently, the combination of CFD and CM modeling has been presented to couple the accuracy of hydrodynamic CFD simulations with the simplicity and speed of compartmented modelings [24, 131–133]. As shown in Fig. 5, the approach can be applied for describing concentration gradients in industrial-scale bioreactors, calculating the inter-compartmental fluxes from CFD velocity fields. Characteristically, turbulent liquid flows are computed by CFD first, followed by the implementation of net mean and turbulent flow rates in the compartment approach. The simplicity of the approach even allows complex genome-scale kinetic models to be used.

Likewise for gene regulatory models, fluid flow simulations must be validated based on independent experimental data. However, experimental observation of large-scale hydrodynamics is often lacking, which limits the comparison with the predicted flow patterns.

4.2. Hydrodynamic Modeling Linked to GRN Models

The physiological state of microorganisms and its impact on growth and product formation is the result of a complex interaction between the cellular environment and the cells. Large-scale studies have shown that homogeneous culture conditions are difficult to establish, nevertheless process engineering and bioreactor design may aim to create the least heterogeneous impact possible [134].

So far, large-scale simulations almost entirely focused on the integration of metabolism kinetics. They basically mirror the instantaneous cellular response on environmental changes [21, 143, 144]. However, cells react in a multi-response, multi-layer fashion also comprising the on- and offset of transcriptional regulation programs. Such responses are triggered in poorly mixed zones and are propagated into well mixed zones [14, 99]. Initiation and execution may be spatially disconnected which differs fundamentally from the metabolic responses studied so far.

To investigate the consequences of environmental heterogeneities, proper modeling frameworks should link local variations with cellular and subcellular kinetics.

The tool of choice is CFD simulation which can link the interaction of cellular activities with local environments [22, 135, 136].

Regarding the Euler-Euler approach, the liquid phase and the microorganisms are considered as a continuum. A continuum is a continuous system which does not allow erratic changes [137]. However,

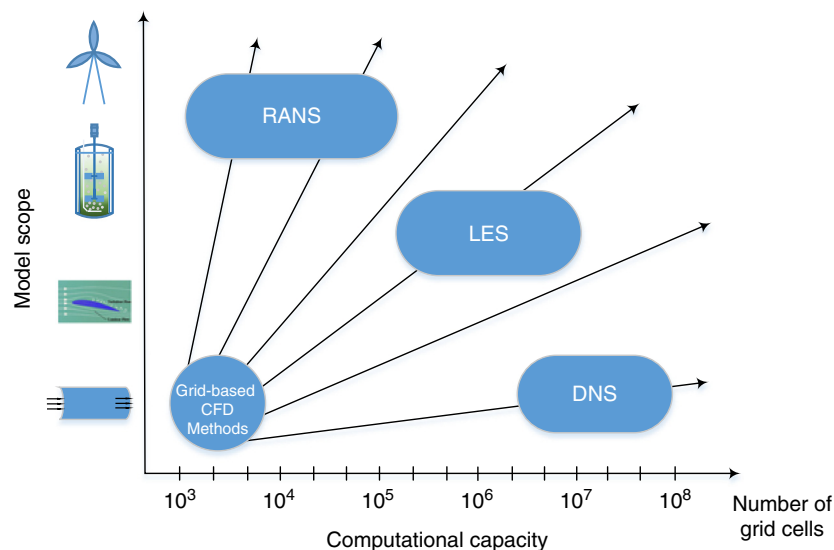


Fig. 3. Different approaches of turbulence models, regarding the model scope and fields of application, as well as the corresponding computational capacity required. RANS: Reynolds-averaged Navier-Stokes equation, LES: Large Eddy Simulation, DNS: Direct Numerical Simulation.

Table 3
Comparison of CM and CFD model approaches.

	CFD		Compartmentation
	Euler-Euler	Euler-Lagrange	
Prediction purpose	Reactor heterogeneity	Cell-Environment Interaction	Reactor design
Computational effort	High	High	Low
Level of detail	High	High	Low
Prediction accuracy of flow regime	High	High	Low
Single cell tracking	No	Yes	Yes
Integrable model size	Coarse-grained small-scale	Coarse-grained small-scale	Genome-scale
Amount of particles	High	Low (<10%)	High

microorganisms are individual in their behavior and therefore the continuum description is a greatly simplified assumption. As a result, the continuum approach leads to a lack of individual responses of the cells.

Conventional Euler-Euler approaches of two-phase flow scenarios can be extended considering Population Balance Equations (PBEs) with unstructured kinetic growth models [138–141]. PBEs are used to model population adaptation dynamics considering nutrient gradients inside large-scale bioreactors. In general, Euler-Euler approaches in combination with PBEs are suited to model particle (cell) swarms that follow flow patterns in the reactor [142].

However, an inherent limitation of PBEs is that the incorporation of a detailed kinetic network leads to massive computational effort because of high dimensional distribution functions that need to be solved.

Additionally, no information on the level of single particles, such as their lifelines and history can be obtained with this approach.

This limitation can be tackled by using the Euler-Lagrange approach, which tracks the fate of each particle (cell) individually. The Lagrangian implementation requires detailed metabolic models of the cell, e.g. to describe the transport processes across the cellular membrane, via substrate uptake rates and product excretion rates. For simplification, massless cells are often used which are described via Monod-like black-box models. Such cells are assumed to travel along the flow fields thereby experiencing substrate gradients.

Notably, the cellular environment is typically ‘frozen’, i.e. fundamental cellular reactions are implemented in the Euler continuum and traveling cells only respond to the given hydrodynamic and concentration gradients [143]. Pioneering studies have been performed by Lapin et al. [18] and have been elaborated further in many follow-up studies [21, 143–145].

Such studies clearly outline that cells are subject to repetitive and fast changes which in turn create heterogeneity within the population.

However, the computational effort for the spatial resolution of the conservation equations is high, requiring smart compositions of the computational grid and, for simplifying, assumptions to solve the numerical problem. Recently, Chen et al. [146] used a rather simple CM approach to simulate a syngas fermentation of *Clostridium ljungdahlii*. As a result, they could show that multi-compartment approaches, even if not widely used yet, give good results regarding the interaction of rather complex cells with their environment. Thus, in situations where a simple model structure meets the requirements of the modeling purpose, non-essential details should be avoided since they will unnecessarily prolong the modeling process.

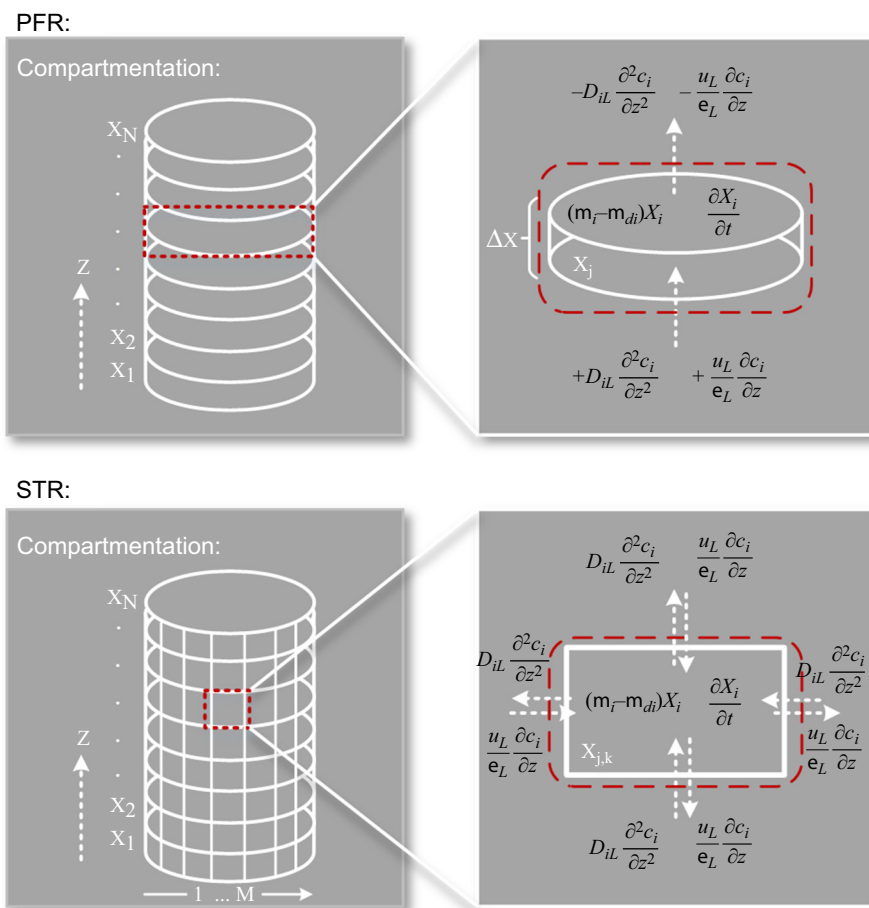


Fig. 4. Compartmentation of a PFR and STR. Each section is homogeneously mixed and represented by a partial differential equation that usually has an accumulation term, a source/sink term due to bacterial growth and death, and convection and diffusion terms (in/out) which describe the environmental conditions. Other than the PFR, the STR needs finer discretization due to tangential mixing.

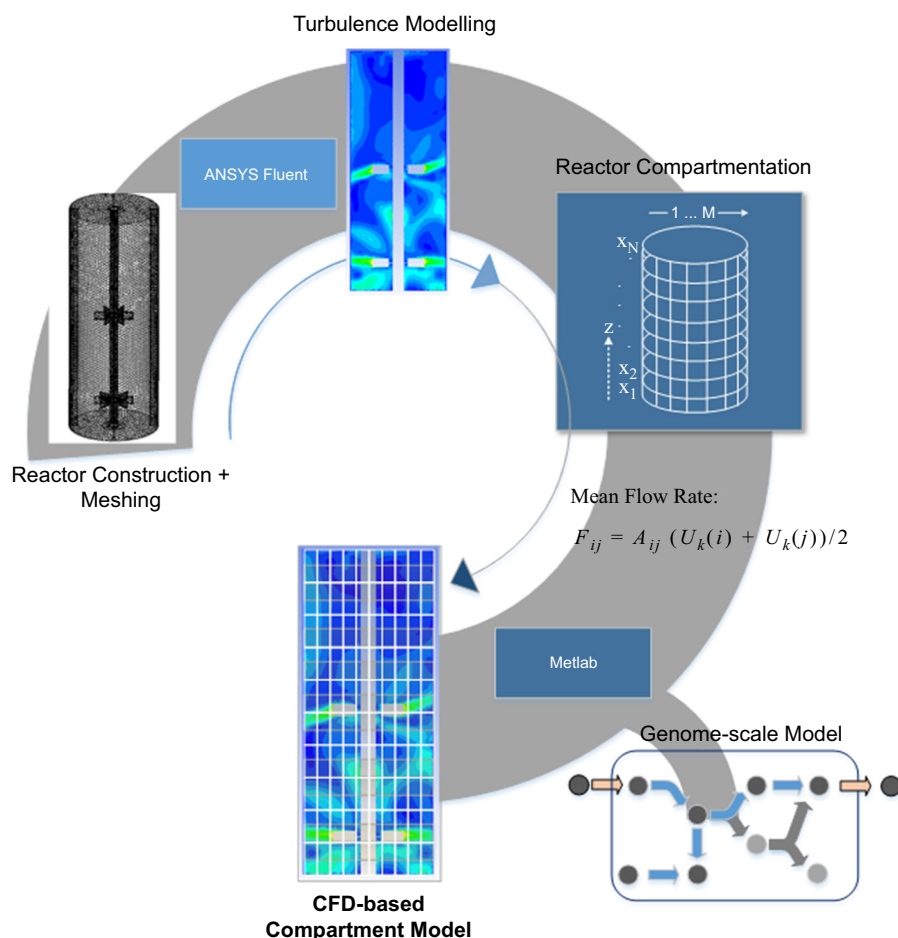


Fig. 5. CFD-based compartment model set-up steps. (1) Construction and meshing of the reactor, (2) CFD simulation of the turbulent flow (velocity components, dissipation rate, turbulent kinetic energy), (3) Definition of reactor compartmentation, (4) Definition of mean flow F_{ij} between cell i and cell j , where $U_k(i)$ is the velocity component and A_{ij} is the area of the face between the two cells, (5) Incorporation of genome-scale model. Common simulation frameworks: Step (1) & (2): ANSYS Fluent, Step (3), (4) & (5): Matlab.

5. Conclusion and Perspectives

Understanding the function of cellular behavior under varying conditions requires the development of computational approaches that incorporate gene regulatory models as well as environmental perturbation simulations relying on reliable experimental evidence.

On the one hand, due to efficient large-scale simulations and stimulus/response experiments, experimental findings have revealed a complex organization of regulatory response in the cell and improved the understanding of several regulatory processes. To further expand this understanding, development towards single cell resolution techniques has evolved. Although this is at the very beginning, this topic has significant potential for further developments regarding reactor design and genetic engineering towards robust strains.

On the other hand, numerical gene regulatory models based on ODE systems or modeling on a single molecular level with stochastic algorithms in combination with hydrodynamic simulations provided a broad and detailed insight into the regulatory mechanisms of microorganisms inside large-scale bioreactors. But due to a lack of large-scale experimental data, many regulation theories are still based to some extent on empirical observations. To date, hydrodynamic simulations as well as kinetic cellular models are available with different scales of complexity which favors the usability regarding the computational effort.

It could also be shown that a combination of already existing methods is often advantageous, such as CFD-based compartment models, providing the possibility of combining genome-scale models with hydrodynamic simulations.

Based on the extended variety and good results of cellular and hydrodynamic modeling approaches and the availability of reliable experimental data allowing detailed insight into cellular mechanisms, the time is right to use and combine these methods to predict the large-scale performance of microbial producers.

However, the above discussion has highlighted the need for knowledge-based process scale-up by elucidating the putative contributions of modeling. The contribution of numerical simulations also warrants further investigation with *in vivo* experiments that incorporate large-scale conditions and single cell resolution. The development towards automated high resolution processes and the detection of single cell behavior is a promising trend. This review shows that the basis to predict *in silico* large-scale performance of microbial producers is given. As a result, robust strains, as well as reactor design parameters and optimized cultivation conditions for more efficient processes can be developed.

Conflict of Interest

None declared.

References

- [1] Bailey J. Toward a science of metabolic engineering. *Science* 1991;252(5013): 1668–75.
- [2] Vallino JJ, Stephanopoulos G. Metabolic flux distributions in *Corynebacterium glutamicum* during growth and lysine overproduction. *Biotechnol Bioeng* 1993; 41(6):633–46.

- [3] Jacob F, Monod J. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* 1961;3(3):318–56.
- [4] Kitano H. Computational cellular dynamics: a network-physics integral. *Nat Rev Mol Cell Biol* 2006;7(3):163–4.
- [5] Westerhoff HV, Palsson BO. The evolution of molecular biology into systems biology. *Nat Biotechnol* 2004;22(10):1249–52.
- [6] Lee KH, Park JH, Kim TY, Kim HU, Lee SY. Systems metabolic engineering of *Escherichia coli* for L-threonine production. *Mol Syst Biol* 2007;3(1):149.
- [7] Park JH, Lee SY. Towards systems metabolic engineering of microorganisms for amino acid production. *Curr Opin Biotechnol* 2008;19(5):454–60.
- [8] Becker J, Zelder O, Häfner S, Schröder H, Wittmann C. From zero to hero. Design-based systems metabolic engineering of *Corynebacterium glutamicum* for L-lysine production. *Metab Eng* 2011;13(2):159–68.
- [9] Wittmann C, Lee SY. Systems metabolic engineering. Springer Science & Business Media; 2012.
- [10] Oosterhuis N, Groesbeek N, Olivier A, Kossen N. Scale-down aspects of the gluconic acid fermentation. *Biotechnol Lett* 1983;5(3):141–6.
- [11] Oosterhuis N, Kossen N. Dissolved oxygen concentration profiles in a production-scale bioreactor. *Biotechnol Bioeng* 1984;26(5):546–50.
- [12] Neubauer P, Häggström L, Enfors SO. Influence of substrate oscillations on acetate formation and growth yield in *Escherichia coli* glucose limited fed-batch cultivations. *Biotechnol Bioeng* 1995;47(2):139–46.
- [13] Buchholz J, Graf M, Freund A, Busche T, Kalinowski J, Blombach B, et al. Co2/hco3-perturbations of simulated large scale gradients in a scale-down device cause fast transcriptional responses in *Corynebacterium glutamicum*. *Appl Microbiol Biotechnol* 2014;98(20):8563–72.
- [14] Löffler M, Simen JD, Jäger G, Schäferhoff K, Freund A, Takors R. Engineering *E. coli* for large-scale production—strategies considering atp expenses and transcriptional responses. *Metab Eng* 2016;38:73–85.
- [15] Löffler M, Simen JD, Müller J, Jäger G, Laghrami S, Schäferhoff K, et al. Switching between nitrogen and glucose limitation: unraveling transcriptional dynamics in *Escherichia coli*. *J Biotechnol* 2017;285:2–12.
- [16] von Wulffen J, Ulmer A, Jäger G, Sawodny O, Feuer R. Rapid sampling of *Escherichia coli* after changing oxygen conditions reveals transcriptional dynamics. *Gene* 2017;8(3):90.
- [17] Bylund F, Collet E, Enfors SO, Larsson G. Substrate gradient formation in the large-scale bioreactor lowers cell yield and increases by-product formation. *Bioprocess Eng* 1998;18(3):171–80.
- [18] Lapin A, Müller D, Reuss M. Dynamic behavior of microbial populations in stirred bioreactors simulated with euler-lagrange methods: traveling along the lifelines of single cells. *Ind Eng Chem Res* 2004;43(16):4647–56.
- [19] Lapin A, Schmid J, Reuss M. 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* 2006;61(14):4783–97.
- [20] Wang G, Tang W, Xia J, Chu J, Noorman H, Gulik WM. Integration of microbial kinetics and fluid dynamics toward model-driven scale-up of industrial bioprocesses. *Eng Life Sci* 2015;15(1):20–9.
- [21] Haringa C, Tang W, Deshmukh AT, Xia J, Reuss M, Heijnen JJ, et al. Euler-lagrange computational fluid dynamics for (bio) reactor scale down: an analysis of organism lifelines. *Eng Life Sci* 2016;16(7):652–63.
- [22] Delvigne F, Takors R, Mudde R, Gulik W, Noorman H. Bioprocess scale-up/down as integrative enabling technology: from fluid mechanics to systems biology and beyond. *Microb Biotechnol* 2017;10(5):1267–74.
- [23] Takors R. Scale-up of microbial processes: impacts, tools and open questions. *J Biotechnol* 2012;160(1):3–9.
- [24] Delafosse A, Collignon ML, Calvo S, Delvigne F, Crine M, Thonart P, et al. Cfd-based compartment model for description of mixing in bioreactors. *Chem Eng Sci* 2014;106:76–85.
- [25] Delvigne F, Destain J, Thonart P. A methodology for the design of scale-down bioreactors by the use of mixing and circulation stochastic models. *Biochem Eng J* 2006;28(3):256–68.
- [26] Neubauer P, Junne S. Scale-down simulators for metabolic analysis of large-scale bioprocesses. *Curr Opin Biotechnol* 2010;21(1):114–21.
- [27] Lemoine A, Maya Martínez-Iturralde N, Spann R, Neubauer P, Junne S. Response of *Corynebacterium glutamicum* exposed to oscillating cultivation conditions in a two- and a novel three-compartment scale-down bioreactor. *Biotechnol Bioeng* 2015;112(6):1220–31.
- [28] Xu B, Jahic M, Blomsten G, Enfors SO. Glucose overflow metabolism and mixed-acid fermentation in aerobic large-scale fed-batch processes with *Escherichia coli*. *Appl Microbiol Biotechnol* 1999;51(5):564–71.
- [29] Egli T. On multiple-nutrient-limited growth of microorganisms, with special reference to dual limitation by carbon and nitrogen substrates. *Antonie Van Leeuwenhoek* 1991;60(3–4):225–34.
- [30] Oldiges M, Lütz S, Pflug S, Schroer K, Stein N, Wiendahl C. Metabolomics: current state and evolving methodologies and tools. *Appl Microbiol Biotechnol* 2007;76(3):495–511.
- [31] Teleki A, Sánchez-Kopper A, Takors R. Alkaline conditions in hydrophilic interaction liquid chromatography for intracellular metabolite quantification using tandem mass spectrometry. *Anal Biochem* 2015;475:4–13.
- [32] Pfizenmaier J, Junghans L, Teleki A, Takors R. Hyperosmotic stimulus study discloses benefits in atp supply and reveals mirna/mrna targets to improve recombinant protein production of cho cells. *Biotechnol J* 2016;11(8):1037–47.
- [33] Matuszczyk JC, Teleki A, Pfizenmaier J, Takors R. Compartment-specific metabolomics for cho reveals that atp pools in mitochondria are much lower than in cytosol. *Biotechnol J* 2015;10(10):1639–50.
- [34] Fernie AR, Trethewey RN, Krotzky AJ, Willmitzer L. Metabolite profiling: from diagnostics to systems biology. *Nat Rev Mol Cell Biol* 2004;5(9):763.
- [35] Winder CL, Dunn WB, Schuler S, Broadhurst D, Jarvis R, Stephens GM, et al. Global metabolic profiling of *Escherichia coli* cultures: an evaluation of methods for quenching and extraction of intracellular metabolites. *Anal Chem* 2008;80(8):2939–48.
- [36] Fiehn O. Metabolomics, the link between genotypes and phenotypes. *Functional Genomics*. Springer; 2002. p. 155–71.
- [37] Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* 1999;17(10):994.
- [38] Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature* 2003;422(6928):198.
- [39] Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, et al. The transcriptional landscape of the yeast genome defined by rna sequencing. *Science* 2008;320(5881):1344–9.
- [40] Nookaew I, Papini M, Pornputtapong N, Scalinati G, Fagerberg L, Uhlén M, et al. A comprehensive comparison of rna-seq-based transcriptome analysis from reads to differential gene expression and cross-comparison with microarrays: a case study in *Saccharomyces cerevisiae*. *Nucleic Acids Res* 2012;40(20):10084–97.
- [41] Wang Z, Gerstein M, Snyder M. Rna-seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 2009;10(1):57–63.
- [42] Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, et al. A survey of best practices for rna-seq data analysis. *Genome Biol* 2016;17(1):13.
- [43] Scheer M, Grote A, Chang A, Schomburg I, Munaretto C, Rother M, et al. Brenda, the enzyme information system in 2011. *Nucleic Acids Res* 2010;39(Suppl_1):D670–6.
- [44] Salgado H, Gama-Castro S, Martínez-Antonio A, Díaz-Peredo E, Sánchez-Solano F, Peralta-Gil M, et al. Regulondb (version 4.0): transcriptional regulation, operon organization and growth conditions in *Escherichia coli* k-12. *Nucleic Acids Res* 2004;32(1):D303–6 suppl.
- [45] Keseler IM, Mackie A, Peralta-Gil M, Santos-Zavaleta A, Gama-Castro S, Bonavides-Martínez C, et al. Ecocyc: fusing model organism databases with systems biology. *Nucleic Acids Res* 2013;41(D1):D605–12.
- [46] Wittig U, Kania R, Golebiewski M, Rey M, Shi L, Jong L, et al. Sabio-rk—database for biochemical reaction kinetics. *Nucleic Acids Res* 2011;40(D1):D790–6.
- [47] Sundararaman N, Ash C, Guo W, Button R, Singh J, Feng X. itap: integrated transcriptomics and phenotype database for stress response of *Escherichia coli* and *Saccharomyces cerevisiae*. *BMC Research Notes* 2015;8(1):771.
- [48] Rapaport F, Khanin R, Liang Y, Pirun M, Krek A, Zumbo P, et al. Comprehensive evaluation of differential gene expression analysis methods for rna-seq data. *Genome Biol* 2013;14(9):3158.
- [49] Hecker M, Lambeck S, Toepfer S, Van Someren E, Guthke R. Gene regulatory network inference: data integration in dynamic models, a review. *Biosystems* 2009;96(1):86–103.
- [50] Banf M, Rhee SY. Computational inference of gene regulatory networks: Approaches, limitations and opportunities. *Biochimica et Biophysica Acta (BBA)–Gene Regulatory Mechanisms* 2017;1860(1):41–52.
- [51] Love M, Anders S, Huber W. Differential analysis of count data—the DESeq2 package. *Genome Biol* 2014;15:550.
- [52] Conesa A, Nueda MJ, Ferrer A, Talón M. masigpro: a method to identify significantly differential expression profiles in time-course microarray experiments. *Bioinformatics* 2006;22(9):1096–102.
- [53] Simen JD, Löffler M, Jäger G, Schäferhoff K, Freund A, Matthes J, et al. Transcriptional response of *Escherichia coli* to ammonia and glucose fluctuations. *Microb Biotechnol* 2017(10):858–72.
- [54] Brown DR, Barton G, Pan Z, Buck M, Wigneshweraraj S. Nitrogen stress response and stringent response are coupled in *Escherichia coli*. *Nat Commun* 2014;5.
- [55] Liu X, Yang S, Wang F, Dai X, Yang Y, Bai Z. Comparative analysis of the *Corynebacterium glutamicum* transcriptome in response to changes in dissolved oxygen levels. *J Ind Microbiol Biotechnol* 2017;44(2):181–95.
- [56] Caspeta L, Flores N, Pérez NO, Bolívar F, Ramírez OT. The effect of heating rate on *Escherichia coli* metabolism, physiological stress, transcriptional response, and production of temperature-induced recombinant protein: A scale-down study. *Biotechnol Bioeng* 2009;102(2):468–82.
- [57] Michalowski A, Siemann-Herzberg M, Takors R. *Escherichia coli* hgt: Engineered for high glucose throughput even under slowly growing or resting conditions. *Metab Eng* 2017;40:93–103.
- [58] Müller S, Davey H. Recent advances in the analysis of individual microbial cells. *Cytometry Part A* 2009;75(2):83–5.
- [59] Broger T, Odermatt RP, Huber P, Sonnleitner B. Real-time on-line flow cytometry for bioprocess monitoring. *J Biotechnol* 2011;154(4):240–7.
- [60] Brognaux A, Han S, Sørensen SJ, Lebeau F, Thonart P, Delvigne F. A low-cost, multiplexable, automated flow cytometry procedure for the characterization of microbial stress dynamics in bioreactors. *Microb Cell Factories* 2013;12(1):100.
- [61] Dusny C, Fritzsche FSO, Frick O, Schmid A. Isolated microbial single cells and resulting microcolonies grow faster in controlled environments. *Appl Environ Microbiol* 2012;78(19):7132–6.
- [62] Taheri-Araghi S, Jun S. Single-cell cultivation in microfluidic devices. *Hydrocarbon and Lipid Microbiology Protocols: Single-Cell and Single-Molecule Methods* 2016:5–16.
- [63] Westerwalbesloh C, Grünberger A, Kohlheyer D, von Lieres E. Modeling inhomogeneities across cultivation chamber arrays in single-cell cultivation devices; 2016.
- [64] Bennett MR, Hasty J. Microfluidic devices for measuring gene network dynamics in single cells. *Nat Rev Genet* 2009;10(9):628–38.
- [65] Lemoine A, Delvigne F, Bockisch A, Neubauer P, Junne S. Tools for the determination of population heterogeneity caused by inhomogeneous cultivation conditions. *J Biotechnol* 2017;251:84–93.

- [67] Bossert M, Kracht D, Scherer S, Landstorfer R, Neuhaus K. Improving the reliability of rna-seq: Approaching single-cell transcriptomics to explore individuality in bacteria. Information-and Communication Theory in Molecular Biology. Springer; 2018. p. 181–98.
- [68] Bailey JE. Mathematical modeling and analysis in biochemical engineering: past accomplishments and future opportunities. Biotechnol Prog 1998;14(1):8–20.
- [69] Nielsen J, Nikolajsen K, Villadsen J. Structured modeling of a microbial system: I. a theoretical study of lactic acid fermentation. Biotechnol Bioeng 1991;38(1):1–10.
- [70] Tang W, Deshmukh AT, Haringa C, Wang G, van Gulik W, van Winden W, et al. A 9-pool metabolic structured kinetic model describing days to seconds dynamics of growth and product formation by *Penicillium chrysogenum*. Biotechnol Bioeng 2017;114:1733–43.
- [71] Chassagnole C, Noisommit-Rizzi N, Schmid JW, Mauch K, Reuss M. Dynamic modeling of the central carbon metabolism of *Escherichia coli*. Biotechnol Bioeng 2002;79(1):53–73.
- [72] Klipp E, Nordlander B, Krüger R, Gennemark P, Hohmann S. Integrative model of the response of yeast to osmotic shock. Nat Biotechnol 2005;23(8):975.
- [73] Guido NJ, Wang X, Adalsteinsson D, McMillen D, Hasty J, Cantor CR, et al. A bottom-up approach to gene regulation. Nature 2006;439(7078):856.
- [74] Brandon M, Howard B, Lawrence C, Laubenbacher R. Iron acquisition and oxidative stress response in *aspergillus fumigatus*. BMC Syst Biol 2015;9(1):19.
- [75] Chou IC, Voit EO. Recent developments in parameter estimation and structure identification of biochemical and genomic systems. Math Biosci 2009;219(2):57–83.
- [76] Erickson DW, Schink SJ, Patsalo V, Williamson JR, Gerland U, Hwa T. A global resource allocation strategy governs growth transition kinetics of *Escherichia coli*. Nature 2017;551(7678):119–23.
- [86] Bolouri H, Davidson EH. Modeling transcriptional regulatory networks. BioEssays 2002;24(12):1118–29.
- [87] Kremling A, Bettenbrock K, Gilles ED. Analysis of global control of *Escherichia coli* carbohydrate uptake. BMC Syst Biol 2007;1(1):42.
- [88] Lemuth K, Hardiman T, Winter S, Pfeiffer D, Keller M, Lange S, et al. Global transcription and metabolic flux analysis of *Escherichia coli* in glucose-limited fed-batch cultivations. Appl Environ Microbiol 2008;74(22):7002–15.
- [89] Hardiman T, Lemuth K, Siemann-Herzberg M, Reuss M. Dynamic modeling of the central metabolism of *E. coli*—linking metabolite and regulatory networks. Systems Biology and Biotechnology of *Escherichia coli*. Springer; 2009. p. 209–35.
- [90] Machado D, Costa RS, Ferreira EC, Rocha I, Tidor B. Exploring the gap between dynamic and constraint-based models of metabolism. Metab Eng 2012;14(2):112–9.
- [91] Khodayari A, Maranas CD. A genome-scale *Escherichia coli* kinetic metabolic model k-ecoli457 satisfying flux data for multiple mutant strains. Nat Commun 2016;7.
- [92] Thomas R. Boolean formalization of genetic control circuits. J Theor Biol 1973;42(3):563–85.
- [93] Kauffman S, Peterson C, Samuelsson B, Troein C. Random boolean network models and the yeast transcriptional network. Proc Natl Acad Sci 2003;100(25):14796–14799.
- [94] Davidich M, Bornholdt S. The transition from differential equations to boolean networks: a case study in simplifying a regulatory network model. J Theor Biol 2008;255(3):269–77.
- [95] Wang RS, Saadatpour A, Albert R. Boolean modeling in systems biology: an overview of methodology and applications. Phys Biol 2012;9(5):055001.
- [96] Qian H, Elson EL. Single-molecule enzymology: stochastic michaelis–menten kinetics. Biophys Chem 2002;101:565–76.
- [97] Turner TE, Schnell S, Burrage K. Stochastic approaches for modeling in vivo reactions. Comput Biol Chem 2004;28(3):165–78.
- [98] Chandrasekaran S, Price ND. Probabilistic integrative modeling of genome-scale metabolic and regulatory networks in *Escherichia coli* and *Mycobacterium tuberculosis*. Proc Natl Acad Sci 2010;107(41):17845–17850.
- [99] Nieß A, Löffler M, Simen JD, Takors R. Repetitive short-term stimuli imposed in poor mixing zones induce long-term adaptation of *E. coli* cultures in large-scale bioreactors: experimental evidence and mathematical model. Front Microbiol 2017;8:1195.
- [100] Karlebach G, Shamir R. Modeling and analysis of gene regulatory networks. Nat Rev Mol Cell Biol 2008;9(10):770–80.
- [101] Machado D, Costa RS, Rocha M, Ferreira EC, Tidor B, Rocha I. Modeling formalisms in systems biology. AMB Express 2011;1(1):45.
- [102] Barberis M, Beck C, Amoussouvi A, Schreiber G, Diener C, Herrmann A, et al. A low number of sic1 mrna molecules ensures a low noise level in cell cycle progression of budding yeast. Mol Biosyst 2011;7(10):2804–12.
- [103] Kremling A, Heermann R, Centler F, Jung K, Gilles E. Analysis of two-component signal transduction by mathematical modeling using the kdpd/kdpe system of *Escherichia coli*. Biosystems 2004;78(1):23–37.
- [104] Costa RS, Machado D, Rocha I, Ferreira E. Critical perspective on the consequences of the limited availability of kinetic data in metabolic dynamic modeling. IET Syst Biol 2011;5(3):157–63.
- [105] Mochizuki A. An analytical study of the number of steady states in gene regulatory networks. J Theor Biol 2005;236(3):291–310.
- [106] Margolin AA, Nemenman I, Basso K, Wiggins C, Stolovitzky G, Dalla Favera R, et al. Aracne: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. BMC Bioinformatics 2006;7(1):S7.
- [107] Degenring D, Froemel C, Dikta G, Takors R. Sensitivity analysis for the reduction of complex metabolism models. J Process Control 2004;14(7):729–45.
- [108] Almquist J, Cvijovic M, Hatzimanikatis V, Nielsen J, Jirstrand M. Kinetic models in industrial biotechnology—improving cell factory performance. Metab Eng 2014;24:38–60.
- [109] Lillacci G, Khammash M. Parameter estimation and model selection in computational biology. PLoS Comput Biol 2010;6(3):e1000696.
- [110] Hardiman T, Lemuth K, Keller MA, Reuss M, Siemann-Herzberg M. Topology of the global regulatory network of carbon limitation in *Escherichia coli*. J Biotechnol 2007;132(4):359–74.
- [111] Faria JP, Overbeek R, Xia F, Rocha M, Rocha I, Henry CS. Genome-scale bacterial transcriptional regulatory networks: reconstruction and integrated analysis with metabolic models. Brief Bioinform 2013;15(4):592–611.
- [112] Ma S, Kemmeren P, Gresham D, Statnikov A. De-novo learning of genome-scale regulatory networks in *S. cerevisiae*. PLoS One 2014;9(9):e106479.
- [113] Reed JL, Vo TD, Schilling CH, Palsson BO. An expanded genome-scale model of *Escherichia coli* k-12 (i jr904 gsm/gpr). Genome Biol 2003;4(9):R54.
- [114] O'Brien EJ, Lerman JA, Chang RL, Hyde DR, Palsson BO. Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. Mol Syst Biol 2013;9(1):693.
- [115] Liu JK, O'Brien EJ, Lerman JA, Zengler K, Palsson BO, Feist AM. Reconstruction and modeling protein translocation and compartmentalization in *Escherichia coli* at the genome-scale. BMC Syst Biol 2014;8(1):110.
- [116] Thiele I, Fleming RM, Que R, Bordbar A, Diep D, Palsson BO. Multiscale modeling of metabolism and macromolecular synthesis in *E. coli* and its application to the evolution of codon usage. PLoS One 2012;7(9):e45635.
- [117] Bordbar A, Monk JM, King ZA, Palsson BO. Constraint-based models predict metabolic and associated cellular functions. Nat Rev Genet 2014;15(2):107.
- [118] Lerman JA, Hyde DR, Latif H, Portnoy VA, Lewis NE, Orth JD, et al. In silico method for modeling metabolism and gene product expression at genome scale. Nat Commun 2012;3:929.
- [119] Gonçalves E, Bucher J, Ryll A, Niklas J, Mauch K, Klamt S, et al. Bridging the layers: towards integration of signal transduction, regulation and metabolism into mathematical models. Mol Biosyst 2013;9(7):1576–83.
- [120] Ma S, Minch KJ, Rustad TR, Hobbs S, Zhou SL, Sherman DR, et al. Integrated modeling of gene regulatory and metabolic networks in mycobacterium tuberculosis. PLoS Comput Biol 2015;11(11):e1004543.
- [121] Klosik DF, Grimbs A, Bornholdt S, Hütt MT. The interdependent network of gene regulation and metabolism is robust where it needs to be. Nat Commun 2017;8(1):534.
- [122] Arrieta-Ortiz ML, Hafemeister C, Bate AR, Chu T, Greenfield A, Shuster B, et al. An experimentally supported model of the bacillus subtilis global transcriptional regulatory network. Mol Syst Biol 2015;11(11):839.
- [123] Klipp E. Modeling dynamic processes in yeast. Yeast 2007;24(11):943–59.
- [124] Ahlstedt H, Lahtinen M. Calculation of flow field in a stirred tank with rushton turbine impeller. Proceedings of the Third CFX International Users Conference, Chesham, UK, vol. 30. ; 1996. p. 91–108.
- [125] Hoekstra A, Derksen J, Van Den Akker H. An experimental and numerical study of turbulent swirling flow in gas cyclones. Chem Eng Sci 1999;54(13–14):2055–65.
- [126] Hartmann H, Derksen J, Montavon C, Pearson J, Hamill I, Van den Akker H. Assessment of large eddy and rans stirred tank simulations by means of Ida. Chem Eng Sci 2004;59(12):2419–32.
- [127] Reynolds O. Iv. on the dynamical theory of incompressible viscous fluids and the determination of the criterion. Phil Trans R Soc Lond A 1895;186:123–64.
- [128] Fluent A. 12.0 theory guide. ; 5Ansys Inc.; 2009
- [129] Rodi W. Turbulence modeling and simulation in hydraulics: A historical review. J Hydraul Eng 2017;143(5):03117001.
- [130] Vrábel P, van der Lans RG, van der Schot FN, Luyben KCA, Xu B, Enfors SO. Cma: integration of fluid dynamics and microbial kinetics in modeling of large-scale fermentations. Chem Eng J 2001;84(3):463–74.
- [131] Bezzo F, Macchietto S, Pantelides C. General hybrid multizonal/cfd approach for bioreactor modeling. AIChE J 2003;49(8):2133–48.
- [132] Guha D, Dudukovic M, Ramachandran P, Mehta S, Alvare J. Cfd-based compartmental modeling of single phase stirred-tank reactors. AIChE J 2006;52(5):1836–46.
- [133] Le Moulec Y, Gentric C, Potier O, Leclerc J. Comparison of systemic, compartmental and cfd modeling approaches: application to the simulation of a biological reactor of wastewater treatment. Chem Eng Sci 2010;65(1):343–50.
- [134] Lara AR, Galindo E, Ramírez OT, Palomares LA. Living with heterogeneities in bioreactors. Mol Biotechnol 2006;34(3):355–81.
- [135] Kelly WJ. Using computational fluid dynamics to characterize and improve bioreactor performance. Biotechnol Appl Biochem 2008;49(4):225–38.
- [136] Noorman H. An industrial perspective on bioreactor scale-down: What we can learn from combined large-scale bioprocess and model fluid studies. Biotechnol J 2011;6(8):934–43.
- [137] Schmalzriedt S, Jenne M, Mauch K, Reuss M. Integration of physiology and fluid dynamics. Process Integration in Biochemical Engineering. Springer; 2003. p. 19–68.
- [138] Morchain J, Gabelle JC, Cockx A. A coupled population balance model and cfd approach for the simulation of mixing issues in lab-scale and industrial bioreactors. AIChE J 2014;60(1):27–40.
- [139] Heins AL, Lencastre Fernandes R, Gernaey KV, Lantz AE. Experimental and in silico investigation of population heterogeneity in continuous *Saccharomyces cerevisiae* scale-down fermentation in a two-compartment setup. J Chem Technol Biotechnol 2015;90(2):324–40.
- [140] Bouguettoucha A, Balanne B, Amrane A. Unstructured models for lactic acid fermentation—a review. Food Technol Biotechnol 2011;49(1):3.
- [141] Pateraki C, Almqvist H, Ladakis D, Lidén G, Koutinas AA, Vlysidis A. Modeling succinic acid fermentation using a xylose based substrate. Biochem Eng J 2016;114:26–41.
- [142] Wang T, Wang J, Jin Y. Population balance model for gas–liquid flows: Influence of bubble coalescence and breakup models. Ind Eng Chem Res 2005;44(19):7540–9.

- [143] Kuschel M, Siebler F, Takors R. Lagrangian trajectories to predict the formation of population heterogeneity in large-scale bioreactors. *Bioengineering* 2017;4(2):27.
- [144] Haringa C, Deshmukh AT, Mudde RF, Noorman HJ. Euler-lagrange analysis towards representative down-scaling of a 22 cubic meters aerobic *S. cerevisiae* fermentation. *Chem Eng Sci* 2017;16(7):652–63.
- [145] Westerwalbesloh C, Grünberger A, Stute B, Weber S, Wiechert W, Kohlheyer D, et al. Modeling and cfd simulation of nutrient distribution in picoliter bioreactors for bacterial growth studies on single-cell level. *Lab Chip* 2015;15(21):4177–86.
- [146] Chen J, Gomez JA, Höffner K, Phalak P, Barton PI, Henson MA. Spatiotemporal modeling of microbial metabolism. *BMC Syst Biol* 2016;10(1):21.