

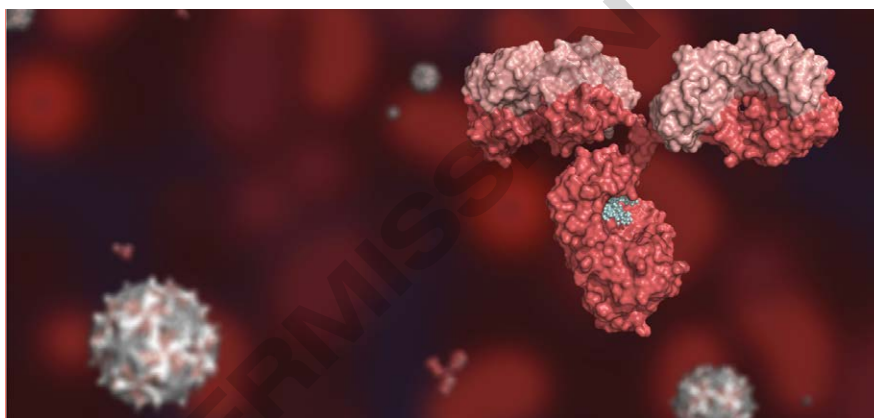
Modeling of Biotherapeutic Manufacturing Processes

Notes from a November 2023 Webinar

Josh Abbott with Cleo Kontoravdi and Dong-Yup Lee

The International Biomanufacturing Network (IBioNe) is a group of organizations that is dedicated to spreading knowledge about biomanufacturing. IBioNe is sponsored by the US National Science Foundation (NSF Award# 2114716: principal investigator Michael Betenbaugh and co-principal investigator Seongkyu Yoon) and serves other NSF-sponsored organizations, including the Advanced Mammalian Biomanufacturing Innovation Center (AMBIC) and the Membrane Science Engineering & Technology Center. According to the IBioNe website, the network seeks to serve as “a catalyst for technology innovation in biomanufacturing with increased biomanufacturing training and workforce development opportunities globally, accelerating discoveries and developments of lifesaving drugs and vaccines.” By sharing information and training industry professionals around the world, IBioNe seeks to help make life-saving treatments available and affordable.

In November 2023, Seongkyu Yoon (co-PI of IBioNe, and chair of workforce development committee) hosted a webinar featuring Cleo Kontoravdi (researcher from Imperial College in London, UK) and Dong-Yup Lee (associate professor of the School of Chemical Engineering at Sungkyunkwan University), who Betenbaugh introduced as experts in mathematical modeling of biotherapeutic production processes (2). Kontoravdi began by discussing modeling of cell-culture processes.



[HTTP://STOCK.ADOBE.COM](http://stock.adobe.com)

MODELING UPSTREAM PROCESSES

According to Kontoravdi, modeling is useful even during the first stages of process development. It can aid both the consolidation of multiomics data sets and their systematic analysis, which can lead to better experiment design.

Modeling often is used to control a cell-culture process and ensure production of a specific range of glycan structures. Protein glycosylation takes place in the endoplasmic reticulum and the Golgi apparatus of a mammalian cell. Such posttranslational modifications rely on the right concentrations of enzymes as well as the transport proteins that bring in substrates produced metabolically in cytosol. Protein synthesis depends both on the nutrients that scientists provide to a cell and on central carbon metabolism and the energy that process generates.

Glycosylation is influenced by metabolic function, enzyme regulation, and an intricate network of reactions occurring in cellular organelles. At the

same time, recombinant proteins compete for resources with host-cell proteins (HCPs), and the loss of nutrients can generate an array of glycans. To control an overall process, biomanufacturers need to account for the many variables that can affect glycosylation, which is difficult to manage exclusively through experimentation. That is where modeling comes in.

Modeling enables Kontoravdi's team to describe mathematically how process conditions affect metabolism and substrate availability. Process engineers can translate such information into predictions about what is happening inside a Golgi apparatus and how the rate of protein production might affect glycan maturity.

From there, Kontoravdi's team can assess whether a process generated an appropriate glycomic profile or requires further optimization. Such assessments can be done partially in silico using insights from the modeling. The team

can also account for heterogeneities at the population level and observe them inside a bioreactor. At pilot and production scales, for example, heterogeneity might affect the availability of substrates and dissolved oxygen and how such factors affect cell metabolism and product quality.

It's helpful to optimize cell-culture behavior using bioreactor-environment models that can be used at a small or large scale. Large-scale models enable scientists to develop computational fluid dynamics (CFD) packages that describe events happening throughout a large-scale bioreactor. Then, analysts can decide whether cells are behaving as a homogeneous population, which is called an *unsegregated biophase*. If the cell population is heterogeneous, it can be segregated into different subpopulations.

Biomanufacturers also can use models to decide if they should introduce structural-biology principles — e.g., by observing the organization and transport of different organelles. Analysts also can observe cells under a *black-box model*, focusing on only the materials flowing in and out of cells rather than their internal components. Multiple types of models can be used to fit different needs. For instance, analysts can create combinations of segregated or unsegregated models based on structured or unstructured data. The most complex models are used for heterogeneous cell populations and can describe intracellular components and organelles. Forming and validating such a model requires an abundance of data gathered under different conditions. That can be a demanding process.

Kontoravdi said that her team sometimes observes random effects in biological systems, meaning that deterministic models are not entirely appropriate. Instead, the team considers stochastic effects, which add to a model's complexity in providing a high-fidelity representation of a system.

Models can be dynamic and able to describe what a population does across an entire culture or production run. But they can also be static and describe a cell's behavior only at a particular time. Models can be designed to describe

single cells or entire populations. They can be black boxes as described above, or they can enable a developer to observe an entire culture environment. They can describe homogeneous or heterogeneous populations and consider the homo- or heterogeneity of an extracellular environment. If an extracellular environment is homogeneous, CFD could be incorporated to account for heterogeneities.

BUILDING A MODEL

Kontoravdi said that before developing a model, it is important to decide why you are building it and how you will use it. A segregated and structured model is appropriate for consolidating multiomics data sets and for analyzing cellular behavior.

But when performing on-line control in some bioprocesses, you might prefer a simple model that is unsegregated and unstructured. Consider the data that are available and whether they need augmentation to form a reliable model. Performing experiments without consideration for data curation may result in data sets of poor quality.

In *dynamic modeling*, biomanufacturers use system excitation that enables them to tease out the dynamics and individual contributions of various process inputs. Modelers must be included in experiment design. It's important to understand what data can be measured and how frequently that can happen.

Models can be either mechanistic or data-driven. *Mechanistic models* hypothesize a relationship among variables in a given data set, yielding differential equations that can be used to predict process outcomes. Such models can be expanded to include critical process parameters (CPPs) and critical quality attributes (CQAs). They typically require extensive experimentation and intracellular measurements to validate the results, and the availability of a time-course data set is essential for dynamic representation. Mechanistic models are useful for quantifying trade-offs and informing operating strategies. Unfortunately, they can be difficult to parameterize, are typically nonlinear, and are expensive to develop.

HYBRID MODELS

combine the best of both worlds and can augment the capabilities of kinetic models.

Data-driven models provide an alternative. Many complex machine-learning (ML) algorithms are available, but their utility relies upon the size of a given dataset. Such models enable developers to gain understanding of correlations, and although they allow for interpolation, they are not reliable for extrapolation. On the positive side, little biological expertise is needed to develop data-driven models quickly, and they are useful for integrating on-line measurements. For example, data-driven modeling is essential when transforming on-line Raman-spectroscopy data into feedback measurements for automated process control.

Mechanistic-Modeling

Approaches: Within the mechanistic modeling domain, biomanufacturers typically use either a kinetic or stoichiometric approach. *Kinetic models* are dynamic and often treat cells as black boxes. However, developers can incorporate known measurements. Such models are specific to the system at hand.

Stoichiometric models are generically applicable sets of equations that can be customized given the right data. In theory, they can integrate nearly limitless amounts of biological and process data and can be expanded beyond metabolic functions, which is usually the first point for which they are developed. Stoichiometric models can be used to create custom cell-line models by integrating information about a product protein and manufacturing process, usually in combination with kinetic or data-driven components.

Hybrid models are a third option that combine the best of both worlds and can augment the capabilities of kinetic models, which assume that the expression of enzymes and transfer proteins is constant. But that is not the

Studies have helped researchers **IDENTIFY** the most energetically **COSTLY** HCPs, which were knocked out to make a recombinant product and create a cleaner feedstock.

case throughout a manufacturing run. It is possible to incorporate ML components to use data-driven methodology that accounts for enzyme regulation inside a Golgi apparatus.

ADVANTAGES OF A HYBRID APPROACH

Kontoravdi's team conducted seven experiments with different feeding strategies using galactose and uridine. They tried to influence the glycosylation and galactosylation of a monoclonal antibody (mAb), and when analyzing their findings, they noticed a large discrepancy between the predicted and experimental results.

Kontoravdi explained that such discrepancies can occur when using a purely kinetic approach because the model will not account for unmeasured events. Although her team understood what was happening with their intracellular nucleotide sugars, they didn't know what was happening with their enzyme levels. But because they knew the outputs, the team was able to build a neural network to account for enzymatic activity, and they slotted the ML component into a kinetic model. That hybrid model performed better because it accounted for genetic-regulation events that otherwise could not be accounted for, and it reduced the mean absolute error by about 30% when compared with the initial kinetic model. That combined approach can make reliable models to solve intricate problems.

Biomanufacturers can use genome-scale models to engineer better cells. Studies have helped researchers identify the most energetically costly HCPs, which were knocked out to make a recombinant product and create a cleaner feedstock. They also can be used for process engineering.

During a successful cell-line experiment, Kontoravdi's group used a genome-scale model and historical data on amino-acid use to invent a strategy for optimizing cell-specific productivity. The team later refined the model and selected scenarios that can be implemented in a laboratory to make higher-producing clones. Most scenarios contained a reasonable number of genetic interventions for abundant amino-acid creation and suppression of biomass growth.

Kontoravdi explained a strategy that her team used to create a leucine abundance within a cell line by knocking out the gene for branched-chain amino-acid aminotransaminase 1 (*BCAT1*) and by overexpressing the *AACS* and *AACS2* genes. The team also lowered the bioreactor temperature to reduce cell growth. That genetically engineered cell line achieved bioreactor titers of >2.6 g/L, which was higher than yields achieved in their previous experiments. The line also maintained high specific productivity through to the end of the culture.

ESACT COURSE ON METABOLIC AND BIOPROCESS MODELING

Lee described a European Society for Animal Cell Technology (ESACT) course on metabolic and process modeling for animal cells. IBioNE first ran the course in October 2023 and plans to do so again in Fall 2024. Lee said that the course covered many considerations for model building — e.g., CPPs, process environment, and media composition — and discussed how those affect cellular behavior and productivity. The course also details uses of and differences between stoichiometric and kinetic models and how to combine their strengths into a hybrid model. Further course topics include CFD for cell bioreactors, chemometrics, and their use to support integration of process analytical technologies.

ONGOING COMMUNITY RESEARCH

Lee concluded by explaining ongoing efforts from his team and global research groups to build a genome-scale model, which he hopes will be available in 2024. He said that an initial goal is to model mammalian cell lines, such as

human embryonic kidney (HEK) and other animal cells, which could be used for adenoassociated virus (AAV) production.

To make the model viable, the researchers first must overcome difficulties establishing reliability. Although their prototype provides an understanding of what is happening inside a cell, Lee's team seeks to further improve cellular behavior and model predictability. They have improved quality by adding regulatory and kinetic data.

Of course, a model's utility is its most important characteristic. To increase that, Lee's team has worked to demonstrate how models can be used to develop basal media. They built a workflow that enabled them to identify a bottleneck during cell culture, then used that to adjust media components.

Technological advances open the possibility that such models could be applied for cell-line engineering and development. Techniques such as genome editing are useful tools, but researchers still need to identify the right genetic targets. Similarly, Lee's team can use their currently available model and multiomics profiling data to identify engineering targets. Furthermore, modern virtual tools can be incorporated to predict and simulate actual systems and their behaviors.

REFERENCES

- 1 *IBioNE: International Biomanufacturing Network*. 2024; <https://www.uml.edu/research/ibione>.
- 2 *IBioNE: News and Events*. 2023; <https://www.uml.edu/research/ibione/news-events.aspx>.

Josh Abbott is associate editor at BioProcess International; josh.abbott@informa.com. *Cleo Kontoravdi* is a researcher at Imperial College London in the UK, and *Dong-Yup Lee* is associate professor at the School of Chemical Engineering at Sungkyunkwan University in Seoul, South Korea.

To share this in PDF or professionally printed form, contact Lisa Payne: lpayne@mossbergco.com, reprints@mossbergco.com; 1-219-561-2036.