Article

Digital Twin for HIV-Gag VLP Production in HEK293 Cells

Alina Hengelbrock 1, Heribert Helgers 1, Axel Schmidt 1, Florian Lukas Vetter 1, Alex Juckers 1, Jamila Franca Rosengarten 2,3, Jörn Stitz 2 and Jochen Strube 1,*

1 Institute for Separation and Process Technology, Clausthal University of Technology, Leibnizstr. 15, 38678 Clausthal-Zellerfeld, Germany; hengelbrock@itv.tu-clausthal.de (A.H.); helgers@itv.tu-clausthal.de (H.H.); schmidt@itv.tu-clausthal.de (A.S.); vetter@itv.tu-clausthal.de (F.L.V.); juckers@itv.tu-clausthal.de (A.J.)
2 Faculty of Applied Natural Sciences, Technische Hochschule Köln, 51368 Leverkusen, Germany; jamila_franca.rosengarten@th-koeln.de (J.F.R.); joern.stitz@th-koeln.de (J.S.)
3 Institute of Technical Chemistry, Leibniz University Hannover, Callinstraße 5, 30167 Hannover, Germany
* Correspondence: strube@itv.tu-clausthal.de

Abstract: The development and adoption of digital twins (DT) for Quality-by-Design (QbD)-based processes with flexible operating points within a proven acceptable range (PAR) and automation through Advanced Process Control (APC) with Process Analytical Technology (PAT) instead of conventional process execution based on offline analytics and inflexible process set points is one of the great challenges in modern biotechnology. Virus-like particles (VLPs) are part of a line of innovative drug substances (DS). VLPs, especially those based on human immunodeficiency virus (HIV), HIV-1 Gag VLPs, have very high potential as a versatile vaccination platform, allowing for pseudotyping with heterologous envelope proteins, e.g., the S protein of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). As enveloped VLPs, optimal process control with minimal hold times is essential. This study demonstrates, for the first time, the use of a digital twin for the overall production process of HIV-1 Gag VLPs from cultivation, clarification, and purification to lyophilization. The accuracy of the digital twins is in the range of 0.8 to 1.4% in depth filtration (DF) and 4.6 to 5.2% in ultrafiltration/diafiltration (UFDF). The uncertainty due to variability in the model parameter determination is less than 4.5% (DF) and less than 3.8% (UFDF). In the DF, a prediction of the final filter capacity was demonstrated from as low as 5.8% (9mbar) of the final transmembrane pressure (TMP). The scale-up based on DT in chromatography shows optimization potential in productivity up to a factor of 2. The schedule based on DT and PAT for APC has been compared to conventional process control, and hold-time and process duration reductions by a factor of 2 have been achieved. This work lays the foundation for the short-term validation of the DT and PAT for APC in an automated S7 process environment and the conversion from batch to continuous production.

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 (HEK293); human immunodeficiency virus (HIV); virus-like particles (VLPs)

1. Introduction

The human immunodeficiency virus (HIV) is a retrovirus that is responsible for acquired immunodeficiency syndrome (AIDS) [1]. Due to the immunodeficiency that the disease causes, affected people are at greater risk of acquiring other diseases, which can be prevented by immunization through a vaccine [2]. In more than twenty-five years of research, several vaccine candidates have been developed, but they proved to be ineffective [3,4]. Therefore, further vaccines need to be developed [4], and virus-like particles for antigen representation have turned out to be a promising approach [5,6].
VLPs are multiprotein or membrane structures and mimic the organization and structure of real viruses. Unlike viruses, they lack the viral genome. Consequently, VLPs are replication-incompetent particles providing a high safety profile. No virus inactivation is required [7], thus qualifying them as ideal vaccine candidates.

The three viral polyproteins, Gag, Pol, and Env, form HIV particles. These are surrounded by a lipid layer and carry the RNA genomes within the Gag-formed capsid core. In mature and infectious HIV particles or virions, the precursor proteins are proteolytically processed to functional subunits, e.g., the Gag protein is mainly cleaved into its matrix (MA), capsid (CA), and nucleocapsid (NC) subunits. In contrast, VLPs may also be immature HIV-derived particles formed by the uncleaved Gag precursor proteins not maturing at all but still surrounded by a lipid layer of the host cell [8].

Unlike soluble antigens, which must be injected with adjuvants to elicit a protective immune response, VLPs can stimulate a better cellular and humoral immune response [9,10]. VLPs are very efficiently taken up by antigen-presenting cells (APCs) due to their particulate nature and repetitive structure, eliciting both humoral and cellular immune responses [8,11,12].

VLPs can be produced using various expression systems such as bacterial, yeast, insect, mammalian, and plant cells, respectively. Large-scale approaches for the production of HIV-1 Gag VLPs have primarily utilized the baculovirus expression system and insect cell lines [13–16].

Mammalian cells can produce more complex enveloped VLPs such as HIV-1-Gag VLPs, but this is associated with lower productivity [11,17]. Human embryonic kidney 293 cells (HEK293) are particularly well suited for the production of VLPs in mammalian cells. This is due to their good genetic manipulability, their ability to grow in suspension, and the resulting possibility of cultivation at high cell densities [18]. Furthermore, these 293F suspension cultures are already widely used for the production of many virus-based products such as viral vaccines and most viral vectors in the industry [19–22]. This leads to a rapid acceptance of 293 cells in the industry [18].

Few methods for the production of HIV-based VLPs in mammalian cells—particularly for suspension cultures—have been described in the literature [18]. Unlike non-enveloped VLPs, enveloped nanoparticles such as HIV-1-Gag VLPs are very sensitive to shear stress, pH variation, and osmotic pressure, rendering their production and purification especially challenging [8,23].

Quality-by-Design (QbD) methods can be used to establish the relationship between process parameters and the relevant quality characteristics of the product. Furthermore, QbD-based process development is becoming the standard in the biopharmaceutical industry and is also required by regulatory authorities. To ensure the quality target product profile (QTPP), a control strategy needs to be developed as part of QbD-based process development, which requires the spanning of design spaces [24,25]. Design spaces can be defined via validated process models to avoid OOS (out-of-specification) batches [26,27].

Advanced process control can also be realized based on a validated process model developed as a “digital twin” (DT) [28,29]. By applying the holistic QbD approach, consistent product quality can be ensured from development, through piloting to production [27,30]. Process models can be used for the real-time prediction of quality attributes. They allow changes to the process even after submission when optimizing it [31]. To achieve this, digital twins are needed that enable a digital representation of the process.

A digital twin can be divided into five levels, which are shown in Figure 1. The first three stages represent purely digital models whose level of detail increases as the stage progresses. The first stage is a steady-state model that describes the process by means of time-independent mass and energy balances. The steady-state model is used in the first design phase of the process for the initial optimization and calculation procedures [32–35]. The second stage of a digital twin is achieved by extending the steady-state model with accumulation terms and the system dynamics. By deriving all variables of interest over time, the model becomes time-dependent and consequently dynamic. It is used for the iden-
tification of optimal operating conditions, scaling planning, and process control [32,36–38]. The third stage, as the final stage of a digital model, is the validated model, which is validated against process data. This considers phenomena such as inhibition, which increases the number of states. Furthermore, equipment conditions, such as working capacity and hydraulics, are considered. The next stage, the digital shadow, is a validated model, which can be executed in real-time based on automatic inputs via a data link with the physical process [39]. The final stage of a digital twin involves model-based control. It enables the online optimization of the process. Through the digitization infrastructure, control structures can be implemented based on the model-based predictions [32].

Figure 1. Levels of a digital twin, starting from a steady-state model to a dynamic model, a validated model, and a digital shadow to a model-based control.

The design space required for QbD-based process development, in which consistent quality can be ensured, can be spanned by experiments or rigorous process models. The workflow of a QbD process is shown in Figure 2 [40,41]. This first includes the definition of a quality target product profile. The QTPPs affect the bioavailability, strength of action, and stability of the drug. Consequently, there is a relationship between the QTPP and the quality, safety, and efficacy of the drug [42]. Characteristics that, when controlled within a certain limit, range, or distribution, result in the desired product quality are referred to as critical quality attributes (CQAs) and must be defined and classified [43–45]. They are the basis for further process development and, as new knowledge is gained about the process or product, must be dynamically adjusted. Experimentation and risk management can be used to determine the CQAs. Risk management includes risk assessment, which should be performed at the beginning of process development [45,46]. Risk analysis is followed by the definition of the design space, which is traditionally done through experiments. Design-of-experiments (DoE) methods are traditionally used to reduce the experimental effort [47]. The final steps of QbD-based process development include the development of a Process Analytical Technology (PAT)-assisted control strategy and continuous improvement, which are considered in detail for the production of HIV-1 gag VLPs by Helgers et al. [23,45,48].
Figure 2. Workflow of model validation based on a QbD-oriented approach. In the first step, the QTPPs are defined. Subsequently, the CQAs are defined and a risk assessment of the influence of various process parameters on the CQAs is carried out. The risk assessment results in a design space for the process parameters to be investigated, which can be examined either via experiments or by means of a rigorous process model. Based on the results, a control strategy is defined, which can be continuously compared online via PAT with the actual state of the system. The strict implementation of this strategy allows continuous process optimization.

By using predictive process models, a design space can be defined in a resource-efficient way and a quantitatively defined and knowledge-based process optimum can be determined. Since they are derived from physical chemistry, they do not lose their validity when the boundaries of the design space are exceeded, which reduces the experimental effort. Thus, process design becomes possible not only on a purely empirical basis but also enables model- and data-based process evaluation [45]. Furthermore, digital twins in combination with a PAT control strategy are the key enabler for automated, continuous biomanufacturing. Thus, the combination of both reduces time-to-market. In addition, this leads to a reduction in the workload of personnel due to lower cleaning and sterilization efforts. A process developed according to QbD guidelines allows better utilization of cost-intensive raw materials such as cultivation medium or feed and reduces the logistical effort [23,28,49–51].

The use of predictive models allows us to define a design space in a resource-efficient way and to determine a quantitatively defined and knowledge-based process optimum. They facilitate a reduction in experimental effort since they are derived from physics-chemistry and thus do not lose their validity when the boundaries of the design space are exceeded. Thus, process design becomes possible not only on a purely empirical basis but also leads to a model- and data-based process evaluation. A prerequisite for the possibility of using predictive models is that they are revealed to be at least as accurate and precise as the experiments they are intended to replace [45]. This can be investigated, for example, using Monte Carlo simulations by comparing the results of simulation studies with experimental data [45].
Therefore, the aim of this study was to develop a digital twin for the process developed and optimized by Helgers et al. [23] to produce HI-VLPs in HEK293 cells as illustrated in Figure 3. To this end, simulation studies were performed for cultivation, cell separation by depth filtration, product purification, and concentration by ultra- and diafiltration (UF/DF) anion exchange chromatography (AEX). In addition, in order to determine their suitability as a digital twin, the models were investigated in terms of accuracy compared to the experimental data and their precision. In addition, the modeling of lyophilization is performed as the formulation step.

Figure 3. Process flow diagram showing the production and subsequent three-step purification of HIV-gag VLPs. Capture is achieved by depth filtration, ultrafiltration/diafiltration (UF/DF) is used for concentration and initial purification, anion exchange chromatography is used for further purification and concentration, and lyophilization is used for the formulation of HIV-gag VLPs.

2. Materials and Methods

The materials and methods used for the simulation of cultivation, depth filtration, ultrafiltration, anion exchange chromatography, and lyophilization are presented below. All simulations were performed using Aspen Custom Modeler (AspenTechnology Inc., Bedford, MA, USA). The materials and methods used to generate the experimental data are described in detail in [23].

2.1. Fed-Batch Cultivation

Fed-batch cultivations were described using a Monod-based model where glucose and glutamine were included as the carbon and nitrogen sources and lactate and ammonium were considered as the major byproducts of glucose and glutamine metabolism, respectively.

The concentration of viable cells, \( X_V \), was described by:

\[
\frac{dX_V}{dt} = \frac{\left( V_{in}X_{V,in} - V_{out}X_{V,out} \right)}{V} + (u - u_d)X_V
\]

where \( \mu \) is the growth rate and \( \mu_d \) is the death rate.

The growth rate was described by a multiplicative Monod equation considering the aforementioned substrates and metabolites:

\[
u = \mu_{max} \cdot \frac{[GLC]}{K_{glc} + [GLC]} \cdot \frac{[GLN]}{K_{gln} + [GLN]} \cdot \frac{K_{lac}}{K_{lac} + [LAC]} \cdot \frac{K_{lamm}}{K_{lamm} + [AMM]}
\]

Lactate and ammonium were considered growth-inhibiting at high concentrations. Cell death was considered to be dependent on lactate and ammonium accumulation.

\[
u_d = k_d \cdot \frac{[LAC]}{K_{Dlac} + [LAC]} \cdot \frac{[AMM]}{K_{Damm} + [AMM]}
\]
The glucose concentration was described by the following equation, in which \( m_{\text{glc}} \) is the maintenance coefficient of glucose:

\[
\frac{d[\text{GLC}]}{dt} = \left( \frac{V_{\text{in}} \cdot [\text{GLC}]_{\text{in}} - V_{\text{out}} \cdot [\text{GLC}]_{\text{out}}}{V} \right) - \left( \frac{u - u_d}{Y_{XV/\text{glc}}} + m_{\text{glc}} \right) \cdot X_V
\]  

(4)

The glutamine concentration was described by:

\[
\frac{d[\text{GLN}]}{dt} = \left( \frac{V_{\text{in}} \cdot [\text{GLN}]_{\text{in}} - V_{\text{out}} \cdot [\text{GLN}]_{\text{out}}}{V} \right) - \left( \frac{u - u_d}{Y_{XV/\text{gln}}} + m_{\text{gln}} \right) \cdot X_V
\]  

(5)

where \( m_{\text{gln}} \) is the maintenance coefficient of glutamine, which is described by the following equation:

\[
m_{\text{gln}} = \frac{a_1 \cdot [\text{GLN}]}{a_2 + [\text{GLN}]}
\]  

(6)

Lactate and ammonium formation were considered proportional to glucose and glutamine consumption, respectively:

\[
\frac{d[\text{LAC}]}{dt} = \left( \frac{V_{\text{in}} \cdot [\text{LAC}]_{\text{in}} - V_{\text{out}} \cdot [\text{LAC}]_{\text{out}}}{V} \right) + Y_{\text{lac/glC}} \left( \frac{u - u_d}{Y_{XV/\text{glC}}} \right) \cdot X_V
\]  

(7)

\[
\frac{d[\text{AMM}]}{dt} = \left( \frac{V_{\text{in}} \cdot [\text{AMM}]_{\text{in}} - V_{\text{out}} \cdot [\text{AMM}]_{\text{out}}}{V} \right) + Y_{\text{amm/glN}} \left( \frac{u - u_d}{Y_{XV/\text{gln}}} \right) \cdot X_V - r_{\text{amm}} \cdot X_V
\]  

(8)

The product formation was considered to be proportional to the viable cell concentration:

\[
\frac{d[\text{VLP}]}{dt} = \left( \frac{V_{\text{in}} \cdot [\text{VLP}]_{\text{in}} - V_{\text{out}} \cdot [\text{VLP}]_{\text{out}}}{V} \right) + Q_{\text{VLP}} \cdot X_V
\]  

(9)

\[
\frac{dV}{dt} = V_{\text{in}} - V_{\text{out}}.
\]  

(10)

Since the process considered in this study is fed-batch cultivation, the outflowing volume flow is zero.

### 2.2. Depth Filtration

Cell separation and the initial purification of the product were performed using Millistak®+ D0HC filter media (Merck Millipore, Burlington, MA, USA). The direct filtration of FB4 and FB5 as well as the filtration of pre-clarified feed (PC1, PC2) was performed at a constant LMH of 60 L·m⁻²·h⁻¹ and the transmembrane pressure (TMP) was recorded.

In depth filtration, the four main blocking mechanisms of standard, complete, intermediate blocking, and cake filtration primarily occur. In addition, there are models which claim to be able to better match the structure of the filter media [52–54].

The TMP describes the pressure difference between retentate (\(p_{\text{ret}}\)) and permeate (\(p_{\text{perm}}\)):

\[
\text{TMP} = p_{\text{ret}} - p_{\text{perm}}
\]  

(11)

Furthermore, the TMP depends on the flux (\(J\)), the filtration resistance (\(R\)), and the permeate viscosity (\(\eta\)). This dependency can be described with the help of the Darcy–Weisbach equation:

\[
\text{TMP} = J \cdot R \cdot \eta
\]  

(12)

The increase in filtration resistance (\(R\)) due to blocking caused by the deposition of particles on and in the filter material can be described by different models. In the formula, the standard blocking model is shown, which describes the blocking due to pore blocking [52,53]:

\[
\frac{R_0}{R} = \left( 1 - \frac{K_s}{2} \cdot v \right)^2
\]  

(13)
A combination of pore blocking, and surface deposit can be described by the intermediate blocking model \[52,53\]:

\[
\ln \frac{R}{R_0} = K_I \cdot v
\]

(14)

The composite media model developed by Zydny et al. considers that the pores are not ideally cylindrical \[54\]:

\[
\frac{R_0}{R} = (1 - \frac{K_{CM}}{2} \cdot v)^3
\]

(15)

In the equations, \(R_0\) describes the membrane resistance present at the start of depth filtration and \(v\) the membrane area’s specific filtrate volume. \(K_S\) is the blocking constant in the standard blocking model, \(K_I\) the same in the intermediate blocking model, and \(K_{CM}\) the same in the composite media model.

2.3. Ultra- and Diafiltration

Experimental materials, methods, and data have previously been published by Helgers et al. \[23\]. The investigated hollow fiber module consists of 60 fibers (0.5 mm inner diameter, 0.2 m length, 190 cm\(^2\) surface area, 0.05 \(\mu\)m pore size). The process model applied is based on work by Grote et al. \[55\]. Filtration is described by the Darcy–Weisbach equation \[56–58\]:

\[
Jv = \frac{\Delta P}{\eta \times R}
\]

(16)

The three major approaches to model flux decline in tangential-flow ultrafiltration are resistance, gel-concentration, and osmotic-pressure models \[59\]. Since the retentate stream is a suspension of virus-like particles, flux decline is best described by the resistance model, where the total resistance \(R\) in Equation (16) is the sum of the initial membrane resistance \(R_m\) and the boundary-layer resistance \(R_{bl}\), which is computed from experiments \[60,61\]. The transmembrane pressure \(\Delta P\) is defined by Equation (17):

\[
\Delta P = \frac{p_{ret,in} - p_{ret,out}}{2} - p_{per}
\]

(17)

For process simulations, all model parameters are randomly varied inside the experimentally observed determination range to show model prediction precision and accuracy. Based on 31 simulations, the experimental results \[23\] are then compared to the simulation range of prediction.

2.4. Anion-Exchange Chromatography

As the capture process for enveloped virus-like particles (VLP), the use of anion-exchange chromatography (AEX) is widely spread \[62–68\]. There are multiple different approaches explored in the literature, such as the use of membrane adsorbers \[63–65,68\], monolithic columns \[65,67\], and standard chromatographic materials \[65,66\]. Based on this selection, after screening experiments, we decided to use POROS™ GoPure™ (0.5 mm \(\times\) 50 mm) (Thermo Fisher Scientific Inc., Waltham, MA, USA). The buffer systems used consisted of a 50 mM HEPES buffer (VWR Chemicals LLC, Radnor, PA, USA) at pH = 7.2 for buffer A and the same buffer with 2 M NaCl (EMSURE®; Merck KGaA, Darmstadt, Germany) for buffer B. The method is a four-step gradient with a length of five column volumes (CV) per modifier concentration. The modifier concentration was 0.3, 0.7, 0.9, and 1.2 M NaCl. Regeneration was done for 10 CV with 100 % buffer B.

The digital twin for anion exchange chromatography was based on the lumped pore diffusion model for chromatography in combination with Langmuir adsorption, as depicted in Equations (18)–(20) \[69–71\]:

\[
\varepsilon_{p,i} \frac{\partial c_{p,i}}{\partial t} + (1 - \varepsilon_{p,i}) \frac{\partial q_i}{\partial t} = \frac{6}{d_p} \frac{(1 - \varepsilon_S)}{\varepsilon_S} k_{eff,i} \cdot (c_{i} - c_{p,i})
\]

(18)

\[
k_{eff,i} = \frac{1}{k_f} + \frac{r_p}{d_p}
\]

(19)
\[ q_i = \frac{q_{\text{max},i} K_{eq,i} c_i}{1 + K_{eq,i} c_i} \] (20)

In the mass balance of the fluid phase (Equation (18)), \( \epsilon_{p,i} \) is the inner porosity for the component, \( c_{p,i} \) is the concentration of component \( i \) in the pores, \( t \) is time, \( q_i \) is the concentration on the resin, \( d_p \) is the mean diameter of the resin particle, \( \epsilon_S \) is the voidage, \( k_{eff,i} \) is the effective mass transport coefficient, and \( c_i \) is the concentration in the continuous phase.

The mass transfer coefficient \( k_{eff,i} \) is given by Equation (19). Here, \( k_{f,j} \) is the film mass transfer coefficient, \( r_p \) is the particle radius, and \( D_{p,i} \) is the pore diffusion coefficient. For the pore diffusion coefficient and the film mass transfer coefficient, the correlations of Carta and Wilson are available and used widely [72,73].

In the Langmuir isotherm, \( q_{\text{max},i} \) is the maximum loading capacity of the component and \( K_{eq,i} \) is the Langmuir coefficient of the component. \( K_{eq,i} \) and \( q_{\text{max},i} \) are related by the Henry coefficient, \( H_i \); see Equation (4) [69]. Salt influence can be described by Equations (22) and (23), defining \( a_1, a_2, b_1, \) and \( b_2 \) as correlation coefficients [74]:

\[ q_{\text{max},i} K_{eq,i} = H_i \] (21)
\[ q_{\text{max},i} = b_1 c_{p,1} + b_2 \] (22)
\[ H_i = a_1 c_{p,1}^{a_2} \] (23)

To more accurately describe the pore diffusion and evaluate a chromatography digital twin for the polishing step, we employed the general rate model for chromatography, as given in Equation (24) [69].

\[ \epsilon_{p,i} \frac{\partial c_{p,i}}{\partial t} + (1 - \epsilon_{p,i}) * \frac{\partial q_i}{\partial t} = 1 \frac{\partial}{\partial r} \left[ r^2 \left( \epsilon_{p,i} D_{p,i} \frac{\partial c_{p,i}}{\partial r} + (1 - \epsilon_{p,i}) D_{s,i} \frac{\partial q_i}{\partial r} \right) \right] \] (24)

Parameters in this study were taken from the supplier’s documents. The Langmuir parameters are determined using the experimental data obtained from the above-described experiment.

2.5. Lyophilization

The lyophilization process is modeled by a one-dimensional sorption sublimation model introduced by [75]. Here, the exact derivation of the proposed model is shown. It calculates the time-dependent product temperature and the residual moisture during the lyophilization process.

Conduction through the frozen layer is the main heat transfer mechanism during primary drying. Ice sublimes at the sublimation interface and flows through the porous dried layer. The energy balance is written as:

\[ \rho_{\text{Product}} \cdot c_{p,\text{apparent}} \cdot \frac{\partial T}{\partial t} = \lambda \cdot \frac{\partial^2 T}{\partial x^2} \] (25)

\( \rho_{\text{Product}} \) describes the density of the product, \( c_{p,\text{apparent}} \) is the apparent heat capacity, \( T \) is the product temperature inside the vial, and \( \lambda \) is the heat conductivity. The phase change at the sublimation interface is implemented by an apparent heat capacity.

In the overall mass, the balance of water, ice, and the dried product is considered. Since the phase change of water is much faster than convection, it controls the overall transport rate:

\[ \frac{\partial m_W}{\partial t} = \left( \rho_{W,g} \cdot \frac{\Delta p}{\eta_W \cdot K} \cdot A_{\text{vial}} \right) \] (26)

where \( m_W \) is the overall water mass, \( \rho_{W,g} \) is the density of the water vapor, \( \Delta p \) is the pressure difference, \( \eta_W \) is the dynamic vapor viscosity of water, \( K \) is the hydraulic flow resistance, and \( A_{\text{vial}} \) is the cross-sectional area of the vial. Heat and mass transfer are coupled by the sublimation enthalpy [76].

The next step in lyophilization is secondary drying. Desorption is the main transport mechanism and is modeled by an Arrhenius approach because of its temperature dependency. The mass balance of the bound water during secondary drying can be written as:
\[ \frac{dw_{bw}}{dt} = - \exp \left( - \frac{\Delta h_{subl}}{R \cdot T_{product}} \right) \cdot \alpha_w \cdot \left( w_{bw} - w_{bw,eq} \right) \]  

(27)

where \( w_{bw} \) is the mass fraction of the bound water in the dried product, \( \Delta h_{subl} \) is the sublimation enthalpy, \( R \) is the gas constant, \( \alpha_w \) is the water activity, and \( w_{bw,eq} \) is the mass share of bound water at equilibrium.

3. Results and Discussion

3.1. Cultivation of HIV-Gag Producing HEK293 Cells

Figure 4 shows the course of the live cell count (a), the glucose and lactate concentration (b), the glutamine concentration (c), and the relative product concentration (d). The experimentally determined concentrations can be described well using the Monod-based model. The largest deviations in the predictions occur at the end of cultivation for the live cell count and the glutamine concentration. The experimental and simulated concentration curves for glucose and lactate are well matched over the entire process time. Product formation was experimentally only detectable from day nine onwards.

![Figure 4](image_url)

Figure 4. Concentration profiles of viable cell concentration (a), glucose (GLC) and lactate (LAC) concentration (b), glutamine concentration (c), and relative product concentration (d). Experimental results are shown as data points; simulation results are shown as lines. In plot (d), experimental data were only available in the last five days of cultivation since the VLP concentration was below the detection limit before then.

3.2. Harvest via Depth Filtration

The blocking constants required for the modeling were determined using linear regression and are shown as an example for FB4 (direct filtration, without pre-clarification)
in Figure 5. In addition to standard blocking (Figure 5a), intermediate blocking (Figure 5b), and the composite media model (Figure 5c), complete blocking and cake formation are listed in the literature as further main blocking mechanisms. However, these mechanisms are not able to reproduce the experimental processes presented here with sufficient accuracy, which has previously been discussed by [23].

**Figure 5.** Characterization of blocking mechanism during constant flux direct (not pre-clarified) depth filtration of FB4 (Millistak® + D0HC) by linear regression. Regression results of standard (a), intermediate (b), and composite media models (c). Black squares represent the experimental pressure values.

The membrane resistance is calculated from the initially present pressure difference via equation. This results in the parameters for the modeling shown in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>$R_0$ ($\text{m}^{-1}$)</th>
<th>$K_S$ ($\text{m}^{-1}$)</th>
<th>$K_I$ ($\text{m}^{-1}$)</th>
<th>$K_{CM}$ ($\text{m}^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FB4</td>
<td>$2.4 \times 10^{11}$</td>
<td>$33.7 \pm 0.01$</td>
<td>$64.2 \pm 0.13$</td>
<td>$27.5 \pm 0.02$</td>
</tr>
<tr>
<td>FB5</td>
<td>$2.3 \times 10^{11}$</td>
<td>$27.9 \pm 0.01$</td>
<td>$53.5 \pm 0.12$</td>
<td>$22.9 \pm 0.02$</td>
</tr>
<tr>
<td>PC1</td>
<td>$6.0 \times 10^{10}$</td>
<td>$7.0 \pm 0.001$</td>
<td>$18.8 \pm 0.01$</td>
<td>$6.3 \pm 0.001$</td>
</tr>
<tr>
<td>PC2</td>
<td>$6.0 \times 10^{10}$</td>
<td>$10.2 \pm 0.003$</td>
<td>$30.2 \pm 0.01$</td>
<td>$9.3 \pm 0.004$</td>
</tr>
</tbody>
</table>

To describe the accuracy and precision of the model, 30 Monte Carlo simulations were performed, combining the parameter errors in a normally distributed and random manner. This obtained a $t$-test confidence interval of 94.97% with a certainty of 99%. The error of
the viscosity, as well as the LMH, were assumed to be 5% each, whereas the error of the membrane resistance and the blocking constants were determined from the experiments.

The experimentally determined pressure curve and the simulation results of the Monte Carlo studies are shown in Figure 6. The harvesting by direct filtration of FB4 (Figure 6a) and FB5 (Figure 6b) can be reproduced well with a deviation of 0.8% (FB4) and 1.4% (FB5) of the achieved filter capacity at the experimentally achieved maximum transmembrane pressure of 1.6 bar (FB4) and 1.9 bar (FB5), respectively, using the standard blocking model. The intermediate model shows the largest deviation with 13.3% (FB4) and 17.0% (FB5).

Figure 6. Increase in pressure over filter capacity during constant flux direct (not pre-clarified) depth filtration of FB4 (a) and FB5 (b) as well as of pre-clarified feed PC1 (c) and PC2 (d). Experimental pressure values (black squares) are compared with the modeling of standard blocking (red lines), intermediate blocking (blue lines), and composite media models (green lines). The dashed lines indicate the maximum/minimum deviation resulting from Monte Carlo simulation studies.

For a pre-clarified feed (PC1, Figure 6c, and PC2, Figure 6d), the smallest deviations of 1.3% (PC1) and 1.4% (PC2) of the maximum filter capacity achieved in the experiment can be achieved using the standard blocking model. However, this deviates most strongly from the experimental results at lower transmembrane bridges, which are decisive for the prediction of the filtration profile in the application as a digital twin. The intermediate blocking model fits this course best and is still sufficiently accurate with a maximum deviation of 4.5% (PC1) of the filter capacity reached at the end.
For both the depth filtration of pre-clarified feed and direct filtration, the modeled pressure curve with the composite media model lies between the standard and the intermediate blocking model. The precision of the models is very accurate, with a deviation of 1.2% (standard) and 1.3% (composite media). With 4.5% the intermediate blocking model shows the largest deviation. Typical experimental reproducibilities are around 5%. The simulation results deviating by less than 4.5% can be assumed as sufficiently precise model predictions.

Figure 7 shows the data points used for the regression to determine the blocking constant and the modeling via the standard blocking model resulting from the parameter determination. Already, from 5.8% of the maximum TMP achieved, the further development of the process can be predicted.

### 3.3. Intermediate Purification and Concentration via Ultra-/Diafiltration

After clarification by depth filtration, volume reduction and the first separation of side components is performed by ultrafiltration/diafiltration. This process step can be divided into two phases. First, a concentration step is performed to reduce the subsequently needed exchange buffer volume as well as process time. Second, buffer exchange and partial purification are achieved by diafiltration with seven diafiltration volumes (DV). Process model predictions are shown in Figure 8 (concentration factor 3) and Figure 9 (concentration factor 7). The digital twin prediction for experiments aligns with the experimental results. The initial strong decline in flux, observed in all experiments, is well described as the slower decrease in flux during the increase in the boundary layer resistance. The model precision, described by the $t$-test derived confidence interval (alpha = 99%), is between 2.1% and 3.8% for concentration factor 3 runs and between 2.4% and 3.3% for concentration factor 7 runs.

### 3.4. Anion-Exchange Chromatography

To model a batch purification with a 180 mL injection, we firstly needed to determine the isotherm parameters as described in Section 2.4. The results of the parameter determination are given in Figure 10a. The experimental results are sufficiently calculated by the digital twin and, as such, enable the scale-up. The laboratory process, employing a 5 mL injection volume, reaches a productivity of $8.87 \times 10^{12}$ particles per mL of column volume and day. A scale-up using up to 180 mL injection volume would require a 35 mL resin
volume with a diameter of 30 mm and a length of 50 mm, which was set to be the same as the column used in the experiment.

![Figure 8](image)

**Figure 8.** Flux $J_v$ over time for experimental (squares) as well as digital twin prediction (red area) for ultrafiltration/diafiltration experiments and concentration factor 3. (a) Experiment 2 (EXP2): 1.5 bar transmembrane pressure (TMP), $3738 \text{ s}^{-1}$ shear rate. (b) Experiment 3 (EXP3): 1.5 bar TMP, $1249 \text{ s}^{-1}$ shear rate. (c) Experiment 7 (EXP7): 0.5 bar TMP, $3738 \text{ s}^{-1}$ shear rate.

To optimize this process, a high salt wash, instead of the first two gradient steps, could be employed. The chromatogram of such a potentially optimized and scaled-up process is depicted in Figure 10b. This scale-up and optimization offer an increase in productivity to $1.89 \times 10^{13}$, which corresponds to a roughly two-fold increase.

### 3.5. Formulation via Lyophilization

In this study, 167 2R vials are filled with 600 µL VLP solution and must be freeze-dried. This scale is manageable with an Epsilon 2-4 LSC-plus [77]. The main excipient is trehalose. Since the main excipient and the equipment scale are comparable, the two model parameters, vial heat transfer and hydraulic flow resistance, are adopted from the literature [75]. The edge vials receive a higher heat input during lyophilization due to radiation [78]. This so-called edge effect leads to a drying batch heterogeneity. Edge vials receive the highest heat input whereas the center vial receives heat mainly due to conduction. It is important both to keep the product temperature for each vial below the collapse temperature and to remove all ice during primary drying.

The lyophilization process for different vials is shown in Figure 11.
Figure 9. Flux Jv over time for experimental (squares) as well as digital twin prediction (red area) for ultrafiltration/diafiltration experiments and concentration factor 7. (a) Experiment 1 (EXP1): 1.5 bar transmembrane pressure (TMP), 3738 s\(^{-1}\) shear rate. (b) Experiment 5 (EXP5): 0.5 bar TMP, 3738 s\(^{-1}\) shear rate. (c) Experiment 6 (EXP6): 0.5 bar TMP, 1249 s\(^{-1}\) shear rate. (d) Experiment 8 (EXP8): 1.5 bar TMP, 1249 s\(^{-1}\) shear rate.

The freeze-drying protocol has been adopted from the literature [79]. During primary drying, the shelf temperature is raised from \(-40\,^\circ\text{C}\) to \(-5\,^\circ\text{C}\) and a chamber pressure of 0.045 mbar is set. During secondary drying, the shelf temperature is increased to 40 \(^\circ\text{C}\) and the pressure is lowered to 0.007 mbar. The initial product temperatures lie at \(-40\,^\circ\text{C}\) and gradually approach the shelf temperature. The higher heat input of the edge vial leads to an increased product temperature compared to the center vial. The primary drying process of the edge vial is completed in 12.5 h, while the center vial needs an additional 3 h. The endpoint of the primary drying [80] can be determined as the point where the product and shelf temperatures are the same. The primary drying process can be considered completed and the secondary drying begins. The product temperature of the hottest vial is near the shelf temperature because bound water is already desorbed in primary drying. The final residual moisture of the vials is about <2%.
4. Discussion

With the digital twins presented in Section 3, it is possible to reproduce the physical process in real-time and, in combination with PAT, to predict the further course of the process and readjust it based on the control strategy within the design space, as illustrated in Figure 12.

During cultivation, the instantaneous VLP, glucose, lactate, and other minor component concentrations can be determined by using modern PAT detectors (especially Raman, FTIR, and MALS/DLS) [48,81–83] and the corresponding modeling [81]. The digital twin can use this process information to adjust feeding and other process parameters, such as gassing, as needed and provide important information for downstream, such as optimal harvest time, total cell density TCD at harvest time, and product concentration, purity, and expected size distribution.
Figure 12. Process flowsheet including APC by digital twins and optimized PAT. Critical Process Parameter (CPP), Well-Controlled Critical Process Parameter (WC-CPP), Key Process Attribute (KPA), particle size distribution (PSD), Viable Cell Density (VCD), transmembrane pressure (TMP), virus-like particle (VLP).
Product concentration and size distribution determine the blocking mechanism and the associated maximum filter capacity [84,85]. Knowledge of this variable makes it possible to ensure that a sufficient filter area is available while avoiding oversizing. Product concentration is also an important process parameter for UF/DF to compensate for variations prior to AEX to a fixed target concentration and thus loading volume.

By using the digital twin during depth filtration, the differential pressure measurement can consistently predict the maximum filter capacity currently expected. This makes it possible to avoid holding times due to filter replacement or yield loss if the entire cultivation broth cannot be filtered due to fixed scheduling [86].

The digital twin in the UF/DF makes it possible to predict the further development of the filter resistance via the measured permeate flow and thus determine the expected endpoint [61]. Furthermore, the shear rate can be adjusted to ensure the timely loading of the AEX with the desired product concentration. This makes it possible to set the shear rate only as high as necessary and thus to work as gently as possible for the product [87].

In the context of QbD-based process development, shear stress as a Critical Process Parameter (CPP) as part of the control strategy would ensure that the process is operating within the proven acceptable range (PAR) [41,88]. Digital twins, in combination with PAT—for example, FTIR—create the flexibility to set the exchange volumes in the UF/DF based on the actual desalination and purity levels. This flexibility cannot be achieved with conventional process development and offline analytics, where fixed exchange volumes are set. Achieving optimal product purity or saving exchange buffers can therefore only be enabled by a digital twin and QbD-based process design.

For the AEX step, real-time product analysis using MALS/DLS in combination with the digital twin offers the possibility to dynamically adjust the cut points for the product fraction to achieve the optimum between yield and purity/concentration.

The digital twin also offers advantages in terms of critical raw materials, equipment technology, and process control [89]. In combination with PAT strategies, it also provides the basis for the transition from batch to continuous operation [90].

The most cost-intensive raw materials in the process, such as feed-in cultivation, can be dosed in a tailored ratio. This enables potential savings that directly reduce the cost of goods, making the process more profitable or more affordable for the healthcare system [91]. The independence from supply bottlenecks increases and a stable market supply is made possible.

In terms of choices of equipment technology, the use of digital twins in combination with online CIP/SIP measurement techniques (such as real-time total organic carbon, TOC, and bioburden, as well as conductivity) [92] offers the freedom to move away from the commonly used single-use technology. This is made possible by automated execution and validation [93]. This eliminates the simple and fast switch of campaigns, which is the main advantage of single-use technology [94]. There are even benefits in terms of sustainability [95]. The recently reported difficulties in the supply of single-use technology can be circumvented [96].

The step from batch-wise to continuous production is made possible last but not least by linking modern PAT-based control strategies with the digital twin. Only this link enables the realization of the main advantages of continuous production, such as real-time release-testing (RTRT) for a shorter time-to-market combined with lower batch failure rates [28,51,97].

The advantages of automated production enabled by PAT for APC and digital twins become particularly clear when this is compared with the conventional process based on offline analyses, as depicted in Figure 13. Ideally, the ongoing prediction of the digital twin enables subsequent process steps to be performed with minimal hold-time. This is secured by real-time PAT [23] in combination with a QbD-based control strategy to ensure operation within the NAR. Design Spaces NAR and PAR are available as part of a QbD-based process development, which was demonstrated for other systems [48,80]. This combination makes it possible for the process presented here to perform the entire downstream within one
working day. In contrast, the conventional process flow requires time-consuming offline analytics, resulting in two working days for the downstream process. This additionally means longer holding times for the product, which increase the risk of quality degradation for sensitive products such as enveloped VLPs [98].

Figure 13. Optimized process schedule enabled by PAT for APC and DT (a) vs. conventional process schedule based on offline analytics (b).

5. Conclusions

For the production and purification of HIV-1 Gag VLPs in HEK293 cells, this study demonstrates the applicability of digital twins in the context of QbD-based process development.

Based on the digital twin, optimized control allows for an optimized feeding strategy as well as the prediction of important KPA such as yield and VLP concentration as well as particle size distribution by MALS/DLS.

In depth filtration, it was shown that the early detection of real filter capacity is possible from as low as 5.8% (9 mbar) of the maximum TMP. The predicted capacity at this point is 50 L/m² and deviates only 1% from the real capacity of 49.5 L/m².

In UF/DF, the predicted product concentration can be used to determine the necessary process time. The maximum deviation in the prediction is less than 3.8%.

This allows the AEX to be loaded with a constant VLP concentration. The digital twin in the AEX can be used for a scale-up, which could increase productivity by up to a factor of 2.

Another important benefit of digital twins in combination with APC and PAT is the optimization of the process schedule. By avoiding hold times between process steps, the downstream can be fully started within one working day instead of two. This reduces the hold time until lyophilization by a factor of 2.

These advantages in raw materials, equipment technology, and the move to continuous production are made possible by the use of digital twins for APC with PAT.

Design Spaces NOR and PAR are available as part of QbD-based process development. Through real-time PAT measurement, the extension of the PAT portfolio with MALS/DLS in combination with digital twins for APC enables the automation of the process up to RTRT.

In summary, this study shows the applicability of the digital twins as process models with a maximum deviation of 4.5% from the experimental data. In combination with real-time PAT and a QbD-based control strategy, it is possible to improve the process schedule by a factor of two. The next steps in 2022 are the validation of the shown optimization potentials by linking PAT and the digital twins as well as the conversion from batch to continuous production in a Siemens S7 process environment.
Author Contributions: Conceptualization, J.S. (Jochen Strube); software, process, and analytics experiments, A.H., H.H., A.S., F.L.V. and A.J.; VLP cell line development, J.F.R. and J.S. (Jörn Stitz); writing—original draft preparation, A.H., H.H., A.S., F.L.V. and A.J.; writing—review and editing, A.H., H.H., A.S., F.L.V., A.J., J.F.R., J.S. (Jochen Strube) and J.S. (Jörn Stitz); supervision, J.S. (Jochen Strube) and J.S. (Jörn Stitz); project administration, J.S (Jochen Strube). All authors have read and agreed to the published version of the manuscript.

Funding: A part of this work was supported by a grant from the German Federal Ministry of Education and Research, funding program Forschung an Fachhochschulen, contract number 13FH767IA6 to J.S. (Jörn Stitz). The authors acknowledge financial support from the Open Access Publishing Fund of the Clausthal University of Technology.

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.

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

References

1. Iqbal, N.; Karaca, Y. Complex Fractional-Order HIV Diffusion Model Based on Amplitude Equations with Turing Patterns and Turing Instability. Fractals 2021, 29, 2140013. [CrossRef]
2. Yasmin, H. Effect of vaccination on non-integer dynamics of pneumococcal pneumonia infection. Chaos Solitons Fractals 2022, 158, 112049. [CrossRef]


66. Pereira Aguilar, P.; Schneider, T.A.; Wetter, V.; Maresch, D.; Ling, W.L.; Tover, A.; Steppert, P.; Jungbauer, A. Polymer-grafted chromatography media for the purification of enveloped virus-like particles, exemplified with HIV-1 gag VLP. *Vaccine* 2019, 37, 7070–7080. [CrossRef]  


85. Taylor, N.; Ma, W.; Kristopeit, A.; Wang, S.-C.; Zydney, A.L. Enhancing the performance of sterile filtration for viral vaccines and model nanoparticles using an appropriate prefILTER. *J. Membr. Sci.* **2022**, *620*, 120264. [CrossRef]


