Integrated Design of Biopharmaceutical Manufacturing Processes: Operation Modes and Process Configurations for Monoclonal Antibody Production

Sara Badr, a Kozue Okamura, a Nozomi Takahashi, a Vera Ubbenjans, b Haruku Shirahata, a Hirokazu Sugiyama a*

a Department of Chemical System Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, 113-8656, Tokyo, Japan
b Faculty of Mechanical Engineering, RWTH Aachen University, Kackertasche 9, 52072 Aachen, Germany
sugiyama@chemsys.t.u-tokyo.ac.jp

Abstract
The market for monoclonal antibodies (mAb)s is rapidly expanding. They are however expensive, with complicated lengthy production processes. Process models for two bottleneck operating units up- and downstream of the production chain are presented. The models are validated using pilot scale and experimental literature data. Dynamic simulations of batch, continuous, and mixed mode operations are used to compare economic performance and productivity. A qualitative discussion of risk factors and possible operational hurdles is presented. Results show that continuous operations have a significant advantage in terms of processing times, but not regarding operating costs. To realize a cost advantage, careful consideration of process conditions and parameters is required. These simulations can be used to map the design space and show the range of favourable operating modes and conditions in a thorough sensitivity analysis study.

Keywords: Process design, Continuous operation, Techno-economic assessment, Upstream, Downstream

1. Introduction
Monoclonal antibodies (mAb)s are dominating the biopharmaceutical market, which is the fastest growing sector in the pharmaceutical industry (Butler and Meneses-Acosta, 2012; Walsh, 2018). They are used to treat, among others, certain types of cancer and auto-immune diseases and often with less side effects than available alternative treatments. The market for monoclonal antibodies has been rapidly expanding and is expected to double again over the next five years (Grilo and Mantalaris, 2019). These drug products are often expensive and have complicated lengthy production processes. To facilitate the expansion of the mAb market, to face the competition from biosimilars, and to ensure the availability of the required demand volume, the production process needs to be optimized to allow for higher production flexibility and lower costs. The production of mAbs is typically conducted through upstream cell cultivation processes and downstream purification processes to obtain the desired mAb product. Figure 1 shows a typical mAb production process.
Generally, continuous production of biopharmaceuticals can offer significant advantages in terms of increased production efficiency. However, continuous production can also suffer from higher operational problems. Many factors affect the performance of the production process. This work aims to explore the available design options and map favourable production modes at different production scales and sizes. Process models are developed and validated with experimental data from literature and pilot scale runs in Japan. Two process units: the main cultivation unit, and the capture chromatography unit are chosen as the main focus of this work. These units represent bottleneck operations in terms of time and cost requirements. Dynamic simulations of batch, continuous, and mixed mode operations are used to compare economic performance and productivity. The impact of different process configurations and operating modes is evaluated in terms of operating costs and production time with a qualitative discussion of risk factors and possible operational hurdles.

2. Process modelling and simulation

Upstream processes comprise activities of cell cultivation. Three cultivation modes are commonly employed: conventional batch cultivation; fed-batch, where nutrients are fed throughout the production batch to maintain glucose concentrations thus increasing batch periods and productivity; and perfusion mode where the cultivation media is constantly fed and the reactor output is removed at the same rate with cell retention in the system achieving continuous operation. Multi-column processes, e.g. periodic counter current chromatography (PCC) with alternating column loading and regeneration stages can be used to simulate continuous flow in downstream capture chromatography processes. Simulated moving bed (SMB) operation can reduce the required number of columns to two (Angarita et al., 2015; Karst et al., 2018). Interconnected columns can thus provide longer loading times and higher column utilizations.

Production data has been obtained from a pilot scale research facility in Japan for the production of mAbs from Chinese hamster ovary (CHO) cells. Different models for the large-scale cultivation and chromatography have been fitted against the production data and the appropriate models were selected and further validated with more experimental data from literature. The obtained parameters were then used to simulate and assess further production scenarios.

2.1. Cultivation models

Cultivation models were developed based on mass balances of key process components and Monod type model for cell growth and death rates as previously presented in literature.
Integrated design of mAb production processes: operation modes & configurations

(Kornecki and Strube, 2018; Xing et al., 2010). The models were adjusted based on pilot scale fed-batch production data and then validated with perfusion mode experimental data from Xu and Chen (2016) and from Zhang et al. (2015), which employed a temperature downshift within the process operation. Equations (1-7) show the perfusion model equations used:

\[
\begin{align*}
\frac{d(V_X)}{dt} &= (\mu - \mu_d) V X_V - F_{\text{bleed}} X_V \\
\frac{d(VP)}{dt} &= Q_P V X_V - (F_{\text{harvest}} + F_{\text{bleed}}) P \\
\frac{d(V[GLC])}{dt} &= -\left(\frac{\mu - \mu_d}{V_{X/V}} + m_{\text{glc}}\right) V X_V + F_{\text{in}} C_{\text{in}} + F_{\text{suppl}} C_{\text{suppl}} \\
&\quad - (F_{\text{harvest}} + F_{\text{bleed}})[GLC] \\
\frac{d(V[LAC])}{dt} &= Y_{\text{lac/gluc}} V X_V - (F_{\text{harvest}} + F_{\text{bleed}})[LAC] \\
\frac{d(V)}{dt} &= F_{\text{in}} + F_{\text{suppl}} - F_{\text{harvest}} + F_{\text{bleed}} \\
\mu &= \mu_{\text{max}} \frac{[\text{GLC}]}{K_{\text{glc}}[\text{GLC}]} \left(\frac{K_{\text{lac}}}{K_{\text{lac}} + [\text{LAC}]}ight) \\
\mu_d &= k_d \frac{[\text{LAC}]}{K_{D_{\text{lac}}}[\text{LAC}]} \left(\frac{K_{D_{\text{glc}}}}{K_{D_{\text{glc}}} + [\text{GLC}]}ight)
\end{align*}
\]

where, \(X_V\) is the viable cell density and \(V\) is the culture volume. \(F_{\text{bleed}}, F_{\text{harvest}}, F_{\text{in}},\) and \(F_{\text{suppl}}\) are the bleed, harvest, main, and supplementary feed flow rates, respectively. \(P, [\text{GLC}],\) and \([\text{LAC}]\) represent the antibody product, glucose and lactate concentrations, respectively. \(C_{\text{in}}\) and \(C_{\text{suppl}}\) are glucose concentrations in the main and supplementary feed streams, respectively. \(\mu\) and \(\mu_d\) are cell growth and death rates respectively, while \(\mu_{\text{max}}\) and \(k_d\) are their corresponding maximum values. \(Y\) is the yield coefficient, where \(Y_{x/y}\) for example represents the change in the value of \(x\) with respect to variations in values of \(y\). \(m_{\text{glc}}, K_{\text{lac}}, K_{\text{glc}}, K_{D_{\text{lac}}}, K_{D_{\text{glc}}}\) are the Monod model parameters representing the glucose maintenance coefficient, the lactose and glucose half maximum rate concentrations for cell growth and cell death, respectively.

The pilot facility implemented single-use equipment for its runs. Changeover time was therefore calculated to be 1 day and changeover costs represented the costs of new reactor bags used. Run durations were assumed to be 7, 14, and 60 days for batch, fed-batch, and perfusion operations, respectively. The operating cost was calculated as the cost of media (basal and feed) and utilities (water, gas, and electricity) in addition to changeover costs (cost of reactor bags).

2.2. Capture chromatography models

The lumped kinetic model was used for the chromatography column (Felinger and Guiochon, 2004; Guélat et al., 2016). Equations (8-9) show the model equations:

\[
\epsilon \frac{\partial c_i}{\partial t} + (1 - \epsilon) \frac{\partial q_i}{\partial t} + u_{sf} \frac{\partial c_i}{\partial x} - u_{sf} a_x \frac{\partial^2 c_i}{\partial x^2} = 0
\]
\[
\frac{\partial q_i}{\partial t} = k_m (q^e_i - q_i) \tag{9}
\]

where, \(c_i\) and \(q_i\) are the concentrations of component \(i\) in the mobile liquid phase and in the resin, respectively. \(\varepsilon\) is the total porosity. \(x\) is the distance travelled along the column length, and \(u_{sf}\) is the superficial velocity. \(d_{ax}\) is the axial dispersion coefficient. \(k_m\) is the lumped mass transfer coefficient, while \(q^e_i\) is the equilibrium concentration of component \(i\) in the resin.

Model parameters were fitted using data obtained from the pilot scale facility. Optimal superficial velocities and column lengths are determined based on input conditions from upstream operations in terms of total volume to be treated and product titer. In the pilot facility, approximately 200 loadings are assumed per column before it is disposed of. The required cost of resin is then accordingly calculated. Operating time for batch columns include: loading, washing, elution, and regeneration. Multiple column operations (PCC and SMB) take into account the parallel execution of the columns, but add the interconnected loading time.

3. Results and discussion

Figure 2 shows the model fit to two experimental perfusion runs. The figure shows the viable cell density, Glucose and lactate concentrations, in addition to the mAb product. In the experiment of Zhang et al. (2015) the total produced mAb was reported, while in the experiment of Xu and Chen (2016), the produced titer. Zhang et al. (2015) implemented a temperature shift in the experiment. Lowering temperatures in perfusion culture to inhibit cell growth has been shown to enhance productivity. The model developed and used here is independent of operating conditions. Therefore, three different phases are modelled for such an operation, the initial batch phase, the initial growth phase in perfusion mode, and the production phase after a temperature downshift. Model performance is reported independently for each phase. The model results match well to the experimental values, especially for viable cell density and produced mAb concentrations/amounts. Lactate concentrations are not well modelled in the perfusion growth phase though. Nevertheless, the fed-batch model (fitted to the pilot scale data) performed better for lactate prediction yielding an \(R^2\) value of 0.75.

Figure 3 shows the simulated results of a comparison of the different cultivation modes in terms of operating costs and production time in a 200 L reactor. Perfusion simulation assumed conditions closer to those presented by Xu and Chen (2016), which resulted in a higher titer compared to Zhang et al. (2015). No scale-up effects were assumed in this work. Perfusion mode shows a definite advantage in terms of productivity. However, due to the higher costs of media in perfusion, it can also lead to higher costs with longer operation times as compared to the fed-batch operation. Further analysis of this result can show the sensitivity of the produced results to operating conditions and model assumptions.

Figure 4 shows the results of the comparison of batch and a 3-column PCC operation of chromatography columns in terms of the cost of resin needed and the processing time for a small, pilot and large-scale operation at an inlet titer of 3 g/L. PCC columns were assumed to be one third of the length of the batch. The higher column utilization observed in PCC can be seen at pilot and larger scales as benefits in operating costs. Continuous column operation once again shows a major advantage in processing time.
Integrated design of mAb production processes: operation modes & configurations

Since column dimensions were not changed among scales, the processing time/g-mAb remains the same. The total processing time will then accordingly change with the scale; or the cost can change as more columns become needed at larger scales.

For the operation of an integrated continuous process, flowrates and productivities of upstream and downstream units need to be unified. Intermediate buffer tanks will be required to maintain flexibility, or in the case of implementing hybrid mode operations. Considerations of operational difficulties might also be encountered at larger scales for continuous operation. In cultivation mode, clogging might impede longer operation times, while in chromatography valve reliability will be a key issue for smooth operations.

Figure 3: comparison of operating modes in terms of (a) operating cost and (b) production time using a 200 L reactor volume.
4. Conclusions

This work presents models developed for the main cultivation unit. The models are fitted and validated using pilot scale fed-batch data and experimental scale perfusion data from literature. Models for chromatography units are also fitted using pilot scale data. The produced models are used to simulate and compare different operating modes in terms of operating costs and processing times. Continuous operation was shown to have a clear advantage in terms of productivity, but the advantage in terms of costs still needs to be fully investigated and verified. A sensitivity analysis to model assumptions is still required to map out and characterize the design space.

Acknowledgements

This work was supported by the Japan Agency for Medical Research and Development (AMED) [grant No. JP19ae0101064, JP17ae0101003, JP18ae0101058].

References