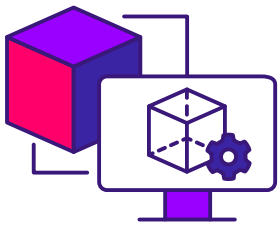
Basetwo

The first end-to-end platform to operationalize digital twins.

Use Case Ebook

Background

Basetwo is a venture-backed software company focused on empowering engineers to **operationalize digital twins**



Johnson & Johnson INNOVATION | JLABS

What sets Basetwo apart from existing Digital Twin solutions?

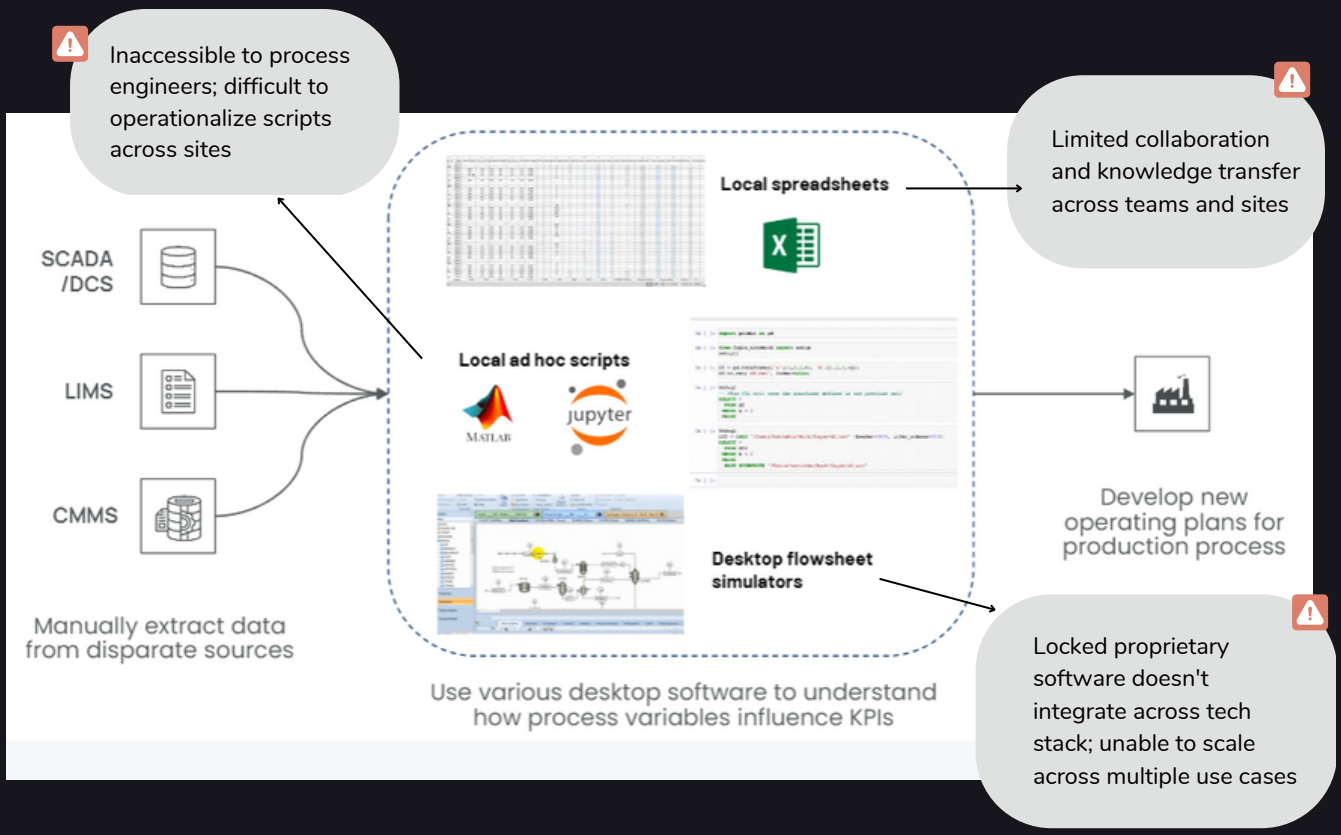
Basetwo is the first end-to-end **TwinOps** platform for the pharmaceutical industry.

TwinOps is focused on the lifecycle of taking digital twins from design to production, and then providing the infrastructure to maintain and monitor them once operationalized. It is inherently collaborative, often comprising process modellers, data scientists, site engineers, and IT.

- Integrated
- Collaborative
- Modular
- Scalable

Introduction

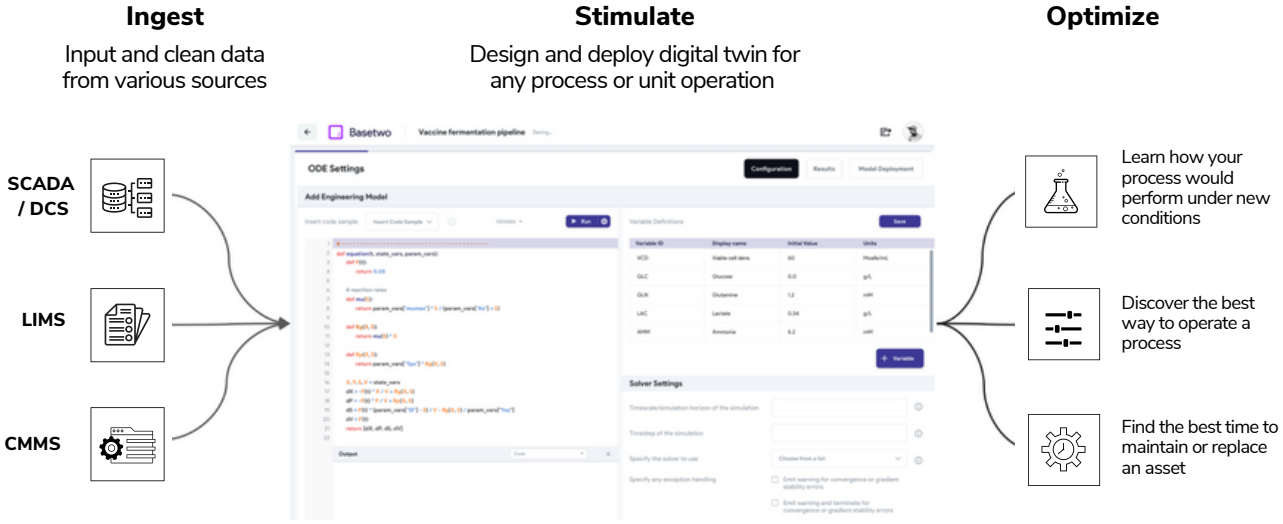
Today, engineers are over-reliant on **disconnected** workflows to improve production performance. Disconnected digital workflows **limit production efficiency** and scalability while introducing business risk. On the other hand, productionizing digital twins in an industrial, regulated environment is difficult. This ebook will introduce readers to Basetwo's unique digital twin solutions alongside their engagement models. In the latter part, we will dive into 5 use cases of how Basetwo supports the pharmaceutical and biotech industry in various ways.



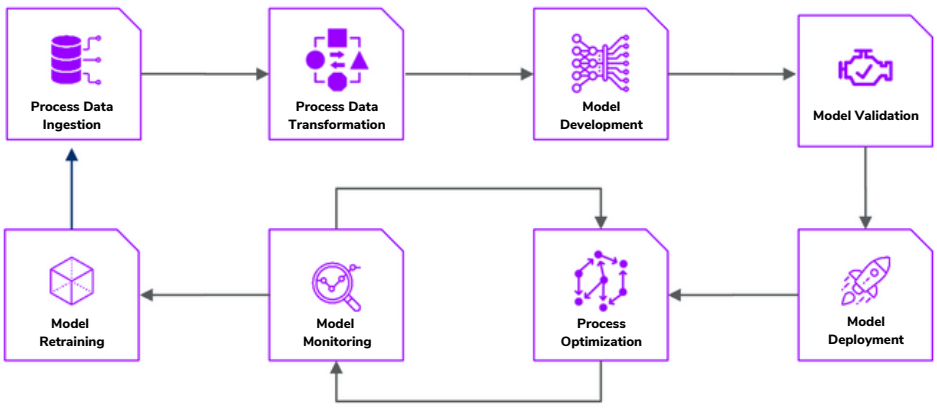
What is TwinOps

TwinOps is a software engineering practice focused on the **lifecycle** of taking digital twins from design to production, and then providing the infrastructure to maintain and monitor them once operationalized.

With Basetwo, engineers can reliably build, validate, and deploy digital twins of their processes



TwinOps is inherently collaborative, often comprising process modellers, data scientists, site engineers, and IT.



USE CASE 1

Optimization of monoclonal antibody titers in GS-NS0 bioreactors

01 The Challenge

Optimizing antibody titers in mammalian cGMP bioreactors

Importance of biological yield

Biological yield is a key indicator of upstream fermentation process performance that directly contributes to the productivity of each manufacturing run as it determines the amount of product available for downstream processing.

Complications when trying to improve biological yield

Biological yield itself is a complex function of various critical process parameters (CPPs), including measured variables, such as the concentration of dissolved gases, substrates, temperature, pH, and unmeasured variables, such as bioreactor fluid dynamics.

The interaction between these CPPs is driven by convoluted underlying cellular and process dynamics, wherein several variables may interact in a synergistic or non-linear manner that is difficult to characterize through conventional process analytical techniques or simple mathematical models built in Excel or MATLAB.

The challenge manufacturers face

Process analysts and engineers who deeply understand underlying process dynamics are a valuable but expensive and limited resource who are often unable to provide feedback for all batches in real-time, around the clock. Further, these experts may not have access to the tools that allow them to diagnose and tackle manufacturing issues due to the noisy nature of the data and lack of valuable on-line measurements for key variables.

Therefore, model-based optimization of biological yield is currently either very complex or costly and therefore not worthwhile to pursue despite the potential gains in biological yield available.

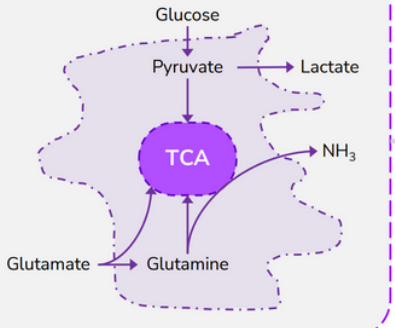
2 The Basetwo Solution

Leveraging an advanced cellular model to dynamically optimize titer

Development of a cellular model

Inputs into the model included results from nutrient and metabolite analysis of GS-NS0 cells and supplemented by quantified cell counts and antibody concentration. A modified version of the Arrhenius equation (1) was created to account for the lag phase using the principles of Monod kinetics (2). These equations were leveraged to create an early cellular model wherein mAb concentration can be isolated and maximized **(Figure 1)**.

Figure 1: Cellular model depiction



Adapted from Kiparissides et al. (2015)

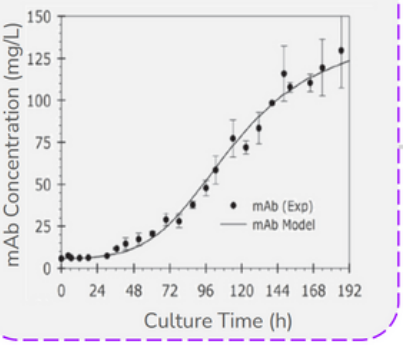
1) $$u_{MAX,\phi} = u_{MAX}\, exp\left(-\frac{K_{LAG}\cdot\delta_{NORM}}{X_{T,NORM}}\right)$$

2) $$u = u_{MAX}\frac{[S]^n}{K_S^n + [S]^n}$$

Model calibration and validation

Outputs of the early cellular model were compared to data from several fed-batch bioreactors to refine the weight of variables within the model through dynamic optimization. This formulation aims to maximize mAb concentration as a function of the established process model above combined with constraints for feed and total media volume and various initial conditions for cell density and media composition to reflect a variety of industrial conditions. The final validated model required **half** of the experimental variables to describe antibody production near perfectly (**Figure 2**). Key variables included coefficients for glucose, lactate, and glutamate metabolism and corresponding concentrations of these metabolites.

Figure 2: Model results versus experiments



Adapted from Kiparissides et al. (2015)

Model-based optimization of antibody titer

Maximization of antibody titer was achieved using model-based predictions to determine an optimal feeding schedule at various initial concentrations of antibody common to a fed-batch bioreactor. Relative to conventional optimization, model-based optimization of the feeding schedule achieved a **30% increase** in final mAb concentration.

03 Business Impact

Increased antibody titer and reduced process variability

The advanced cellular model presented in this use case is able to accurately predict and improve mAb production using a smaller set of commonly measured variables relative to traditional feed optimization approaches, reducing process variability and improving productivity

Advanced bioreactor models available through Basetwo are able to accurately predict bioreactor process variables and lead to an improvement in mAb titer of **30%**

Primary Reference: Kiparissides et al. 2015. On the model-based optimization of secreting mammalian cell (GS-NS0) cultures. Biotechnology & Bioengineering, 112(3):536-48.

USE CASE 2

Model-based optimization of Protein A resin service life

01 The Challenge

Optimizing antibody titers in mammalian cGMP bioreactors

Importance of resin lifespans

Protein A resins are crucial to mAb manufacturing and are nearly 50% more expensive than non-protein-based stationary phases. Manufacturing-scale columns are in the order of one million dollars per packing, and resin aging results in performance loss and limits the lifetime around 50-100 cycles, incurring significant manufacturing costs.

The complexities of maximizing resin lifespan

Optimization of Protein A resin lifespan is typically approached in a conservative manner by adding a safety margin to the number of cycles a column can process. Fixing the number of cycles at manufacturing scale can bear considerable risks, due to lot-to-lot variability, differences in feed material, column storage, packing quality, or other unforeseen events that can influence resin lifespan, often pushing the column lifespan below the safety buffer established by manufacturers.

The challenge manufacturers face

Resin aging is a complex phenomenon and is difficult to consistent and reliably predict without forsaking extended use of the column due to manufacturing risk associated with poor resin performance. Process engineers may not have access to the tools or models that allow them to diagnose and tackle resin lifespan issues due to the noisy nature of the data and lack of on-line or frequent measurements for key variables, such as dynamic binding capacity.

Therefore, maximization of Protein A resin lifespan is currently very complex and difficult to pursue despite the potential gains in process performance available.

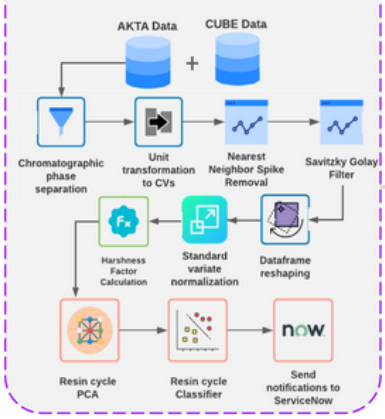
2 The Basetwo Solution

Leveraging a process model to track and simulate resin aging

Generation and processing of experimental data

Resin lifetime cycling studies were performed on Contichrom CUBE Combined and ÄKTA Avant 150 systems consisting of equilibration, loading, three wash steps, elution, and two cleaning in place and neutralization cycles. Elution fractions and breakthrough curves were measured for the desired mAb alongside various impurities. Experimental data were cleaned and processed according to the workflow in **Figure 1** to allow for input of data into chromatographic models.

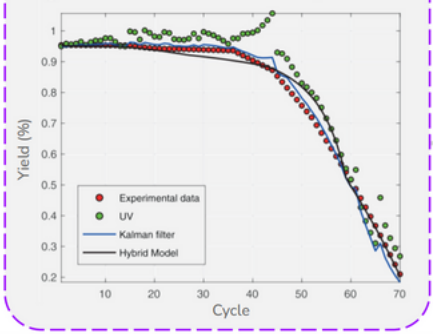
Figure 1: Data processing workflow in Basetwo



Development of hybrid and on-line model

Using the post-processed data, a hybrid model (1) is developed using a lumped kinetic model with column aging parameters (2) and calibrated according to experimental breakthrough curves. A separate model was also developed using the Kalman filter approach which translates on-line UV signals into mAb concentrations that are normally only available offline. Both models were fit to several column lifetime experimental datasets under various commercial conditions that affected aging behaviors.

Figure 2: Model results versus experiments



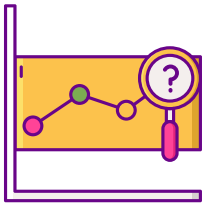
1) $$\frac{dc}{dt} = -v\frac{dc}{dx} + D_L\frac{d^2c}{dx} - \varphi\frac{dq}{dt}$$

2) $$q_{sat} = q^0_{sat}\ \theta_1$$
$$k_{max} = k^0_{max}\ \theta_2$$

Model performance and predictive ability

Maximization of antibody titer was achieved using model-based predictions to determine an optimal feeding schedule at various initial concentrations of antibody common to a fed-batch bioreactor. Relative to conventional optimization, model-based optimization of the feeding schedule achieved a **30% increase** in final mAb concentration.

03 Business Impact



Predictive resin maintenance and better resin ROI

The advanced hybrid models presented in this use case are able to accurately predict resin performance over their entire lifespan and even enable on-line maintenance using UV measurements. These models are also able to forecast resin performance, allowing for predictable maintenance scheduling and better overall return on investment for each resin used.

Advanced hybrid models available through Basetwo are able to accurately **predict** and **forecast** resin performance over their **entire lifespan** using **on-line** variables to vastly simplify maintenance

Primary Reference: Feidl et al. 2020. Model based strategies towards protein A resin lifetime optimization and supervision. J Chromatography A, 1625, p.461261.

USE CASE 3

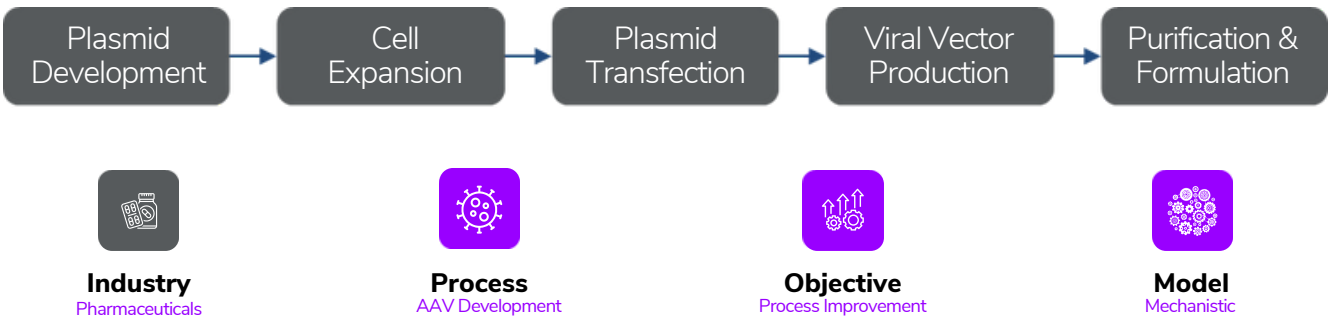
AI in Cell & Gene Therapy

Optimization of recombinant adeno-associated virus (AAV) development in HEK293 cells

01 Background

Currently, providing recombinant adeno-associated virus (AAV) supply for clinical trials and commercial demand remains challenging due to a combination of high required dosing and relatively low process efficiency. One of the most widely used protocols to produce AAVs is transient transfection of HEK293 cells. Improvement of AAV yields can significantly improve clinical supply, reducing a key bottleneck when conducting clinical trials leveraging AAV therapeutics.

Overview of Adeno-associated virus (AAV) manufacturing workflow



The Challenge 2

AAV yield is a complex function of various process parameters that are poorly understood despite decades of research. **Only 5% to 30% of AAV capsids produced by cells contain the desired therapeutic element, creating a significant downstream purification burden.**

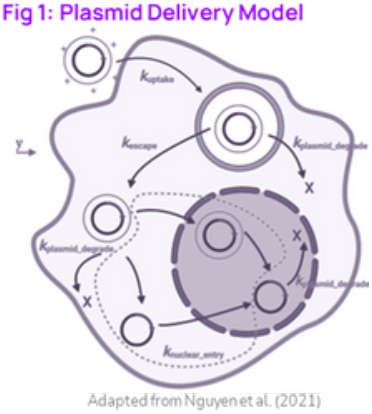
The interaction between AAV manufacturing parameters is driven by convoluted underlying cellular and capsid dynamics, wherein several variables may interact in a non-linear manner difficult to characterize through conventional process analytical techniques or simple mathematical models built in Excel or MATLAB. This makes it difficult to improve production performance or to alleviate process bottlenecks.

03 Digital Twin Development

A **mechanistic model** was leveraged to virtually simulate the AAV manufacturing processes and identify new process conditions that can overcome current manufacturing bottlenecks.

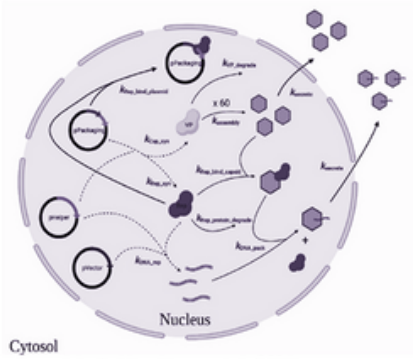
A mathematical **mass balance process model** was developed to describe the kinetic behaviour of transfection and other AAV development steps. The model contains two parts that capture the trafficking of plasmid to the nucleus **(Figure 1)** and replication **(Figure 2)**.

Fig 1: Plasmid Delivery Model



Adapted from Nguyen et al. (2021)

Fig 2: Viral Vector Synthesis Model



Adapted from Nguyen et al. (2021)

In total, 14 kinetic parameters were leveraged in the model which account for gene delivery, viral production, protein synthesis, protein degradation, protein binding, packaging, and secretion (Table 1). The definitions of these parameters are described in the appendix.

Table 1: Kinetic model parameters estimated from experimental data & literature

k				
uptake	*escape*	*nuclear_entry*	*plasmid_degrade*	*Cap_syn*
Rep_syn	*VP_degrade*	*Rep_protein_degrade*	*DNA_rep*	*assembly*
secrete	*DNA_pack*	*rep_bind_plasmid*	*rep_bind_capsid*	-

Once kinetic parameters were established, a family of 18 ordinary differential equations (ODE) describing the mass balance of key process processes were developed. An example ODE describing the dynamics of the endosomal vesicle complex is below:

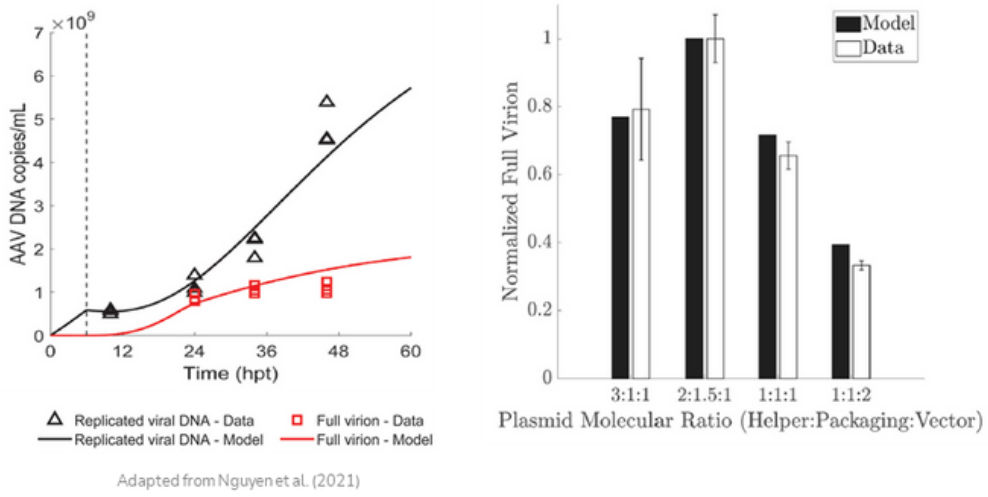
$$\frac{d(EndosComplex)}{dt} = k_{uptake}(ExComplex) - \left(k_{escape} + k_{plasmid\_degrade} + \mu\right)(EndosComplex)[0,1]$$

04 Digital Twin Results

The kinetic parameters were determined using process data collected from historical transfections. When data to inform any parameters were limited or experimentally infeasible, estimates were obtained from literature for related processes and validated with small-scale experiments using maximum likelihood estimation when possible. If no information was available for a given parameter, structural analysis and prior parameter knowledge was utilized to make parameter assumptions.

ODEs were solved by a stiff ODE solver. Briefly, a multi-start approach with 1,000 initial conditions and an interior-point algorithm was used to identify global minima by minimizing the objective function, defined as the error between experimental and simulated results.

Fig 3: Digital Twin vs. Experimental Results



Adapted from Nguyen et al. (2021)

Mechanistic model simulations showed good agreement with experimental data **(Figure 3).** Simulation results show total plasmid uptake peaked 20h post-transfection and plasmid concentration in the nucleus peaked at 45h. Of note, only 5% of the initial plasmid added to the cell culture is transfected into the cells, and only 0.6% of the total plasmid input is eventually trafficked into the nucleus, revealing important process bottlenecks.

05 Business Impact

The digital twin revealed AAV production was significantly influenced by:

1.Plasmid trafficking

2.Rep proteins that repress gene expression

3.DNA replication during first 24 hours post-transfection

This allows researchers to focus efforts to improve production performance and accelerate process development of AAVs crucial for clinical supply for gene therapy manufacturers.

The biological mechanism of AAV synthesis captured by the digital twin is largely unchanged for other transfection methods, including across cell lines, meaning **the digital twin can easily be adapted to other production methods**, improving AAV development across the organization.

Overall, leveraging the digital twin unlocks digital experimentation and hypothesis testing to help stakeholders evaluate and reveal process bottlenecks that can be investigated and addressed to improve process development and improve final AAV yield and quality.

Advanced process models available through Basetwo accurately predict process dynamics in AAV development, facilitating identification of process bottlenecks, accelerating process development, and advance gene-therapy manufacturing.

Primary Reference: Nguyen et al. Mechanistic model for production of recombinant adeno-associated virus via triple transfection of HEK293 cells. Mol Ther Methods Clin Dev. 2021.

Digital Twins in API Development

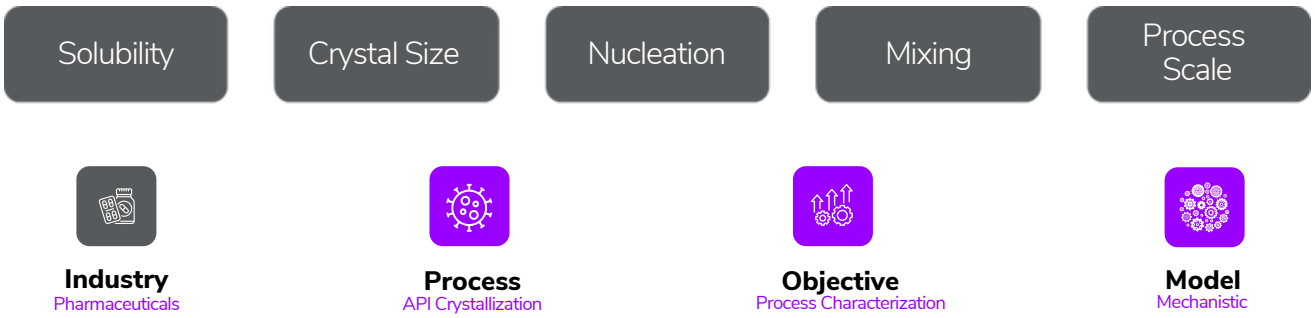
Characterization and prediction of API crystallization scale-up leveraging a population balance model

USE CASE 4

01 Background

Crystallization represents the most widely used method to separate API from its mixture and is the major operation of API manufacturing where the control strategies of API synthesis converge. Successful crystallization needs to meet a variety of purity criteria and ensure appropriate physicochemical critical quality attributes (e.g. crystal form, particle size distribution) are optimal as to not affect downstream processing and the final composition of drug formulations.

Overview of factors influencing crystallization performance



The Challenge 2

Particle size is considered an important powder property that influences downstream performance. However, particle size and process parameters such as crystallizer size, baffle type, seed load, temperature profile, antisolvent addition rates, and agitation rate, can all impact the final product particle size.

The interaction between particle size is driven by underlying thermodynamics, wherein several variables may interact in a non-linear manner difficult to characterize through conventional process analytical techniques or simple mathematical models built in Excel or MATLAB. This makes it difficult to predict or understand particle size when scaling up or down or under novel process conditions

03 Digital Twin Development

A **population balance model** was created with the goal of describing the kinetic behaviour the crystallization processes, including supersaturation, nucleation, and other crystallization steps. The structure of the kinetic model was informed by experiments conducted on **lab scale (10 g)** crystallization runs designed to evaluate the influence of various process parameters on particle size distribution, crystal growth, and nucleation during crystallization.

The general population balance equation adapted from literature takes into account inlet and outlet streams and can be simplified to a mass balance model for batch crystallization:

$$\frac{dC}{dt} = -3k_v\rho_c G\int_0^\infty nL^2\,dL - \frac{C}{V}Q$$

1. Lab-scale experiments
Collect data on process parameters and particle size

2. Create process model
Kinetic model leveraging process & reactor details

3. Estimate kinetic parameters
Fit kinetic equations to match lab data

4. Simulate in novel conditions
Predict particle size, growth rate, nucleation rates

Where C is the API concentration in solution, **k(v)** is the volumetric shape factor, **ρ(c)** is the true density of the crystals, and Q is the flow rate of the antisolvent. Values were informed by simulated single-crystal structure and literature when they could not be estimated from experiments and simulations.

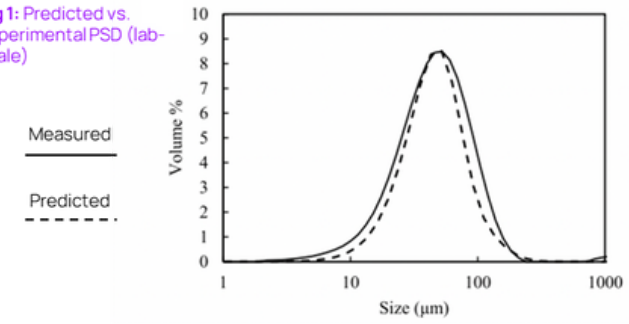
Solubility, particle size distribution, and crystallization parameters were estimated from lab-scale experiments and simulated experiments at **plant scale (35 kg)** using a variety of reactor capacities, impeller types, diameters, and speeds.

04 Digital Twin Results

Solubility data for the API were collected from HPLC and compared to modeled estimates as a function of temperature and antisolvent. Predicted solubility showed good agreement with measured data. After solubility was validated, crystallization was then modelled using secondary nucleation kinetic models considering particle-impeller collisions since crystallization runs were seeded.

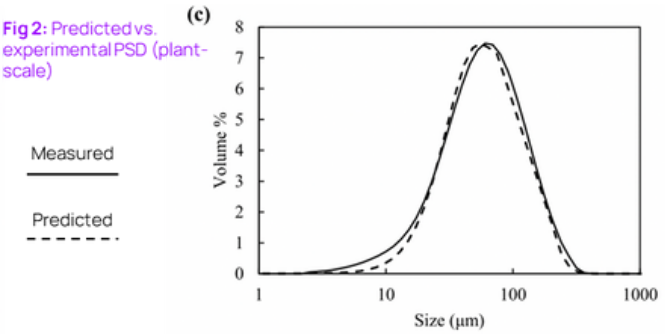
Overall, model predictions were in good alignment with both lab-scale **(Figure 1)** and plant-scale **(Figure 2)** batches. Several process parameters were observed to have an influence on particle size distribution, including supersaturation, impeller pumping number, impeller power number, and energy dissipation rate and were well captured in the final model for almost all experiments performed.

Fig 1: Predicted vs. experimental PSD (lab-scale)



Impeller	Impeller diameter	Speed	Power number	Pumping number	Pump to power
Retreat curve	36 in	75 rpm	0.27	0.37	1.37

Fig 2: Predicted vs. experimental PSD (plant-scale)



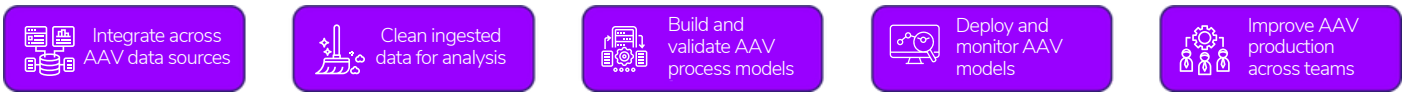
Impeller	Impeller diameter	Speed	Power number	Pumping number	Pump to power
Three blade pitch turbine	23.6 in	60 rpm	0.88	0.72	0.82

05 Business Impact

The digital twin was able to predict particle size from crystallization processes conducted under different conditions, different scales, and in different reactors with different impeller types. Process modelling indicated that secondary nucleation strongly influences the particle size in the API crystallization process. **Agitator type and speed were found to have a strong influence on the resulting particle size due to their influence on secondary nucleation**, while parameters that influence growth, such as seed load, were observed to have more minor effect on particle size.

Overall, leveraging the digital twin unlocks digital experimentation and hypothesis testing to help stakeholders evaluate and reveal process bottlenecks that can be investigated and addressed to improve process development and improve final API yield and quality. This model helps to provide a **more robust particle size control strategy** by employing fundamental crystallization kinetics to predict particle size.

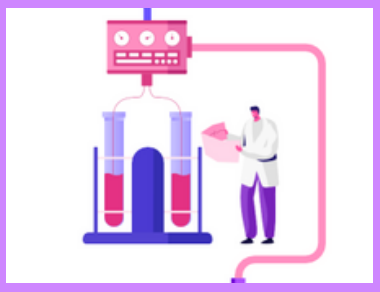
Advanced process models available through Basetwo are able to accurately predict process dynamics in API development, facilitating identification of process bottlenecks, accelerating process development, and advancing API manufacturing.



Primary Reference: Rosenbaum et al. Population Balance Modeling To Predict Particle Size Distribution upon Scale-Up of a Combined Antisolvent and Cooling Crystallization of an Active Pharmaceutical Ingredient. Org. Process Res. Dev. 2019, 23, 2666−2677.

USE CASE 5

Optimization of bioreactor titers through virtual “what-if” analysis



01 The Challenge

Optimizing cell culture productivity in bioreactors

Importance of cell culture productivity

Cell culture productivity often refers to product titer as a function of viable cell count and is a complex function of several key process parameters such as pH, dissolved gasses, and metabolite concentration, making productivity difficult to measure in isolation. Lactate concentration in particular directly contributes to the productivity of each manufacturing run as it impedes protein manufacturing at higher levels, reducing the amount of product available for downstream processing.

The complexities of maximizing cell culture productivity

The interaction between the critical process parameters that determine the productivity of a given batch is driven by convoluted underlying cellular and process dynamics, wherein several variables may interact in a synergistic or non-linear manner that is difficult to characterize through conventional analytical approaches or simple mathematical models built in Excel or MATLAB. Introducing metabolic shifts to optimize productivity (i.e., controlling lactate production or inducing lactate consumption) has been achieved through control strategies such as dynamic feeding to control glucose levels.

The challenge manufacturers face

Process analysts and engineers who deeply understand underlying process dynamics are a valuable but expensive and limited resource who are often unable to provide feedback for all batches in real-time, around the clock. Further, these experts may not have access to the tools that allow them to diagnose and tackle manufacturing issues due to the noisy nature of the data and lack of valuable on-line measurements for key variables.

Therefore, model-based optimization of cell culture productivity is currently either very complex or costly and therefore not worthwhile to pursue despite the potential gains available.

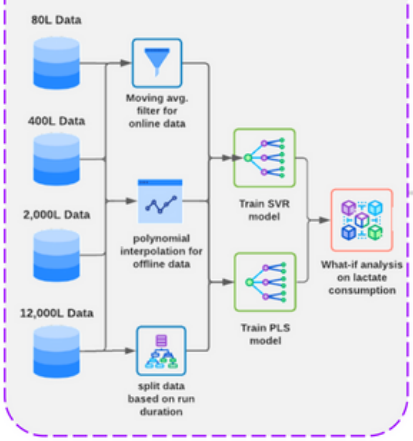
2 The Basetwo Solution

Using a model to reveal cellular patterns and virtually experiment

Generation and processing of experimental data

Data from 243 production-scale bioreactor runs producing a recombinant IgG molecule using a single CHO cell line were used for analyses. Three seed batch processes were included (80L, 400L, and 2000L), as well as a 12,000 L fed-batch bioreactor. Experimental data containing approximately 30 parameters at both inoculum and production scales were cleaned and processed according to the workflow in **Figure 1**, reproducible on the Basetwo platform.

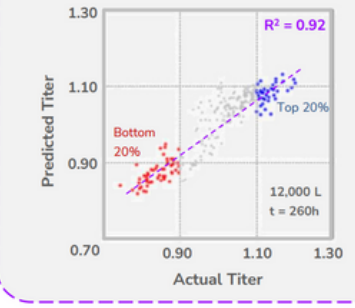
Figure 1: Data processing workflow in Basetwo



Development of vector regression model

Using the post-processed data and available parameters, a support vector regression (SVR) algorithm aimed at predicting final titer and lactate production at commercial scale is developed and cross-validated using data from all scales. Model performance was assessed according to correlation coefficients and root mean square error with actual titer. The SVR model was able to accurately predict antibody titer at commercial scale over a 260-hour fermentation period **(Figure 2)**.

Figure 2: Model titer versus actual titer



Adapted from Le et al. (2012)

Application of the model for "what-if" analyses

Leveraging the validated SVR model, several manufacturing insights were obtained using "what-if" analyses that varied virtual experimental conditions which were later validated experimentally. For example, the model revealed that final antibody titer can be accurately predicted from the first 70h of fermentation time. Relatedly, cells appeared to experience a persistence metabolic shift dependent on early glucose levels: batches with low glucose consumption had reduced lactate production and higher final product titers. As a result, corrective measures to suppress glucose consumption early on in fermentation were shown to increase final product titer.

03 Business Impact

Greater process understanding and better bioreactor ROI

The models presented in this use case are able to accurately predict bioreactor titer at multiple scales and allow for virtual experimentation to enable resource-free movement within the design space. This not only enables greater process performance, but also deeper process understanding by revealing process trends through accurately simulated “what-if” analyses.

Advanced models available through Basetwo are able to accurately **predict** bioreactor titer and enable **resource-free virtual experimentation** to **maximize process performance** and **understanding**

With an accessible web-based interface, Basetwo allows users to:

Integrate with multiple data sources and in-house spreadsheets

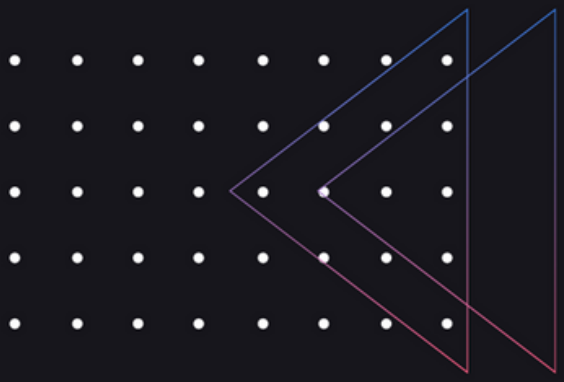
Clean ingested data for analysis

Collaborate with performance reports across departments

Easily **build** and **validate** advanced ML and process models

Deploy and **monitor** models to enhance manufacturing SOPs

Primary Reference: Le et al. 2012. Multivariate analysis of cell culture bioprocess data—lactate consumption as process indicator. Journal of biotechnology, 162(2-3), pp.210-223.



Interested in learning more?

Reach out!

basetwo.ai contact@basetwo.ai basetwoai basetwoai

Appendix

Appendix 1 - Full List of Model Parameters

Table 1. Model parameters

Parameter	Process	Value (95% confidence interval)	Unit	Evaluating method	Literature range
k_{uptake}	DNA uptake from medium into cell cytosol	2.78×10^{-2} (2.06×10^{-2}, 3.75×10^{-2})	h^{-1}	fit to literature measurement of total plasmid[20]	exponential decay of 9×10^{4}[19]
		1.19×10^{-3} (1.00×10^{-3}, 1.40×10^{-3})		fit to in-house measurement of vDNA	6.527×10^{-3} to 4.2×10^{-1} [18,25]
k_{escape}	endosomal escape	6.00×10^{-1}	h^{-1}	parameter for PEI 25K[18]	–
$k_{nuclear\_entry}$	nuclear entry	4.30×10^{-3} (2.4×10^{-3}, 7.7×10^{-3})	h^{-1}	fit to literature measurement of nuclear plasmid[20]	8.75×10^{-2} to 3.54[18,25]
$k_{plasmid\_degrade}$	plasmid degradation	1.95×10^{-2} (1.09×10^{-2}, 3.48×10^{-2})	h^{-1}	fit to literature measurement of total plasmid[20]	0.4621 to 0.8318[26]
k_{Cap_syn}	capsid protein synthesis	6.50×10^{4} (5.28×10^{4}, 8.00×10^{4})	molecule/cell/h	fit to in-house measurement of vDNA, total capsid, and full virion	–
k_{Rep_syn}	Rep protein synthesis				
$k_{VP\_degrade}$	VP degradation	2.77×10^{-1}	h^{-1}	literature	0.23 to 0.35[27]
$k_{Rep\_protein\_degrade}$	Rep protein degradation	2.45×10^{-2}	h^{-1}	literature	0.01 to 6[28–30]
k_{DNA_rep}	transgene rescue and replication	3.10×10^{-7} (1.52×10^{-7}, 6.32×10^{-7})	$cell^2/molecule^2/h$	fit to in-house measurement of vDNA	–
$k_{assembly}$	capsid assembly	1.00×10^{5}	h^{-1}	fast step assumption[31]	–
$k_{secrete}$	capsid secretion from the nucleus to the cytosol	1.00×10^{5}	h^{-1}	fast step assumption	–
k_{DNA_pack}	viral DNA encapsidation	8.18×10^{-2} (1.1×10^{-9}, 6.08×10^{6})	molecule/cell/h	fit to in-house measurement of full virion	–
$k_{Rep\_bind\_plasmid}$	binding of Rep protein on packaging plasmid	3.38×10^{-5}	molecule/cell/h	literature	1.0×10^{4} $M^{-1}s^{-1}$[32,33]
$k_{Rep\_bind\_capsid}$	binding of Rep protein on empty capsid	5.70×10^{-3} (4.10×10^{-3}, 8.00×10^{-3})	molecule/cell/h	fit to in-house measurement of total capsid and full virion	–

Nguyen et al. (2021)

Appendix 2 - Additional Crystallization Information

Table 2. Summary of Reactor Specifications Utilized for Crystallizations

batch	reactor capacity (volume)	impeller type	impeller diameter (in)
lab	250 mL	half-moon blade	1.89
plant 1–4	2140 L	retreat curve impeller	36.00
plant 5–8	685 L	three-bladed pitch blade turbine	23.62

batch	speed (rpm)	N_P (power number)	N_Q (pumping number)
lab	260	4.5	1.4
plant 1–4	75 for batches 1–3, 105 for batch 4	0.27	0.37
plant 5–8	60	0.88	0.72

Table 6. Summary of Kinetic Parameters Used To Fit Lab Data[a]

parameter, symbol, and equation number	value and units	95% confidence interval	standard deviation
effective diffusivity correction factor, α_{AB}, eq 7	1	N/A	N/A
activation energy, E_a, eq 9	23 426 J/mol	50 570	25 270
growth rate constant, k_g, eq 9	1.17×10^{-5} m/s	2.18×10^{-4}	1.09×10^{-4}
supersaturation order, s, eq 9	3.82	1.71	0.85
natural log of rate constant for crystal–impeller collisions, k_a, eq 9	19.35	4.07	2.04
size above which crystals undergo attrition, L_{min}, eq 10	132 μm	161	80.5
secondary nucleation order with respect to supersaturation, n_s, eq 10	3	N/A	N/A

[a]Confidence interval is for a two-sided interval centered around the estimated parameter value.

Primary Reference: Rosenbaum et al. Population Balance Modeling To Predict Particle Size Distribution upon Scale-Up of a Combined Antisolvent and Cooling Crystallization of an Active Pharmaceutical Ingredient. Org. Process Res. Dev. 2019, 23, 2666–2677.

Reference

1. Primary Reference: Kiparissides et al. 2015. On the model-based optimization of secreting mammalian cell (GS-NS0) cultures. Biotechnology & Bioengineering, 112(3):536-48.

2. Primary Reference: Feidl et al. 2020. Model based strategies towards protein A resin lifetime optimization and supervision. J Chromatography A, 1625, p.461261.

3. Primary Reference: Nguyen et al. Mechanistic model for production of recombinant adeno-associated virus via triple transfection of HEK293 cells. Mol Ther Methods Clin Dev. 2021.

4. Primary Reference: Rosenbaum et al. Population Balance Modeling To Predict Particle Size Distribution upon Scale-Up of a Combined Antisolvent and Cooling Crystallization of an Active Pharmaceutical Ingredient. Org. Process Res. Dev. 2019, 23, 2666–2677.

5. Primary Reference: Le et al. 2012. Multivariate analysis of cell culture bioprocess data—lactate consumption as process indicator. Journal of biotechnology, 162(2-3), pp.210-223.