processes

MDPI

Article

Towards Autonomous Process Control—Digital Twin for HIV-Gag VLP Production in HEK293 Cells Using a Dynamic Metabolic Model

Heribert Helgers [1], Alina Hengelbrock [1], Jamila Franca Rosengarten [2,3], Jörn Stitz [2], Axel Schmidt [1] and Jochen Strube [1,*]

[1] Institute for Separation and Process Technology, Clausthal University of Technology, Leibnizstr. 15, 38678 Clausthal-Zellerfeld, Germany
[2] Faculty of Applied Natural Sciences, Technische Hochschule Köln, 51368 Leverkusen, Germany
[3] Institute of Technical Chemistry, Leibniz University Hannover, Callinstraße 5, 30167 Hannover, Germany
* Correspondence: strube@itv.tu-clausthal.de

Abstract: Despite intensive research over the last three decades, it has not yet been possible to bring an effective vaccine against human immunodeficiency virus (HIV) and the resulting acquired immunodeficiency syndrome (AIDS) to market. Virus-like particles (VLP) are a promising approach for efficient and effective vaccination and could play an important role in the fight against HIV. For example, HEK293 (human embryo kidney) cells can be used to produce virus-like particles. In this context, given the quality-by-design (QbD) concept for manufacturing, a digital twin is of great importance for the production of HIV-Gag-formed VLPs. In this work, a dynamic metabolic model for the production of HIV-Gag VLPs was developed and validated. The model can represent the VLP production as well as the consumption or formation of all important substrates and metabolites. Thus, in combination with already described process analytical technology (PAT) methods, the final step towards the implementation of a digital twin for process development and design, as well as process automation, was completed.

Keywords: digital twin (DT); advanced process control (APC); quality-by-design (QbD); process analytical technology (PAT); real-time-release testing (RTRT); human embryonic kidney 293 cells (HEK293); human immunodeficiency virus (HIV); virus-like particles (VLPs)

Citation: Helgers, H.; Hengelbrock, A.; Rosengarten, J.F.; Stitz, J.; Schmidt, A.; Strube, J. Towards Autonomous Process Control—Digital Twin for HIV-Gag VLP Production in HEK293 Cells Using a Dynamic Metabolic Model. *Processes* **2022**, *10*, 2015. https://doi.org/10.3390/pr10102015

Academic Editor: Tao Sun

Received: 20 August 2022
Accepted: 28 September 2022
Published: 5 October 2022

1. Introduction

1.1. HIV

The human immunodeficiency virus (HIV) is responsible for the acquired immunodeficiency syndrome (AIDS) and causes about two million deaths per year, which is why intensive research has been conducted within the last three decades to develop a vaccine. However, clinical trials revealed disappointing efficacies, and yet no vaccine against HIV has been approved by regulatory authorities [1].

In particular, a major challenge is the genetic and phenotypic variability of HIV type 1 (HIV-1), as it facilitates evasion from both antibody and T-cell responses [2]. Reasons for this are the diversity of ornamental HIV-1 strains, the frequent occurrence of mutations, and the rapid spread of intracellular DNA reservoirs. The latter remains undetectable to the immune system for a long time [3].

In general, several vaccine classes, which can be divided into live and inactivated vaccines, are available for vaccine development. Despite some advances in the tools to induce antibody and T-cell responses, live attenuated vaccines and attenuated natural infections have been the only partially successful approaches against HIV-1 to date [4,5]. A prerequisite for a preventive vaccine against HIV is that it elicits the immune response in the form of the generation of T cells, which prevent the spread of the virus with almost no

time delay after infection. A vaccine that induces both a broadly neutralizing antibody and a T-cell response could be a possible approach [6].

Virus-like particles (VLPs) have many of the same desirable properties as live attenuated vaccines while being significantly less risk-associated and have shown promise as an approach to antigen presentation [7–9]. The reason for the high safety profile is the absence of any viral genome or components of it, making VLPs replication-incompetent—thus, avoiding the risk of infecting the vaccinees—and also rendering viral inactivation obsolete [10]. The efficient uptake of VLPs, caused by their repetitive and particulate structure, by antigen-representing cells (APCs) stimulates both antibody and T-cell responses [11,12]. This is already being exploited for VLP-utilizing vaccination against human papillomavirus (HPV), which is considered, amongst others, a major trigger of cervical carcinoma in women [13]. Another relatively new but promising vaccine class is so-called virus-like vaccines (VLV), which are a combination of virus-like particles and replication-deficient viral vectors [14].

1.2. QbD-Based Process Development

The development of efficient manufacturing processes for biopharmaceuticals is increasingly supported by in silico methods [15–17]. By applying mechanistic, physicochemical-based process models, wet-lab experiments can be reduced, and more efficient process development can be achieved [18,19]. In this context, the quality-by-design (QbD) concept should be the focus of the entire process development and optimization [20] (Figure 1). Consistent implementation of the principles captured in ICH Q8 to Q12 guidelines [21–26] reveals knowledge-based causality between critical process parameters and product quality and ensures consistent product quality in manufacturing by adhering to a predefined design space [27,28]. In order to span the design space, the quality target product profile (QTTP) is first defined, from which the critical quality attributes (CQAs) emerge [29,30]. In process development, classical process parameters such as productivity are defined as CQAs in addition to toxicity, bioavailability, etc. [31]. A risk analysis can then be performed, which can be qualitative in the form of an Ishikawa analysis or quantitative by means of an impact ranking [29,30,32]. The latter also allows a decision to be made as to whether a parameter should be investigated univariately or multivariately [29,30]. This investigation of the parameters to arrive at the design space is performed via a statistical Design of Experiments (DoE). In addition to experimental process development, a DoE can also be used for model validation, firstly to determine the influence of the investigated factors on the CQAs and secondly to determine a design space [33].

Operation within the design space must be ensured by robust process control. For this purpose, classical control approaches such as PID controllers can be used. However, newer approaches, such as Advanced Process Control (APC) in the form of model-based control, are more efficient and offer the advantage that predictive process control can be ensured (Schmidt et al., 2022). The latter is made possible by the fact that in real time, through the use of process analysis technologies (PAT), information about the critical process parameters can be obtained continuously. This enables real-time release testing (RTRT), eliminating bottlenecks in the production of critical biologics [34]. Furthermore, PAT, in combination with the gained and documented process understanding, enables continuous improvement of the process based on new process data [35,36].

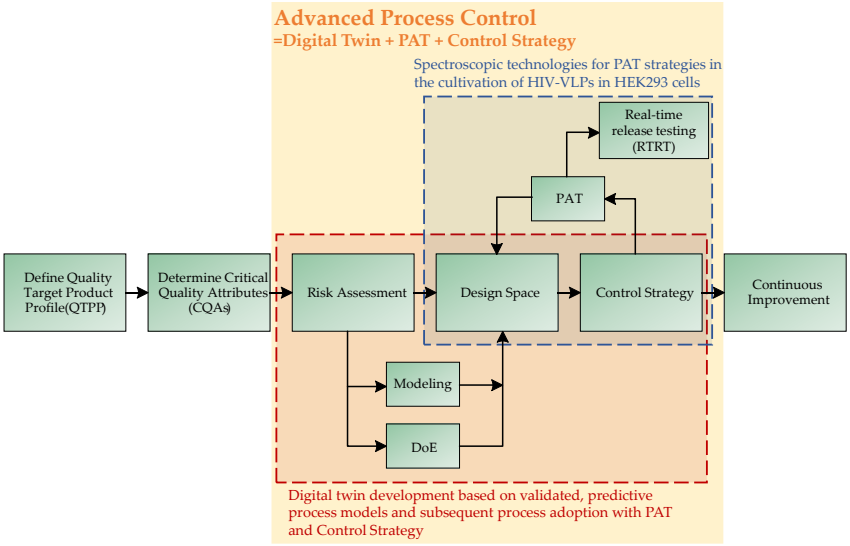


Figure 1. Workflow of model validation based on QbD-based process development. After defining the QTTP and CQAs, a risk assessment of the influence of the defined CQAs should be performed. Based on the risk assessment, a design space can be spanned either by means of a DoE or by modeling approaches. The design space leads to a control strategy, which is assisted by PAT. Furthermore, the implementation of a control strategy allows continuous process optimization [37].

In order to be able to realize the points mentioned above, such as continuous monitoring, control, and optimization of the process, a digital twin is required. The distinction between simple digital models and full digital twins is based on the process depth as well as the extent of information transfer with the physical process. Figure 2 shows the different levels of a digital model, from the time-independent steady-state model to the predictive digital twin that receives data from the process in real time and can control it based on the model [38].

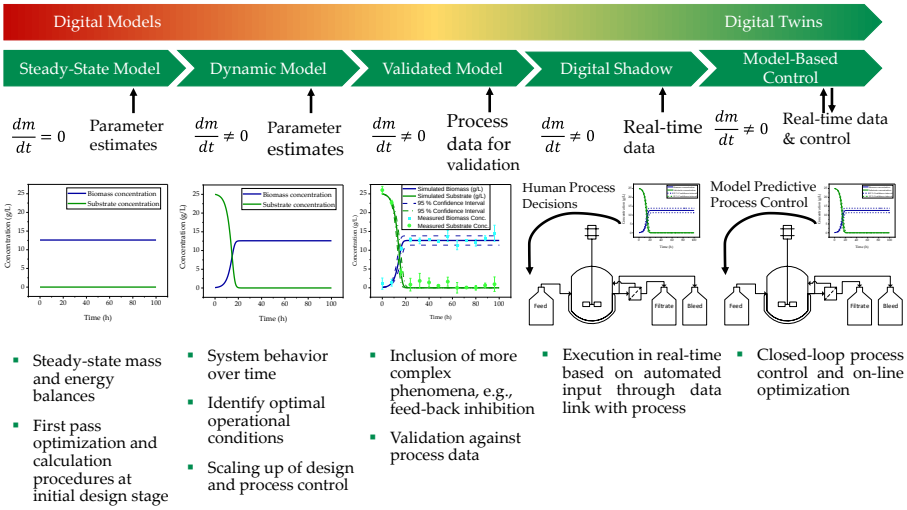


Figure 2. Levels of a digital twin depending on model depth and the amount of information exchanged with the physical process. From a simple steady-state model to a fully developed digital twin for model-based control [39].

1.3. Model Validation

The prerequisite for the use of the model as a digital twin in the QbD-based process is the quantitative and distinctive model validation (Figure 3) [40]. First, the task and application of the dynamic metabolic model must be defined. A verification follows in which it is shown that the model can correctly reproduce processes such as substrate

consumption, cell growth, and VLP formation. Due to the high proportion of Monod-based formation and consumption, special effort must be made to correctly implement stoichiometry. If the results are plausible, the sensitivity is investigated. For this purpose, a DoE can be performed. By means of statistical evaluation, rough design spaces can be generated. For the developed process model to be used as a digital twin for closed-loop process control, it must represent the process sufficiently accurately and precisely. Therefore, the model failures must correspond sufficiently to the real target variables measured in the process (accuracy). Furthermore, robust control requires that the predictions are precise. The final step of model validation is to show that the model in the experimental set under study is at least as accurate and precise as the observation in the physical process [16,41].

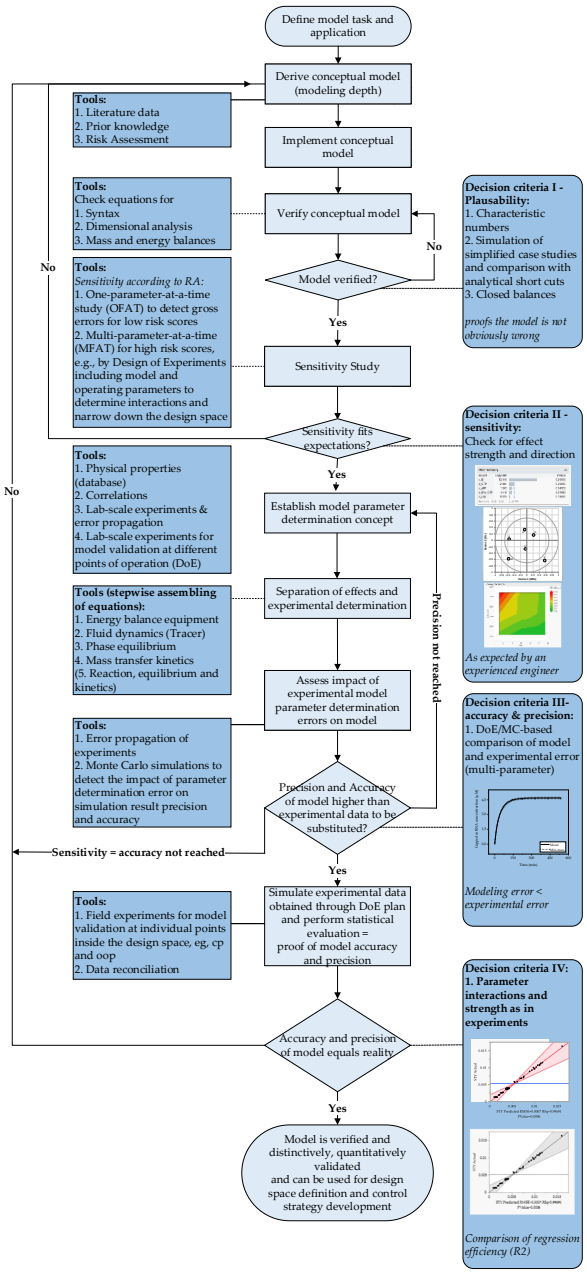


Figure 3. Decision tree for a process model validation according to Sixt et al. Applying this workflow allows a quantitative evaluation of the model quality based on mechanistic and statistical decision criteria. A rigorous execution of the procedure leads to a distinctively and quantitatively validated rigorous process model [39,40].

The aim of this study was to develop a predictive metabolic model for the production of HI-VLPs in HEK293 cells for the process developed and optimized by Helgers et al. [19] that meets the qualitative and quantitative requirements of QbD-compliant model validation and is consequently suitable as both a predictive process model and a basis for a digital twin. For this purpose, the model was first verified to show that it can sufficiently accurately predict the experimental courses. Then, the main influencing factors were identified using sensitivity analysis, and finally, the accuracy of the experimental data and the precision of the model were verified.

2. Materials and Methods

The process model was adapted and validated on data from previously published cultivations [19]. HIV-Gag VLPs were produced on a laboratory scale in 2 L glass bioreactors (Sartorius, Göttingen, Germany). For this purpose, 1 L of SMD medium (CellcaCHO Expression Platform, Sartorius) was inoculated with approximately 0.5×10^6 cells/mL. The temperature was 37 °C, and the pH was controlled to 7.1 by adding 1 M carbonate solution or by gassing with CO_2 as acid. The dissolved oxygen concentration was kept constant at 60% air saturation by gassing with 150 mL/min of air and adding pure oxygen if needed. A segmented three-blade impeller with a blade pitch of 30° was used as the stirrer, which was operated constantly at 433 rpm. An in situ turbidity probe (Transmission, 880 nm, HiTec Zang GmbH, Herzogenrath, Germany) was used for real-time quantification of total cell concentration during bioreactor cultures. After an initial batch phase of typically three days, different feeding strategies were tested. The results shown here depict feeding with reduced glucose concentration. For this purpose, HEK FS (XCell, Bielefeld, Germany) was added from the third day onwards so that a glucose concentration of 3–3.5 g/L was achieved after feeding, whereby the glucose concentration in the feed was 40 g/L. Initial cultivations have shown that reduced glucose concentration is positive for productivity and maximum achievable live cell number concentration [19]. Cultivation was continued for approximately 12 days, feeding once daily as described above.

Samples were taken once daily, each time before feeding, and live and total cell number concentrations were determined by Trypan blue method, using a Cedex XS (Roche, Basel, Switzerland). Subsequently, the sample was centrifuged at 500× *g* for 5 min to achieve separation of cells while minimizing product loss. Samples for subsequent analysis of substrates or metabolites were then filtered through a 0.2 µm filter, whereas samples for product analysis were not treated further. Glucose and lactate concentrations were determined in duplicate using LaboTRACE Compact (TRACE Analytics GmbH, Braunschweig, Germany). The amino acid concentration was measured by RP chromatography (InfinityLab Poroshell HPH-C18; 3.0 × 100 mm; 2.7 µm; Agilent Technologies, Santa Clara, CA, USA) and precolumn derivatization of amino acids with orthophthalic aldehyde (OPA) reagent in basic medium. If necessary, samples were diluted beforehand to ensure complete derivatization. The column was tempered to 40 °C for better separation. The product concentration was determined by SEC.

3. Modelling

3.1. Process Model

All simulations were performed using Aspen Custom Modeler V12 (AspenTechnology Inc., Bedford, MA, USA). A general mass balance model for stirred tank reactors served as the basis for process modeling. This describes the temporal (t) change in mass (m_i) of component i by taking into account the inflowing (index in) and outflowing (index out) concentrations (c_i) as well as the reaction rates (r_i).

$$\frac{dm_i}{dt} = \dot{V}_{in} \cdot c_{i_{in}} - \dot{V}_{out} \cdot c_i \pm r_i \cdot V \quad (1)$$

The reaction rates were described by the equations given in the following section. For the modeling of fed-batch cultivations performed here, the volumetric changes due to sampling

and the addition of feed solution were taken into account, resulting in the volume (V) balance, which is composed of the inlet (index in) and outlet (index out) volume flows ($\dot{V}$).

$$\frac{dV}{dt} = \dot{V}_{in} - \dot{V}_{out} \tag{2}$$

3.2. Modeling of the Intracellular Metabolism of HEK Cells

In general, therefore, for the reaction rate r_i in Equation (1), the expression shown in Equation (3), where v depends on whether the reaction is one without or with activation or inhibition, the incoming v is described by one of the Equations (4)–(6). Here, stoichiometry is taken into account by the stoichiometric coefficient α_i.

For modeling the intracellular metabolism of HEK293F suspension cells, the model for CHO DG44 cells published by Helgers et al. [42] was used as a starting point. The model consists of multiplicative Michaelis–Menten equations, or variants derived from them, representing the inhibition (Equation (5)) or activation (Equation (6)) of a reaction as a consequence of an accumulation of activator ([A]) or inhibitor ([I]), respectively.

$$r_i = \sum_{i=1}^{n} \alpha_i \cdot v_i \tag{3}$$

$$v = \frac{v_{max} \cdot [S]}{K_S + [S]} \tag{4}$$

$$v = \frac{v_{max} \cdot [S]}{K_S \cdot \left(1 + \frac{[I]}{K_I}\right) + S} \tag{5}$$

$$v = \frac{v_{max} \cdot [S] \cdot \left(1 + \frac{\beta \cdot [A]}{\alpha \cdot [K_A]}\right)}{K_S \cdot \left(1 + \frac{[A]}{K_A}\right) + [S] \cdot \left(1 + \frac{[A]}{\alpha \cdot K_A}\right)} \tag{6}$$

K_s, KA, and KI are the Michaelis–Menten constants for the substrate ([S]), activator ([A]), and inhibitor ([I]), respectively.

The specific growth rate μ was also formulated in terms of a multiplicative Monod equation, taking into account the concentrations of glucose, glucose-6-phosphate, and ribulose-5-phosphate as precursors of nucleotides, citrate as a precursor of lipids, and the concentrations of all 20 proteinogenic amino acids. For each substrate ([S_i]), a specific value for the growth-related Monod constant (K_{S_i}) was determined.

$$\mu = \mu_{max} \cdot \prod_{i=1}^{n} \frac{[S_i]}{K_{S_i} + [S_i]} \tag{7}$$

The maximum growth rate of HEK293F cells used here was determined from experimental data and is $0.0282 \pm 0.0033\ h^{-1}$.

The model describes central metabolism and includes glycolysis as well as its activation and inhibition, the formation and consumption of various substrates in the context of abundance metabolism in connection with high glucose consumption rates, the citrate cycle, and anaplerotic reactions of various amino acids. Glutamine and glutamic acid are particularly noteworthy in this regard, as they play an important role as a carbon supplier for the citrate cycle and as a nitrogen source in the metabolism of many animal cell lines.

Data for metabolite consumption and cellular composition were taken from the literature and adjusted [43]. An average molecular mass of 107.5 g mol^{-1} was assumed for the proteins composing the cells. The sequence was used to determine amino acid consumption for Gag synthesis. It was also assumed that all lipids are derived from citrate in the citric acid cycle and that nucleic acid synthesis proceeds as a lumped reaction from ribulose 5-phosphate and glucose 6-phosphate. The requirement of ATP for biomass and product synthesis and a conversion factor of 3.15×10^{-4} gDW 10^{-6} cells were also taken from the literature [44,45].

The model was adapted to the new cell type according to the observed metabolic phenotype as well as the altered product formation mechanism. Multiplicative Michaelis–Menten kinetics was assumed for product formation, taking into account all amino acids weighted by their proportion in the Gag protein, as well as citrate as a precursor of lipids. Furthermore, the formation of virus-like particles by multimerization of Gag protein in the host cell membrane was considered by an appropriate kinetic equation. Therefore, the following changes were made to the initial model:

- SDHH was assumed to be reversible;
- Adjustment of the product formation mechanism;
- Stronger weighting of ammonium as an inhibitor of growth and inducer of cell death since the cells used tolerate high lactate concentrations relatively well;
- Adjustment of several parameter values as detailed in Table 1.

Table 1. Modified parameters and new parameter values.

Parameter.	Value	Unit
$K_{growth_{dNH4}}$	3.38×10^{5}	mM
$K_{A_{AMP/ATP}}$	1.20×10^{2}	mM 10^{-6} cells
$K_{m_{ADP/ATP}}$	3.40×10^{-3}	mM 10^{-6} cells
$K_{m_{ATP}}$	8.35×10^{-5}	mM 10^{-6} cells
$\alpha_{AMP/ATP}$	2.43×10^{-4}	-
$\beta_{AMP/ATP}$	2.20	-
$v_{AA\ to\ SUC_{max}}$	4.61×10^{-4}	mM 10^{-6} cells h^{-1}
$v_{AK_{f,max}}$	1.43×10^{-7}	mM 10^{-6} cells h^{-1}
$v_{AlaTA_{f,max}}$	2.17×10^{-4}	mM 10^{-6} cells h^{-1}
$v_{AlaTA_{r,max}}$	2.4×10^{-5}	mM 10^{-6} cells h^{-1}
$v_{ASTA_{max}}$	5.03×10^{-5}	mM 10^{-6} cells h^{-1}
$v_{GlnT_{f,max}}$	5.15×10^{-4}	mM 10^{-6} cells h^{-1}
$v_{GlnT_{r,max}}$	7.92×10^{-6}	mM 10^{-6} cells h^{-1}
$v_{HK_{max}}$	2.63×10^{-4}	mM 10^{-6} cells h^{-1}
$v_{LDH_{max}}$	8.50×10^{-7}	mM 10^{-6} cells h^{-1}
$v_{LDH_{f,max}}$	1.28×10^{-2}	mM 10^{-6} cells h^{-1}
$v_{LDH_{r,max}}$	6.78×10^{-4}	mM 10^{-6} cells h^{-1}
$v_{leak_{max}}$	1.40×10^{-5}	mM 10^{-6} cells h^{-1}
$v_{PGI_{f,max}}$	8.89×10^{1}	mM 10^{-6} cells h^{-1}
$v_{PGK_{max}}$	1.5×10^{-3}	mM 10^{-6} cells h^{-1}
$v_{PK_{max}}$	4.00×10^{-4}	mM 10^{-6} cells h^{-1}
$v_{resp_{max}}$	3.42×10^{-3}	mM 10^{-6} cells h^{-1}
$v_{SDHH_{f,max}}$	8.50×10^{-6}	mM 10^{-6} cells h^{-1}
$v_{SDHH_{r,max}}$	3.10×10^{-6}	mM 10^{-6} cells h^{-1}

Furthermore, the assumptions made for the initial model still apply [42]:

- Assumption of ideal mixing in the stirred tank reactor, i.e., homogeneous pH, temperature, and concentration profile of chemical species in the liquid phase. Moreover, the pH, pO_2, and temperature were assumed as constant;
- The model is an unsegregated, structured model. This means that an "average cell" was assumed to represent the entire cell population, and the phases of the cell cycle were not considered;
- Limited the number of metabolites, where possible, and reasonable reaction pathways were lumped by only considering the main reaction of interest in the pathway, representing a branch in a metabolic pathway, or that are taken up directly from the medium into the cell. This helps to reduce model complexity without sacrificing the predictive power of the metabolic model;
- Constant enzyme amounts: the maximum reaction rate of an enzyme-catalyzed reaction depends on the enzyme amount. The enzyme amount depends on the transcription and translation rates of the respective genes, which may be induced or inhibited by the substrate and product concentrations and other influencing variables. To fully elucidate such relationships, "omics" data are required, which were not available in the context of this work. Thus, constant enzyme amounts were assumed in this model;
- New cells and virus-like particles were assumed to be directly formed from precursors present in the cell (amino acids, citrate (representing lipids), and R5P (representing nucleotides);
- Components such as vitamins, trace elements, phospholipid precursors, growth factors, etc., were assumed to be present in sufficient amounts;
- The cell volume was assumed to be constant during the cultivation, i.e., the change in concentration of intracellular substrates due to changes in cell volume were not considered;
- The composition of the cells was assumed to be constant.

For a digital twin, process parameters such as fluid dynamics and energy balance are necessary for addition to the accurate description of all concentration profiles. These form the basis for scale-up by accurately predicting non-ideality. In the work described here, standardized laboratory equipment on a 1-L scale was used, in which dynamic fluid imbalances do not play a significant role with respect to mixing and residence time behavior, so non-ideal effects could be neglected here. Approaches to account for residence time and energy balance non-idealities in stirred tank reactors of different scales are described in detail in the literature, e.g., in [46].

The energy balance of the balance space investigated in this study includes the accumulation as the difference between outgoing and incoming thermal energy (Equation (8), with the density ρ_S and the volume Vs of the medium and its specific heat capacity c_p) as well as the power consumption by the stirrer ($\dot{Q}_{St}$, Equation (9)). The ratio of flow resistance to inertial force is represented by the Newton number (Ne). The power consumption of the stirrer also depends on the rpm (n) and the diameter (d_R) of the stirrer, as well as the density of the medium. In order to be able to operate the bioreactor constantly at the ideal temperature T of 37 °C, a temperature control via the bioreactor's double jacket is necessary. The heat supplied or dissipated here ($\dot{Q}_{Cool}$) depends on both the heat transfer coefficient k_W and the exchange surface A (cf. Equation (10)).

$$\rho_S \cdot c_p \cdot V_S \cdot \frac{dT}{dt} = \dot{Q}_{St} - \dot{Q}_{Cool} \quad (8)$$

$$\dot{Q}_{St} = Ne \cdot n^3 \cdot d_R \cdot \rho_S \quad (9)$$

$$\dot{Q}_{Cool} = k_W \cdot A_M \cdot \Delta T \quad (10)$$

A graphical illustration of the model can be seen in Figure 4.

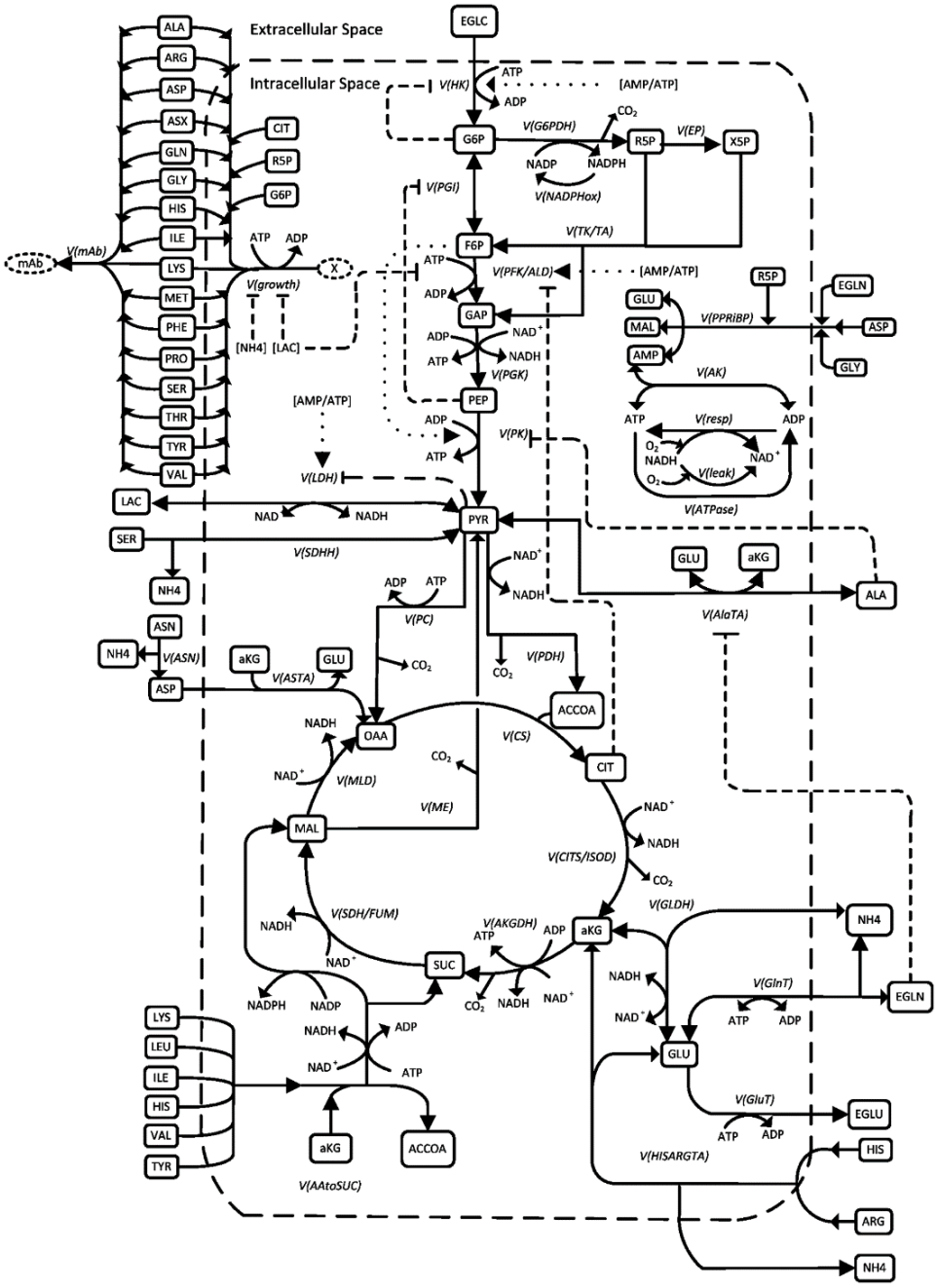


Figure 4. The metabolic network is described by the model. Biochemical reactions are depicted by solid lines, inhibition mechanisms by dashed lines and activation mechanisms by dotted lines. Figure taken from [45].

4. Model Parameter Determination

The parameter values determined for the model adapted to CHO DG44 were not able to describe the cultivation processes for the HEK293 cells sufficiently well. Accordingly, the values of various parameters had to be adjusted. Table 1 gives an overview of this.

Of the total 24 parameters that had to be adjusted, 18 parameters were maximum reaction rates, 4 were Michaelis–Menten constants, and 2 were dimensionless parameters that play a role in the regulation of glycolysis. This observation was consistent with experience from the previous publication as well as the literature [45], where in each case, it was found that the greatest sensitivities in the model emanate primarily from the maximum reaction rates as well as the regulation of glycolysis.

5. Model Validation

5.1. Model Verification

The first step in the model validation process was the verification of the model by checking whether the predictions of the model were plausible (see Figure 3). The correct implementation of the stoichiometry and syntax was checked, and the cell growth, product formation, and substrate consumption were examined for errors (see Section 5.1.1). If the plausibility of the model was given, the sensitivity of the influencing parameters on the system was then examined and discussed in Section 5.1.2. This represents the second decision criterion in the workflow. Monte Carlo simulations can be used to determine the accuracy and precision of the model in Section 5.2.

5.1.1. Plausibility

The aim of the plausibility analysis is to qualitatively compare the simulation results with experimental data. If the effect strengths and directions from the sensitivity analysis are correct, qualitatively similar courses should appear in the simulation results and the experimental data. Figure 5 shows the results of fed-batch cultivation as well as the simulation results. Spikes in the simulation correspond to the concentration changes caused by the daily addition of feed, starting from day three. Samples were taken before feed addition, i.e., experimentally determined concentrations should intersect the simulation results at their low points.

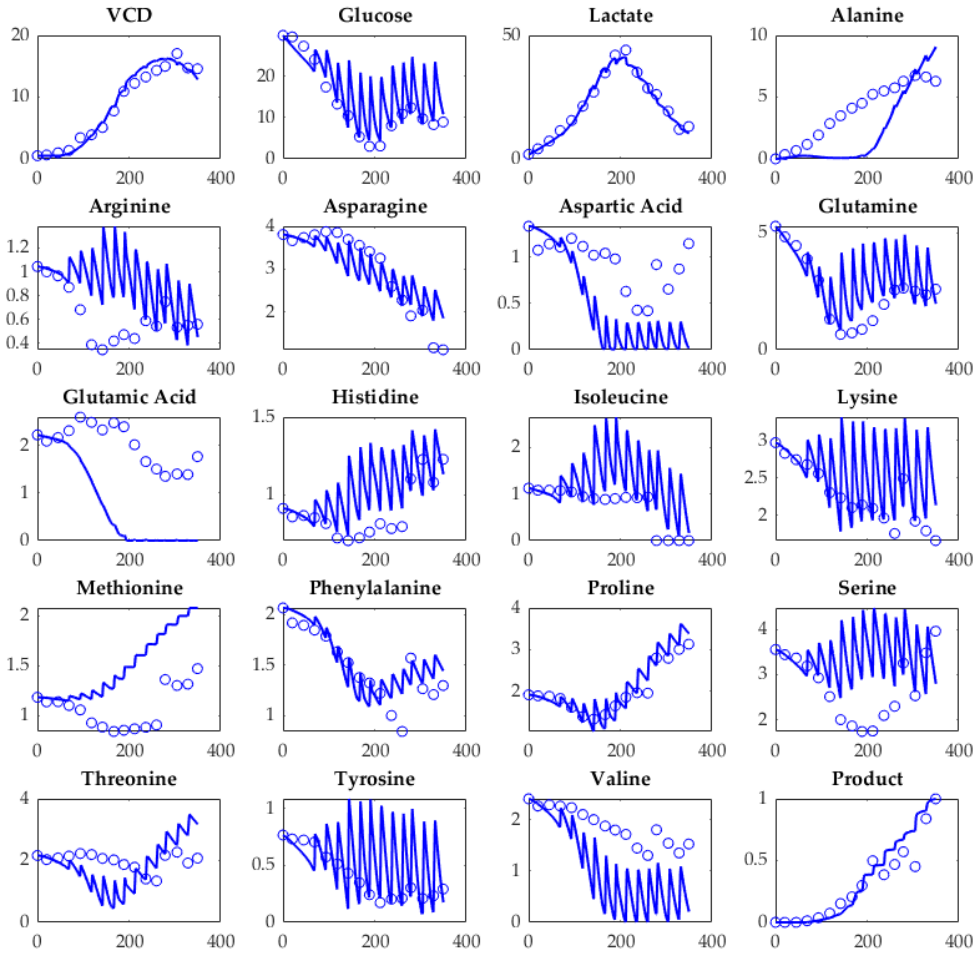


Figure 5. Experimentally measured concentration curves (dots) and simulated concentration curves (solid line) over the processing time in hours. Live cell count (VCD) in 1×10^6 cells/mL, the normalized product concentration is dimensionless, all other concentrations in mM.

The model is able to predict well the concentration curves of live cell count (VCD), VLP (product), glucose, lactate, asparagine, and glutamine concentration, as well as the concentration curves of most amino acids. Moreover, the model is able to correctly predict the switch in the metabolism towards lactate consumption because of limiting glucose concentration after approximately 200 h of process time. The decrease in growth rate and increased cell death after 250 and 300 h of process time, respectively, are correctly predicted by the model. The decreased uptake rate of glucose and glutamine from this point is also correctly predicted. The largest deviations between simulation and experiment are seen for alanine, aspartic acid, and glutamic acid, but with the trend of increasing and decreasing concentrations being correctly predicted. Only the decrease in the concentration of methionine is not correctly predicted in its direction. With this exception, the model can be considered plausible as methionine is not sensitive based on the OFAT plausibility studies.

5.1.2. Sensitivity

After the implementation and verification of the conceptual model, a sensitivity study follows according to the model validation concept described above. The goal of the sensitivity study was to quantify the effect sizes predicted by the model and the directions in which the effects operate. One-factor-at-a-time (OFAT) and multiple-factors-at-a-time studies were used as tools for sensitivity determination. The goal of OFAT studies is to detect gross errors in the model in the form of unexpected effect sizes or directions. OFAT studies are performed for parameters for which a low-risk value was identified in the previous risk assessment. MFAT studies are conducted for those parameters that received a high-risk value in the risk assessment. In order to detect and quantify interactions between these parameters and to make the number of simulations be performed as efficiently as possible, MFAT studies are performed within a statistical design of experiments (DoE). As a decision criterion, the determined sensitivities are compared with already known, comparable sensitivities for corresponding parameters, and empirical values of experienced engineers are taken into account. For the MFAT studies, statistical quantities from the evaluation of the DoE can also be included.

In this study, sensitivities were determined using a full factorial experimental design with a total of 216 simulations. For this purpose, the concentrations of the medium used were varied by ±50%. As an example, the statistical evaluation of the experimental design with respect to the normalized product concentration is shown below. As can be seen in Figure 6, the normalized HIV-Gag VLP concentration can be predicted with an R^2 of >0.998 with a *p*-value of <0.0001 by the applied regression model for the DoE.

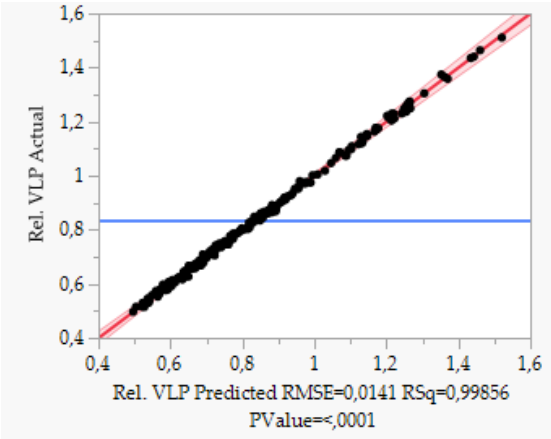


Figure 6. Actual by predicted plot of DoE simulations for normalized HIV-Gag VLP concentration.

Figure 7 shows the effect summaries determined in the evaluation of the DoE for the 15 parameters with the highest significance. It can be seen that the glucose concentration has the greatest influence, followed by the concentrations of isoleucine, glutamic acid, and

several other amino acids and their interaction. The effects all work in the positive direction, i.e., a higher concentration should lead to a higher VLP concentration according to the model prediction. Figure 8 graphically shows the influence of the two parameters with the highest significance.

Source	LogWorth	PValue
Glucose	103,043	0,00000
Isoleucine	75,452	0,00000
Glutamic acid	57,231	0,00000
Glucose*Isoleucine	52,983	0,00000
Arginine	48,588	0,00000
Histidine	43,871	0,00000
Methionine	42,514	0,00000
Aspartic acid	38,074	0,00000
Phenylalanine	21,175	0,00000
Glucose*Arginine	18,710	0,00000
Glycine	18,348	0,00000
Lysine	13,233	0,00000
Glucose*Aspartic acid	12,775	0,00000
Leucine*Threonine	9,595	0,00000
Arginine*Isoleucine	9,466	0,00000

Figure 7. Effect summaries of DoE simulations of the normalized HIV-Gag VLP concentration.

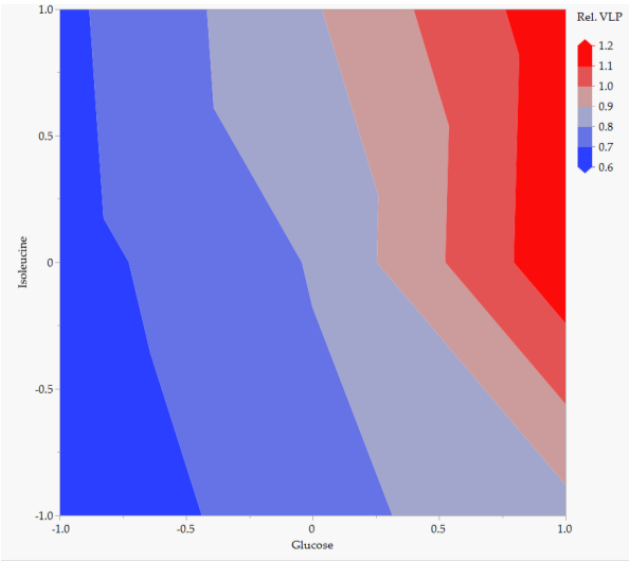


Figure 8. Contour plots for simulations of the normalized HIV-Gag VLP concentration depending on the Isoleucine and Glucose concentration (both depicted as normalized values).

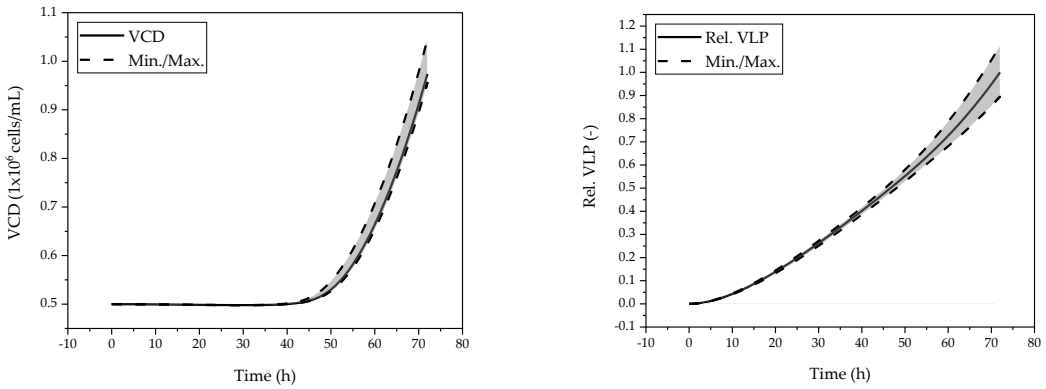
It is important to emphasize that these effects are not generally valid but apply to the specific conditions and the specific medium in the context of the process. Other process conditions, however, may lead to other significant effects. It should also be noted that the goal of DoE is not primarily to optimize the medium but rather to identify particularly sensitive components within the model. Once the model is validated, it can be used for the prediction of optimized cultivation conditions.

In order to reach the last milestone in the model validation process, the model must be able to predict the cultivation process adequately, accurately, and precisely.

5.2. Accuracy and Precision

In addition to the main effects, optimized media concentrations can be determined from the DoE. These are used in the following to investigate the model accuracy using 30 Monte Carlo simulations (cf. Figure 9). The robustness of the model prediction was examined by combining the random, normally distributed deviation of 5% of the media

concentrations. This yields a confidence interval (*t*-test) of >95% with >99% certainty. The results of the model predictions deviate less than 15% from each other and can be assumed to be sufficiently accurate, thus fulfilling the criterion of model accuracy.



(**a**) (**b**)

Figure 9. Monte Carlo Simulations using the Optimal Operating Point (derived from DoE) for initial substrate concentrations with a standard deviation of ±5%. (**a**) Viable cell concentration during initial batch phase, (**b**) normalized HIV-Gag VLP concentration.

6. Discussion

The present study demonstrated the distinct and quantitative validation of a dynamic metabolic model used to simulate fed-batch cultivation of an HIV Gag-VLP-producing HEK293 cell line. The model is based on previous work on CHO DG44 cells [19,42,47] and was adapted to the specificities of the cells used here [19,37,48]. By applying distinct and quantitative statistical methods [40], the model was shown to predict the cultivation progression of HEK293 cells accurately and precisely with regard to experimental data. Thus, the model is suitable as a digital twin for process development and design as well as for model predictive control in the context of an advanced process control strategy. The use of modern PAT detectors such as Raman, FTIR, and MALS/DLS [35,49–54] in combination with developed PLS models [55] allow conclusions to be drawn about critical process parameters of the cultivation-like product (VLP), glucose, lactate, and other by-product concentrations. This real-time generated information can be used by the digital twin for model-based control. In this case, this includes feeding and other process parameters such as gassing and pH adjustment through the addition of base. By predicting the product concentration, it can also generate important information for subsequent process steps. These depend on the optimal harvesting time, the total cell count, the product concentration, the purity, and the expected size distribution of the VLPs.

Furthermore, the model could be used for the development of processes to partially replace costly screening experiments for media optimization with in silico experiments. As a distinct advantage over conventionally used Monod models for describing concentration profiles during cultivation, the model presented here can clearly and comprehensibly contextualize the relationships due to the mechanistic connection of all formation and consumption rates. For example, the model can predict the shift in lactate metabolism with high accuracy in terms of time and concentration, depending on the pyruvate concentration present. Upcoming work will address the further development and implementation of the model as an integrated digital twin to enable model predictive process control in manufacturing. For this, an implementation of hard and software communication is crucial to transfer data measured in real-time by PAT into the model and to be able to make the necessary adjustments for process control.

Author Contributions: Conceptualization, J.S. (Jochen Strube); Funding acquisition, J.S. (Jochen Strube); Investigation, H.H., A.H. and A.S.; Project administration, J.S. (Jochen Strube); Software, H.H. and A.S.; Supervision, J.S. (Jörn Stitz) and J.S. (Jochen Strube); Writing—original draft, H.H., A.H. and A.S.; Writing—review and editing, H.H., A.H., J.F.R., J.S. (Jörn Stitz), A.S. and J.S. (Jochen Strube). All authors have read and agreed to the published version of the manuscript.

Funding: J.F.R. was supported by a grant from the German Federal Ministry of Education and Research, funding program Forschung an Fachhoschulen, contract number 13FH767IA6 to J.S. (Jörn Stitz).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data generated in this study are available from the authors upon reasonable request.

Acknowledgments: The authors would like to acknowledge the fruitful discussions with their Clausthal University of Technology colleagues. The authors acknowledge financial support from the Open Access Publishing Fund of the Clausthal University of Technology.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hammonds, J.; Chen, X.; Zhang, X.; Lee, F.; Spearman, P. Advances in methods for the production, purification, and characterization of HIV-1 Gag-Env pseudovirion vaccines. *Vaccine* **2007**, *25*, 8036–8048. [CrossRef]
2. Wang, H.; Wolock, T.M.; Carter, A.; Nguyen, G.; Kyu, H.H.; Gakidou, E.; Hay, S.I.; Mills, E.J.; Trickey, A.; Msemburi, W.; et al. Estimates of global, regional, and national incidence, prevalence, and mortality of HIV, 1980–2015: The Global Burden of Disease Study 2015. *Lancet HIV* **2016**, *3*, e361–e387. [CrossRef]
3. Haynes, B.F.; Shaw, G.M.; Korber, B.; Kelsoe, G.; Sodroski, J.; Hahn, B.H.; Borrow, P.; McMichael, A.J. HIV-Host Interactions: Implications for Vaccine Design. *Cell Host Microbe* **2016**, *19*, 292–303. [CrossRef] [PubMed]
4. Daniel, M.D.; Kirchhoff, F.; Czajak, S.C.; Sehgal, P.K.; Desrosiers, R.C. Protective effects of a live attenuated SIV vaccine with a deletion in the nef gene. *Science* **1992**, *258*, 1938–1941. [CrossRef] [PubMed]
5. Lifson, J.D.; Rossio, J.L.; Piatak, M.; Parks, T.; Li, L.; Kiser, R.; Coalter, V.; Fisher, B.; Flynn, B.M.; Czajak, S.; et al. Role of CD8(+) lymphocytes in control of simian immunodeficiency virus infection and resistance to rechallenge after transient early antiretroviral treatment. *J. Virol.* **2001**, *75*, 10187–10199. [CrossRef]
6. Williamson, A.-L.; Rybicki, E.P. Justification for the inclusion of Gag in HIV vaccine candidates. *Expert Rev. Vaccines* **2016**, *15*, 585–598. [CrossRef]
7. Minor, P.D. Live attenuated vaccines: Historical successes and current challenges. *Virology* **2015**, *479*, 379–392. [CrossRef]
8. Mohsen, M.O.; Zha, L.; Cabral-Miranda, G.; Bachmann, M.F. Major findings and recent advances in virus-like particle (VLP)-based vaccines. *Semin. Immunol.* **2017**, *34*, 123–132. [CrossRef]
9. Deml, L.; Speth, C.; Dierich, M.P.; Wolf, H.; Wagner, R. Recombinant HIV-1 Pr55gag virus-like particles: Potent stimulators of innate and acquired immune responses. *Mol. Immunol.* **2005**, *42*, 259–277. [CrossRef]
10. Roldão, A.; Mellado, M.C.M.; Castilho, L.R.; Carrondo, M.J.T.; Alves, P.M. Virus-like particles in vaccine development. *Expert Rev. Vaccines* **2010**, *9*, 1149–1176. [CrossRef]
11. Cervera, L.; Gòdia, F.; Tarrés-Freixas, F.; Aguilar-Gurrieri, C.; Carrillo, J.; Blanco, J.; Gutiérrez-Granados, S. Production of HIV-1-based virus-like particles for vaccination: Achievements and limits. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 7367–7384. [CrossRef] [PubMed]
12. Ludwig, C.; Wagner, R. Virus-like particles-universal molecular toolboxes. *Curr. Opin. Biotechnol.* **2007**, *18*, 537–545. [CrossRef] [PubMed]
13. Kirnbauer, R.; Taub, J.; Greenstone, H.; Roden, R.; Dürst, M.; Gissmann, L.; Lowy, D.R.; Schiller, J.T. Efficient self-assembly of human papillomavirus type 16 L1 and L1-L2 into virus-like particles. *J. Virol.* **1993**, *67*, 6929–6936. [CrossRef] [PubMed]
14. Andersson, A.-M.C.; Schwerdtfeger, M.; Holst, P.J. Virus-Like-Vaccines against HIV. *Vaccines* **2018**, *6*, 10. [CrossRef] [PubMed]
15. Chhatre, S.; Farid, S.S.; Coffman, J.; Bird, P.; Newcombe, A.R.; Titchener-Hooker, N.J. How implementation of Quality by Design and advances in Biochemical Engineering are enabling efficient bioprocess development and manufacture. *J. Chem. Technol. Biotechnol.* **2011**, *86*, 1125–1129. [CrossRef]
16. Schmidt, A.; Strube, J. Distinct and Quantitative Validation Method for Predictive Process Modeling with Examples of Liquid-Liquid Extraction Processes of Complex Feed Mixtures. *Processes* **2019**, *7*, 298. [CrossRef]
17. Kornecki, M.; Strube, J. Accelerating Biologics Manufacturing by Upstream Process Modelling. *Processes* **2019**, *7*, 166. [CrossRef]
18. Kornecki, M.; Schmidt, A.; Lohmann, L.; Huter, M.; Mestmäcker, F.; Klepzig, L.; Mouellef, M.; Zobel-Roos, S.; Strube, J. Accelerating Biomanufacturing by Modeling of Continuous Bioprocessing—Piloting Case Study of Monoclonal Antibody Manufacturing. *Processes* **2019**, *7*, 495. [CrossRef]

19. Helgers, H.; Hengelbrock, A.; Schmidt, A.; Rosengarten, J.; Stitz, J.; Strube, J. Process Design and Optimization towards Digital Twins for HIV-Gag VLP Production in HEK293 Cells, including Purification. *Processes* **2022**, *10*, 419. [CrossRef]
20. Beg, S.; Hasnain, M.S.; Rahman, M.; Swain, S. Introduction to Quality by Design (QbD): Fundamentals, principles, and applications. In *Pharmaceutical Quality by Design*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 1–17. ISBN 9780128157992.
21. ICH Expert Working Group. *Pharmaceuticel Development Q8(R2): ICH Harmonised Tripartite Guideline*; ICH Expert Working Group: Geneva, Switzerland, 2009.
22. ICH Expert Working Group. *Riskmanagement (Q9): ICH Harmonised Tripartite Guideline*; ICH Expert Working Group: Geneva, Switzerland, 2005.
23. ICH Expert Working Group. *Pharmaceutical Quality System (Q10): ICH Harmonised Tripartite Guideline*; ICH Expert Working Group: Geneva, Switzerland, 2008.
24. ICH. Q11 Step 5 Development and manufacture of drug substances. In *ICH Quality Guidelines: An Implementation Guide*; Teasdale, A., Elder, D., Nims, R.W., Eds.; Wiley: Hoboken, NJ, USA, 2017; pp. 639–665.
25. European Medicines Agency. *Q12 Step 5 Technical and Regulatory Considerations for Pharmaceutical Product Lifecycle Management*; European Medicines Agency: Amsterdam, The Netherlands, 2017.
26. Aksu, B.; Yeğen, G. Global regulatory perspectives on quality by design in pharma manufacturing. In *Pharmaceutical Quality by Design*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 19–41. ISBN 9780128157992.
27. Rajamanickam, V.; Babel, H.; Montano-Herrera, L.; Ehsani, A.; Stiefel, F.; Haider, S.; Presser, B.; Knapp, B. About Model Validation in Bioprocessing. *Processes* **2021**, *9*, 961. [CrossRef]
28. Alt, N.; Zhang, T.Y.; Motchnik, P.; Taticek, R.; Quarmby, V.; Schlothauer, T.; Beck, H.; Emrich, T.; Harris, R.J. Determination of critical quality attributes for monoclonal antibodies using quality by design principles. *Biologicals* **2016**, *44*, 291–305. [CrossRef] [PubMed]
29. CMC-Vaccines Working Group. *A-VAX: Applying Quality by Design to Vaccines*; Parental Drug Association: Bethesda, MD, USA, 2012.
30. CMC Biotech Working Group. *A-Mab: A Case Study in Bioprocess Development*; CASSS, International Separation Science Society: Emeryville, CA, USA, 2009.
31. Yu, L.X.; Amidon, G.; Khan, M.A.; Hoag, S.W.; Polli, J.; Raju, G.K.; Woodcock, J. Understanding pharmaceutical quality by design. *AAPS J.* **2014**, *16*, 771–783. [CrossRef] [PubMed]
32. Cogdill, R.P.; Drennen, J.K. Risk-based Quality by Design (QbD): A Taguchi Perspective on the Assessment of Product Quality, and the Quantitative Linkage of Drug Product Parameters and Clinical Performance. *J. Pharm. Innov.* **2008**, *3*, 23–29. [CrossRef]
33. Beg, S.; Swain, S.; Rahman, M.; Hasnain, M.S.; Imam, S.S. Application of Design of Experiments (DoE) in pharmaceutical product and process optimization. In *Pharmaceutical Quality by Design*; Beg, S., Hasnain, M.S., Eds.; Elsevier: Amsterdam, The Netherlands, 2019; pp. 43–64. ISBN 9780128157992.
34. Schmidt, A.; Helgers, H.; Lohmann, L.J.; Vetter, F.; Juckers, A.; Mouellef, M.; Zobel-Roos, S.; Strube, J. Process analytical technology as key-enabler for digital twins in continuous biomanufacturing. *J. Chem. Technol. Biotechnol.* **2022**, *97*, 2336–2346. [CrossRef]
35. Helgers, H.; Schmidt, A.; Lohmann, L.J.; Vetter, F.L.; Juckers, A.; Jensch, C.; Mouellef, M.; Zobel-Roos, S.; Strube, J. Towards Autonomous Operation by Advanced Process Control—Process Analytical Technology for Continuous Biologics Antibody Manufacturing. *Processes* **2021**, *9*, 172. [CrossRef]
36. Nayak, A.K.; Ahmed, S.A.; Beg, S.; Tabish, M.; Hasnain, M.S. Application of quality by design for the development of biopharmaceuticals. In *Pharmaceutical Quality by Design*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 399–411. ISBN 9780128157992.
37. Hengelbrock, A.; Helgers, H.; Schmidt, A.; Vetter, F.L.; Juckers, A.; Rosengarten, J.F.; Stitz, J.; Strube, J. Digital Twin for HIV-Gag VLP Production in HEK293 Cells. *Processes* **2022**, *10*, 866. [CrossRef]
38. Udugama, I.A.; Lopez, P.C.; Gargalo, C.L.; Li, X.; Bayer, C.; Gernaey, K.V. Digital Twin in biomanufacturing: Challenges and opportunities towards its implementation. *Syst. Microbiol. Biomanuf.* **2021**, *1*, 257–274. [CrossRef]
39. Helgers, H.; Hengelbrock, A.; Schmidt, A.; Strube, J. Digital Twins for Continuous mRNA Production. *Processes* **2021**, *9*, 1967. [CrossRef]
40. Sixt, M.; Uhlenbrock, L.; Strube, J. Toward a Distinct and Quantitative Validation Method for Predictive Process Modelling—On the Example of Solid-Liquid Extraction Processes of Complex Plant Extracts. *Processes* **2018**, *6*, 66. [CrossRef]
41. Zobel-Roos, S.; Schmidt, A.; Mestmäcker, F.; Mouellef, M.; Huter, M.; Uhlenbrock, L.; Kornecki, M.; Lohmann, L.; Ditz, R.; Strube, J. Accelerating Biologics Manufacturing by Modeling or: Is Approval under the QbD and PAT Approaches Demanded by Authorities Acceptable Without a Digital-Twin? *Processes* **2019**, *7*, 94. [CrossRef]
42. Helgers, H.; Schmidt, A.; Strube, J. Towards Autonomous Process Control—Digital Twin for CHO Cell-Based Antibody Manufacturing Using a Dynamic Metabolic Model. *Processes* **2022**, *10*, 316. [CrossRef]
43. Sheikh, K.; Förster, J.; Nielsen, L.K. Modeling hybridoma cell metabolism using a generic genome-scale metabolic model of Mus musculus. *Biotechnol. Prog.* **2005**, *21*, 112–121. [CrossRef] [PubMed]
44. Nolan, R.P.; Lee, K. Dynamic model of CHO cell metabolism. *Metab. Eng.* **2011**, *13*, 108–124. [CrossRef] [PubMed]
45. Robitaille, J.; Chen, J.; Jolicoeur, M. A Single Dynamic Metabolic Model Can Describe mAb Producing CHO Cell Batch and Fed-Batch Cultures on Different Culture Media. *PLoS ONE* **2015**, *10*, e0136815. [CrossRef]
46. Huter, M.J. Modellunterstützte Prozessauslegung Unterschiedlicher Grundoperationen am Beispiel von kontinuierlicher Ultrafiltration und Absatzweiser Kristallisation. Available online: https://dokumente.ub.tu-clausthal.de/servlets/mcrfilenodeservlet/clausthal_derivate_00001426/db114668.pdf (accessed on 16 August 2022).

47. Helgers, H.; Hengelbrock, A.; Schmidt, A.; Vetter, F.L.; Juckers, A.; Strube, J. Digital Twins for scFv Production in Escherichia coli. *Processes* **2022**, *10*, 809. [CrossRef]
48. Rosengarten, J.F.; Schatz, S.; Wolf, T.; Barbe, S.; Stitz, J. Components of a HIV-1 vaccine mediate virus-like particle (VLP)-formation and display of envelope proteins exposing broadly neutralizing epitopes. *Virology* **2022**, *568*, 41–48. [CrossRef]
49. Kornecki, M.; Schmidt, A.; Strube, J. PAT as key-enabling technology for QbD in pharmaceutical manufacturing A conceptual review on upstream and downstream processing. *Chim. Oggi* **2018**, *36*, 44–48.
50. Santos, R.M.; Kessler, J.-M.; Salou, P.; Menezes, J.C.; Peinado, A. Monitoring mAb cultivations with in-situ raman spectroscopy: The influence of spectral selectivity on calibration models and industrial use as reliable PAT tool. *Biotechnol. Prog.* **2018**, *34*, 659–670. [CrossRef]
51. Abu-Absi, N.R.; Kenty, B.M.; Cuellar, M.E.; Borys, M.C.; Sakhamuri, S.; Strachan, D.J.; Hausladen, M.C.; Li, Z.J. Real time monitoring of multiple parameters in mammalian cell culture bioreactors using an in-line Raman spectroscopy probe. *Biotechnol. Bioeng.* **2011**, *108*, 1215–1221. [CrossRef] [PubMed]
52. Pereira Aguilar, P.; González-Domínguez, I.; Schneider, T.A.; Gòdia, F.; Cervera, L.; Jungbauer, A. At-line multi-angle light scattering detector for faster process development in enveloped virus-like particle purification. *J. Sep. Sci.* **2019**, *42*, 2640–2649. [CrossRef] [PubMed]
53. Berry, B.; Moretto, J.; Matthews, T.; Smelko, J.; Wiltberger, K. Cross-scale predictive modeling of CHO cell culture growth and metabolites using Raman spectroscopy and multivariate analysis. *Biotechnol. Prog.* **2015**, *31*, 566–577. [CrossRef] [PubMed]
54. Biechele, P.; Busse, C.; Solle, D.; Scheper, T.; Reardon, K. Sensor systems for bioprocess monitoring. *Eng. Life Sci.* **2015**, *15*, 469–488. [CrossRef]
55. Kornecki, M.; Strube, J. Process Analytical Technology for Advanced Process Control in Biologics Manufacturing with the Aid of Macroscopic Kinetic Modeling. *Bioengineering* **2018**, *5*, 25. [CrossRef] [PubMed]