International Journal of Pharmaceutics 686 (2025) 126301

Contents lists available at ScienceDirect

International Journal of Pharmaceutics

journal homepage: www.elsevier.com/locate/ijpharm

# End-to-end digital twin software for continuous mRNA manufacturing

Mohammed Aatif Shahab *, Nathan Merica Stover[1], Hasan Al-Mahayni[1], Prakitr Srisuma[1], Jie Wang[1], Sunkyu Shin[1], Anish Dighe[1], Bernhardt L. Trout, Allan S. Myerson, Richard D. Braatz *

*Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA, 02139, USA*

ARTICLE INFO

Dataset link: https://github.com/aatifshahab/mRNAdigitalTwin-public

*Keywords:*
Digital twin
mRNA
Continuous biomanufacturing
Process simulation
Open-source software

ABSTRACT

Messenger RNA (mRNA) vaccines and therapeutics have received intense interest due to their rapid development potential and broad therapeutic applications. However, traditional batch-based manufacturing of these therapeutics faces scalability and consistency challenges. In this work, we present a novel digital twin software specifically tailored for *end-to-end* continuous mRNA production. The software integrates all key unit operations – In Vitro Transcription (IVT), Tangential Flow Filtration (TFF), Continuous Chromatography (CCTC), Lipid Nanoparticle (LNP) formulation, and Freeze-Drying – within a modular, plug-and-play architecture. We have developed a user-friendly graphical interface that enables process engineers to add, remove, and rearrange modules without coding, while each operation is underpinned by mechanistic models. A built-in database system stores the inputs and outputs of each run, facilitating result retrieval and comparisons. Through an illustrative example involving a representative mRNA production scenario, we demonstrate the seamless simulation of all five operations in silico, generating time series and 3D plots that reveal both process behavior and final product characteristics. Additionally, we used Morris-based global sensitivity analysis for each unit to reveal critical input factors and guide process optimization and model validation. Hence, the software can be used to accelerate process optimization, promotes Quality-by-Design (QbD) principles, and provides a platform for future closed-loop control and operator training. This work supports fully integrated, data-driven continuous manufacturing for mRNA therapeutics, with the potential to enhance scalability, reliability, and responsiveness in the face of global healthcare demands.

## 1. Introduction

The COVID-19 pandemic highlighted critical weaknesses in current biopharmaceutical manufacturing practices, especially for novel therapeutics such as mRNA vaccines. Despite promising laboratory results, many vaccine candidates struggled to transition into robust industrial-scale production, resulting in shortages when demand was at its peak. One key reason for these bottlenecks is the traditional reliance on batch processing methods, which are time-consuming, expensive, and challenging to scale due to limited flexibility and process complexities.

A leading approach to overcoming these challenges is the shift toward continuous manufacturing (Coffman et al., 2021), which offers greater flexibility and efficiency than traditional batch processes. However, continuous processing introduces new complexities, including the need for real-time monitoring, precise control of critical parameters, and rapid adjustments to maintain consistent product quality. To address these, industry practitioners are increasingly adopting Quality by Design (QbD) principles, which emphasize a science- and risk-based approach to understand and control manufacturing processes (International Conference on Harmonisation, 2009; Fisher et al., 2016). These principles help define critical quality attributes (CQAs), critical process parameters (CPPs), and design spaces.

Implementing QbD in continuous biomanufacturing is greatly facilitated by digital twins (Leng et al., 2021; Park et al., 2021). A digital twin (DT) is commonly defined as "a virtual representation of a physical system (and its associated environment and processes) that is updated through the exchange of information between the physical and virtual systems" (Grieves, 2014). The benefits of DTs in biomanufacturing are twofold. First, during early-stage development, researchers can use these virtual environments to identify and optimize inputs and operating conditions, reducing the number of physical experiments and accelerating the transition from lab-scale to pilot-scale. Second, once a process is operational at an industrial scale, DTs enable real-time process monitoring and predictive control. They help forecast the

* Corresponding authors.
*E-mail addresses:* moha0095@mit.edu (M.A. Shahab), braatz@mit.edu (R.D. Braatz).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ijpharm.2025.126301
Received 17 August 2025; Received in revised form 13 October 2025; Accepted 20 October 2025
Available online 22 October 2025

future state of the process, identify deviations, and correct them before they impact product quality and yield. This capability directly supports QbD paradigms by enabling continuous improvement throughout the product lifecycle.

These advantages have already begun to manifest in mRNA-based therapeutics manufacturing, across both upstream and downstream processes. Model-based approaches – mechanistic, data-driven, and hybrid models – have been used to capture the dynamics of these workflows. For instance, mechanistic models have been developed to predict mRNA yield during in vitro transcription (IVT), considering factors such as the nucleation and growth of inhibitory byproducts such as magnesium pyrophosphate (Akama et al., 2012; Helgers et al., 2021; Stover et al., 2024). Optimization of continuous IVT processes using these models has demonstrated up to 55% higher mRNA yield and 33% lower truncated mRNA formation (Hengelbrock et al., 2024). In the downstream space, DTs have similarly been used to optimize purification steps, buffer exchanges, and formulation (Schmidt et al., 2021; Messerian et al., 2022).

Despite these advances, there is limited work on truly end-to-end DT platforms that integrate all major unit operations and provide a flexible software interface for configuration and simulation. This paper presents such a platform, designed for continuous mRNA manufacturing. Our contributions include the integration of multiple mechanistic models, a modular software architecture that supports plug-and-play operations, and a graphical user interface (GUI) that allows users to design, run, and visualize the performance of a full mRNA production pipeline without coding.

## 2. Related work

The production of mRNA-based therapeutics starts with an upstream *in vitro* transcription (IVT) reaction, where mRNA is synthesized using nucleotide triphosphates, a linearized DNA template (often plasmid-derived), and an RNA polymerase (commonly T7). After synthesis, downstream purification steps remove enzymes, nucleotides, and other byproducts, resulting in concentrated mRNA. The purified material is then formulated into lipid nanoparticles (LNPs), which protect against degradation and facilitate cellular uptake. Finally, lyophilization can further improve stability and shelf life. In the subsequent sections, we describe each of these unit operations and modeling approaches.

### 2.1. Upstream – in vitro transcription (IVT)

In IVT, a DNA template is transcribed by T7 RNA polymerase in a cell-free system. This approach avoids reliance on living cells and allows operations to be scaled up simply by adjusting reagent volumes and reaction times. Typically, the reaction includes nucleoside triphosphates (NTPs), a suitable promoter sequence on the DNA, and magnesium ions that serve as cofactors for RNA polymerase (Rosa et al., 2021; Whitley et al., 2022). The IVT reaction proceeds through initiation, elongation, and termination phases, which produces a single-stranded mRNA transcript. Inorganic pyrophosphate (PPi) is formed as a byproduct of polymerization and can precipitate if it binds magnesium (Akama et al., 2012; Stover et al., 2024). Pyrophosphatase enzymes are therefore often introduced to degrade PPi and maintain optimal magnesium availability. The IVT reaction may also integrate capping strategies (co-transcriptional or enzymatic) and the addition of a poly(A) tail for enhanced stability and translational efficiency (Rosa et al., 2021). The IVT platform is modular; different mRNA sequences can be synthesized by altering the DNA template. This makes the system highly adaptable for a range of therapeutic targets. IVT can be conducted either in batch mode or with continuous-flow reactors. For instance, Hengelbrock et al. (2024) developed a tubular plug-flow reactor IVT system and showed that segmented continuous operation can significantly increase space–time yield compared to conventional batch processing.

Mechanistic modeling of IVT has evolved to the point of enabling DTs (Hengelbrock et al., 2024). Such models typically capture enzyme kinetics and key side phenomena – PPi precipitation (Akama et al., 2012) and $Mg^{2+}$PPi nucleation/growth with DNA interactions (Stover et al., 2024) – and have been applied to optimize continuous IVT and reduce off-spec species (Hengelbrock et al., 2024; Ziegenhals et al., 2023). For instance, Stover et al. (2024) developed a comprehensive kinetic model that describes the nucleation and growth of $Mg^{2+}$PPi crystals, and their interactions with DNA in solution. This model successfully explained previously unclear observations, including the way pyrophosphatase enhances productivity by degrading pyrophosphate. These models enable predictive control and are key enablers for the DTs of the upstream IVT process.

### 2.2. Downstream processing

After IVT, the reaction mixture contains target mRNA along with residual NTPs, enzymes, template DNA, and potential dsRNA contaminants. To achieve clinical-grade purity, downstream operations are employed to purify the IVT mRNA, which typically include chromatographic techniques and membrane-based filtration.

#### 2.2.1. Chromatography

Chromatography is well-established for mRNA analytical applications and is increasingly being explored for mRNA purification. Various chromatographic modalities have been used, including size exclusion chromatography, anion exchange chromatography, affinity chromatography, core bead chromatography, and their combinations (Schwaminger et al., 2022; Scorza et al., 2021; Von Der Mülbe et al., 2018; Rosa et al., 2021; Zydney, 2016). Among these, affinity chromatography has demonstrated high selectivity and operational robustness in mRNA purification through the use of oligo(dT) resins that selectively bind to the poly(A) tail of mRNA. This enables efficient capture of the target mRNA while allowing impurities to be washed away.

Several continuous configurations have been developed to enhance separation efficiency and operational throughput, including periodic countercurrent chromatography, multicolumn countercurrent solvent gradient purification, and continuous countercurrent tangential chromatography (CCTC) (Zydney, 2016; Dutta et al., 2015; Qu et al., 2024). Unlike batch single-column operation, continuous systems run loading, washing, elution, and regeneration simultaneously across multiple columns, thereby improving recovery and purity while increasing throughput for IVT mRNA (Challener, 2018; Qu et al., 2024). Among these emerging alternatives, CCTC has the unique advantages of true steady-state operation while reducing buffer consumption and eliminating complex valve-switching schemes typically required in multicolumn systems (Dutta et al., 2015; Napadensky et al., 2013; Fedorenko et al., 2020). Its applicability has been demonstrated in the purification of monoclonal antibodies (mAbs) and adeno-associated virus (AAV) vectors for gene therapy (Dutta et al., 2015; ChromaTan, 2024). In the CCTC setup, a resin slurry is continuously circulated through a series of static mixers, which ensure thorough mixing and provide the necessary residence time, and hollow-fiber membrane modules that separate the solid resin from the liquid phase. The resin advances sequentially through key chromatographic steps, while buffers flow in the opposite direction. This countercurrent arrangement achieves continuous multistep equilibration similar to other continuous chromatographic systems, but without reliance on traditional packed-column setups.

Mechanistic chromatography models grounded in mass transfer and binding kinetics are widely used, with mature toolchains for column systems (GoSilico™, Chromulator, CADET) (Guiochon, 2006; Hahn et al., 2023; Gu, 2015; Leweke and von Lieres, 2018). For CCTC, simple lumped models (e.g., irreversible adsorption) enable early-stage sizing and residence-time estimates and have guided yield optimization (Dutta et al., 2015, 2016).

#### 2.2.2. Tangential flow filtration

Tangential flow filtration (TFF), also known as cross-flow ultrafiltration, is a membrane-based separation process widely used in downstream bioprocessing. In TFF, the fluid is circulated parallel to the filter membrane. This cross-flow mechanism minimizes fouling and permits continuous concentration and buffer exchange of the mRNA solution. TFF enables two key operations simultaneously: concentration (retaining the large product molecules while filtering out solvents and small impurities) and diafiltration (buffer exchange by continuously adding fresh buffer while filtrate is removed). However, process performance depends strongly on key variables such as the *trans-membrane pressure* (TMP), the *critical flux* that signals the onset of fouling, and the *diafiltration volume* needed to reach a target buffer exchange (Agrawal et al., 2023). Careful optimization of these parameters has the greatest impact on mRNA yield, processing time, and buffer consumption. An important design factor in TFF is the selection of membrane molecular weight cutoff (MWCO), which determines which components are retained or removed during filtration. For mRNA, ultrafiltration membranes with MWCO in the 100–300 kDa (Behboudi et al., 2025; Nourafkan et al., 2024) range are commonly used.

TFF may appear multiple times in the manufacturing pipeline (Javidanbardan et al., 2025). After chromatography, TFF concentrates dilute mRNA and performs buffer exchange into the formulation buffer; it is also applied upstream to condition IVT reaction mixtures (for instance, exchange to a low-salt binding buffer before chromatography). TFF is likewise employed after the LNP step at laboratory and industrial scales for solvent removal and buffer exchange (Wu et al., 2025). Traditionally, TFF has been performed in batch mode. For instance, batch TFF using a hollow fiber membrane with a molecular weight cutoff of 300 kDa has been shown to be effective in purifying 1025 nt mRNA (Nourafkan et al., 2024). More recently, single-pass TFF (SPTFF) has enabled continuous operation with no recirculation; modules in series reach high concentrations and fast steady state simplifies monitoring (Madsen et al., 2022; Chaubal and Zydney, 2023). Javidanbardan et al. demonstrated SPTFF directly on IVT crude, achieving 10×concentration with 100% mRNA retention over 24 h, substantial removal of proteins and small impurities, no mRNA degradation, and stable transmembrane pressure below the critical flux (Javidanbardan et al., 2025).

The performance of TFF can be described by well-known models of membrane filtration that account for convective transport, diffusive back-transport, and various fouling mechanisms. For mRNA, one key fouling mechanism is adsorption to membrane surfaces (Nourafkan et al., 2024). Studies have applied a classical Hermia fouling model to capture the dynamics of mRNA filtration (Nourafkan et al., 2024). Such models can simulate how filtrate flux declines over time and how product concentration rises in the retentate, given certain membrane and flow parameters. Another study applied a simplified stagnant film model to predict how mRNA concentration builds up along a single-pass TFF module (Fuchs et al., 2023). CFD combined with mass transfer models have been used to model flow and solute transport in TFF modules to optimize channel geometry and minimize dead zones (Kaiser et al., 2022). Data-driven models have also been introduced to support digital twin applications. Jesubalan et al. (2023) integrated a deep neural network with in-line sensor data to predict membrane permeability decline in real time for mAb ultrafiltration. Such approaches are particularly valuable for monitoring fouling and scheduling membrane cleaning in continuous operations, and they could be extended to mRNA TFF systems.

### 2.3. Formulation – lipid nanoparticle (LNP) formulation

LNP formulation involves encapsulating the purified mRNA into lipid-based carriers to create the final drug product form. LNPs are typically nano-sized (~50–150 nm) vesicular structures composed of an ionizable lipid for mRNA complexation and endosomal escape, helper lipids (phospholipids and cholesterol) for structural stability, and PEG-lipids to limit aggregation (Kulkarni et al., 2021, 2018; Byun et al., 2022). Manufacturing proceeds by rapidly mixing an aqueous mRNA stream with an ethanolic lipid stream, where the polarity shift drives self-assembly; microfluidic (e.g., T-junction) and impinging-jet mixers provide the short mixing times needed for consistent particles (Maeki et al., 2022; Devos et al., 2025). Process parameters such as the flow rate ratio of the two phases, total flow rate, lipid composition, and buffer conditions are carefully controlled. After mixing, the solvent (ethanol) is removed and the dispersion is exchanged into the final buffer; this post-mixing solvent removal and buffer exchange are typically performed via tangential flow filtration (TFF) (Wu et al., 2025).

The LNP formulation largely determines the potency of the mRNA drug. The benefits of a well-controlled LNP formulation process include consistent particle size (important for biodistribution and uptake), high mRNA encapsulation efficiency (minimizing wastage of expensive mRNA), and stable composition for shelf-life. Mixing studies show that sub-optimal flow regimes produce larger, more polydisperse particles with poor encapsulation, whereas operating in an optimal regime yields ~100 nm LNPs with nearly all particles containing mRNA. For instance, Ripoll et al. (2022) found that at certain flow rates, micromixers produce larger and more polydisperse LNPs with lower encapsulation efficiency, whereas operating in the optimal mixing regime yielded uniformly small (~100 nm) LNPs with high encapsulation of mRNA.

The formation of LNPs is a complex self-assembly process that involves rapid nucleation and growth of lipid structures upon mixing. Mechanistic modeling of LNP self-assembly is challenging due to the multiscale nature (molecular interactions forming nano-structures within milliseconds of mixing). However, some studies have approached it by modeling the microfluidic mixing environment and the resulting particle size distributions (PSDs). Device-scale computational fluid dynamics (CFD) studies map micromixing regimes (side-by-side vs. highly mixed) that correlate with LNP size and encapsulation and inform mixer designs yielding narrower PSDs (Ripoll et al., 2022; Kimura et al., 2021). At molecular/empirical scales, coarse-grained MD elucidates LNP morphology during self-assembly, while machine-learning on formulation libraries highlights lipid substructures and ratios associated with higher encapsulation and in vivo expression (Wang et al., 2025; Bae et al., 2025).

The CFD-based approaches described above resolve fluid–mixing fields but do not treat the particulate events – nucleation, growth, agglomeration, and breakage – that actually create and transform nanoparticles. To capture how those discrete phenomena shape the evolving size distribution, a *population–balance model* (PBM) can be coupled to the mixing field. A PBM expresses the rate of change of particle number density in each size interval as the sum of *birth* and *death* terms (Ramkrishna, 2000). In LNP self-assembly, birth arises from *nucleation*, where small lipid–mRNA clusters first appear; growth adds mass as additional molecules deposit on existing particles; agglomeration merges two colliding LNPs into a larger one; and breakage fragments oversized particles into smaller ones (Inguva et al., 2025). This allows the PBM to predict how operating variables – mixing intensity, flow-rate ratio, lipid composition, or temperature – translate into the final LNP size distribution, encapsulation efficiency, and stability. However, this PBE framework does not capture non-ideal aggregation phenomena – often referred to as "blob formation" – that arise from transient shear instabilities or local phase separation in microfluidic mixers. These effects can produce large, irregular aggregates that skew the PSD.

### 2.4. Storage and distribution

Lyophilization (freeze-drying) stabilizes biologics that are unstable in liquid form. The mRNA and lipid components are prone to degradation (hydrolysis, oxidation) in aqueous form, and the original COVID-19

mRNA vaccines required ultra-cold storage (−20 °C to −80 °C) with limited shelf-life. Studies have shown that lyophilized mRNA–LNP formulations can remain stable at 2–8 °C or even room temperature for significant periods (Muramatsu et al., 2022). In practice, filled vials are frozen, ice is removed by sublimation under vacuum (primary drying), and residual bound water is removed at a higher temperature (secondary drying). Cryoprotectants and lyoprotectants (for example, sucrose, trehalose) preserve LNP structure during freezing and drying. The process can be run in batch trays or in emerging continuous setups (such as spin freeze-drying) (Lamoot et al., 2023; Meulewaeter et al., 2023). Key parameters such as shelf temperature and chamber pressure directly influence the sublimation rate. Uneven drying, particularly in vials located at the edges or corners of the chamber, further complicates the process (Srisuma et al., 2024).

Consequently, mechanistic models has been developed to optimize primary drying by capturing the dominant heat- and mass-transfer phenomena inside each vial. For instance, Velardi and Barresi introduced a diffusion- and desorption-based framework that treats heat flow through the vial side wall and predicts drying time as a function of shelf temperature, chamber pressure, fill height, and formulation properties (Velardi and Barresi, 2008). Massei and Fissore (Massei and Fissore, 2023) extracted the vial heat-transfer coefficient ($K_v$) and dried-layer resistance ($R_p$) from a single gravimetric test and showed that those parameters translate directly to pilot-scale equipment. Purely conductive–convective descriptions can miss radiative heterogeneity, so Srisuma et al. (2024) introduced a diffuse-gray radiation network that reproduces edge acceleration and guides chamber modifications to homogenize drying. The same radiation-aware formulation extends to continuous spin freeze-dryers (Srisuma et al., 2025).

The literature demonstrates deep mechanistic understanding and even digital twin development for unit operations such as IVT, chromatography, and filtration. Despite the richness of these individual efforts, most software platforms are proprietary, restricted to one or two unit operations, and lack the flexibility to modify process sequences dynamically. To our knowledge, very few – if any – fully open-source platforms exist that integrate both upstream and downstream mechanistic models. No prior work offers a unified software platform for continuous mRNA manufacturing that combines multiple mechanistic models across the entire pipeline, with plug-and-play capability, user-friendly configuration, and end-to-end simulation support.

This work fills that gap by proposing and demonstrating a modular digital twin software that connects models for in vitro transcription, filtration, purification, nanoparticle formation, and lyophilization, all within a web-accessible interface.

## 3. Materials and methods

### 3.1. In vitro transcription (IVT)

We adapted the mechanistic model of the IVT reaction from the batch model of Stover et al. (2024). That model tracks 10 dynamic state variables (DNA, RNA, PPi, NTPs, Mg, phosphate (Pi), and $Mg_2PP_i$ nuclei) and three algebraic constraints that enforce electroneutrality and divalent-ion balances. Reaction rates include (i) transcription initiation/elongation with sequence-length dependence, (ii) nucleation and growth of $Mg_2PP_i$ crystals, (iii) agglomeration of nuclei with DNA, and (iv) pyrophosphatase degradation of $PP_i$. The original ordinary/differential–algebraic system was

$$\frac{dC_{\mathrm{ATP}}}{dt} = -N_A r_{\mathrm{tr}}, \quad \frac{dC_{\mathrm{PP_i}}}{dt} = (N_{\mathrm{all}} - 1) r_{\mathrm{tr}} - r_{\mathrm{precip}} - r_{\mathrm{PPiase}}, \quad \Phi(C) = 0, \tag{1a–c}$$

where $r_{\mathrm{tr}}$ is the sequence-length-dependent transcription rate, $r_{\mathrm{precip}}$ and $r_{\mathrm{PPiase}}$ are the solid-formation and enzymatic degradation rates, and $\Phi$ denotes the charge-balance algebraic equations (full list in Stover et al. (2024)).

To simulate a continuous in-vitro transcription reactor, we recast each mass balance as a continuously stirred tank reactor (CSTR):

$$\frac{dC_i}{dt} = \frac{C_{i,\mathrm{in}} - C_i}{\tau} + r_i(C), \tag{2}$$

with residence time $\tau = V/Q$. For example, the DNA balance becomes

$$\frac{dC_{\mathrm{DNA}}}{dt} = \frac{C_{\mathrm{DNA,in}} - C_{\mathrm{DNA}}}{\tau} - r_{\mathrm{seq}}, \tag{3}$$

where $r_{\mathrm{seq}}$ equals the initiation rate $r_{\mathrm{tr}}$ in batch. All other state derivatives receive the same convective term; the algebraic constraints are unaltered. The remainder of the parameter set, stoichiometry, and rate laws were retained verbatim. Because only two extra quantities – $C_{i,\mathrm{in}}$ and $\tau$ – must be supplied, the reformulation preserves the mechanistic detail model while enabling dynamic computation of IVT CSTRs for the end-to-end digital twin. All balances and algebraic constraints were coded in Julia and solved with the DifferentialEquations.jl package.

### 3.2. Continuous countercurrent tangential chromatography (CCTC)

For mRNA purification using CCTC with oligo(dT) resin, the resin slurry moves countercurrently to the buffer streams through sequential processing steps-binding, washing, elution, stripping, and equilibration (Fig. 1). Given the critical role of binding, washing, and elution in defining the overall yield and purity of mRNA, modeling these steps is essential for process understanding and optimization.

*Binding model.* Binding is the selective adsorption of mRNA from the bulk mobile phase onto the resin. This step primarily determines the overall process yield by defining the amount of mRNA that is initially captured. It also influences mRNA recovery, since the efficiency with which bound mRNA is eluted depends on the effectiveness of the initial binding interaction. This process is typically modeled by mass balances (Guiochon, 2006).

Due to the unique operation and well-mixed nature of the CCTC, axial dispersion in the bulk phase can be neglected. Intraparticle diffusion within the resin particles is considered the dominant mechanism governing the uptake of mRNA. Defining $c_p$ as the concentration of the free mRNA inside the pores of resin particles and $q$ as the concentration of the bound mRNA, the balance is

$$\frac{\partial c_p}{\partial t} + \upsilon \frac{\partial c_p}{\partial x} = \frac{D_{\mathrm{eff}}}{\varepsilon_p} \frac{\partial^2 c_p}{\partial r^2} + \frac{2 D_{\mathrm{eff}}}{\varepsilon_p r} \frac{\partial c_p}{\partial r} - \frac{1 - \varepsilon_p}{\varepsilon_p} \left( \frac{\partial q}{\partial t} + \upsilon \frac{\partial q}{\partial x} \right), \tag{4}$$

where $\varepsilon_p$ is the internal porosity of the resin and $D_{\mathrm{eff}}$ is the effective diffusivity. Adsorption follows Langmuir kinetics,

$$\frac{\partial q}{\partial t} + \upsilon \frac{\partial q}{\partial x} = k_a c_p (q_\infty - q) - k_d q, \tag{5}$$

where $q_\infty$ is the maximum adsorption capacity of the resin and $k_a, k_d$ are the adsorption and desorption constants. A Danckwerts film condition closes the particle surface:

$$\left. \frac{\partial c_p}{\partial r} \right|_{r=R_0} = \frac{k_f}{D_{\mathrm{eff}} \varepsilon_p} \left( c_m - c_p|_{r=R_0} \right),$$

and symmetry gives $\partial c_p / \partial r = 0$ at $r = 0$. The mRNA concentration in the bulk fluid $c_m$ is then modeled as

$$\frac{\partial c_m}{\partial t} + \upsilon \frac{\partial c_m}{\partial x} = -\frac{(1 - \epsilon_c)}{\epsilon_c} \frac{A k_f}{V} \left( c_m - c_p|_{r=R_0} \right), \tag{6}$$

where $\epsilon_c$ denotes the interstitial porosity, $\upsilon$ is the flow velocity, $A$ is the particle surface area, and $V$ is the particle volume. At the column inlet $x = 0$, $c_m = c_{in}$, while $\partial c_m / \partial x = 0$ at the outlet $x = L$.

The above partial differential–algebraic equations describe the transport and sorption processes between mRNA molecules and porous resins within a Eulerian framework, resolved along the axial coordinate $x$. With system parameters, $q$ can be solved numerically and assumed to be preserved to evaluate mRNA recovery after elution.

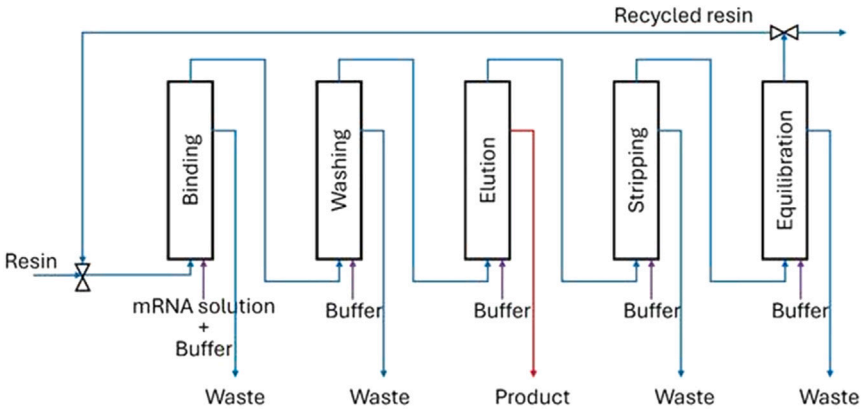


**Fig. 1.** Schematic diagram of the CCTC system.

*Washing and elution performance.* MRNA is selectively bound to oligo(dT) resin under high-salt conditions and eluted in a low-ionic-strength buffer (Flook, 2021). To enable this transition to be controlled in a continuous mode, CCTC utilizes a series of hollow-fiber membrane modules to provide spatial and temporal gradients in buffer conditions and pH throughout the washing and elution steps. However, these gradients can result in local non-uniformities in buffer composition and pH, which may compromise overall mRNA yield and purity by introducing premature desorption and incomplete washing.

By assuming that impurities do not bind to the resin and the elution recovers bound mRNA quantitatively, the step efficiency depends mainly on the operating parameters, such as buffer-to-slurry flow-rate ratio $\gamma = Q_{\text{buffer}}/Q_{\text{slurry}}$ and the number of stages $N$ (Dutta et al., 2016). The fractional impurity removal $R$ (wash) or product yield $Y$ (elution) can be approximated by

$$R \text{ (or } Y) = \frac{\alpha\left(\alpha^{N}-1\right)}{\alpha^{N+1}-1},$$

where

$$\begin{aligned} &\alpha = \gamma S K, \\ &K = \frac{1}{1-\varphi_{\text{ave}}(1-\omega)}, \text{ and} \\ &\varphi_{\text{ave}} = \frac{\varphi_{\text{in}}(\gamma+2)}{2\gamma+2}. \end{aligned} \tag{7}$$

In the equation above, $S$ is defined as the average sieving coefficient and is set to be 1 for mRNA-bounded resins that are fully retained in the retentate, $K$ is a correction factor to account for the resin volume fraction in the column, and $\varphi_{\text{ave}}$ denotes the averaged resin volume fraction. $\varphi_{\text{in}}$ represents the resin volume fraction in the inlet stream, while $\omega$ corresponds to the liquid fraction in the resin (Dutta et al., 2017).

Eq. (7) guides preliminary sizing and theoretical system efficiency, while detailed optimization relies on the result from the binding step. MATLAB's `ode15s` solver was used to solve the discretized ordinary differential Eqs. (4)–(6).

### 3.3. Tangential flow filtration (TFF)

This section discusses the model development of the SPTFF used to concentrate the IVT broth and then use a countercurrent diafiltration (CCDF) section to wash out proteins, NTPs, and salts. The model is designed to predict (i) the outlet concentrations of mRNA, proteins, and small solutes after the single-pass concentration step and after the subsequent countercurrent wash, and (ii) the full axial concentration profiles along the membrane, which are needed for future fouling and scale-up studies. It can be run in either static or vibratory mode, so the influence of membrane vibration on critical flux and protein transmission can be examined directly.

The TFF block therefore has two sub-models: a single concentrator and an $n$-stage CCDF loop, as indicated in Fig. 2.

*Module and species groups.* All calculations use a 100 kDa SANI vibratory hollow-fiber cartridge (Behboudi et al., 2025). Solution components are lumped into: mRNA ($i = 1$), protein ($i = 2$), and NTP + salts ($i = 3$).

*Concentration.* Axial dispersion is neglected, giving the one-dimensional balances

$$A_{\text{cs}}\frac{dC_i}{dt} + q(x)\frac{dC_i}{dx} = C_i(1-s_i)Ja, \tag{8}$$

$$q(x) = q_0 - Jax, \qquad a = \frac{A_{\text{mem}}}{L}, \tag{9}$$

where $A_{\text{cs}}$ is the lumen cross-section $\text{cm}^2$, $q(x)$ is the retentate flow rate $\text{mL}\,\text{min}^{-1}$ at axial position $x$, $q_0$ is inlet flow rate, $C_i(x,t)$ is the bulk retentate concentration of species $i$ $[\text{g}\,\text{L}^{-1}]$, $s_i$ is the transmission coefficient $= C_{i,p}/C_{i,r}$ (permeate/ retentate), $J$ is the imposed permeate flux $\text{m}\,\text{s}^{-1}$, $A_{\text{mem}}$ membrane area, and $L$ is the module length.

We set $s_1 = 0$ (mRNA fully retained) and $s_3 = 1$ (NTP/salts fully transmitted). For proteins,

$$1 - s_2 = R_2 = \frac{(1-X)^S - (1-X)}{X}, \tag{10}$$

where $X = Q_p/Q_0$ is the volumetric conversion and $S$ is a fitted exponent (different for vibratory and static modes) (Behboudi et al., 2025).

*Critical flux.* The maximum sustainable flux was determined by fitting a simple power law for a vibration and a Hill function for the no-vibration case,

$$J_{\text{crit,vibro}} = AQ_f^n, \tag{11}$$

$$J_{\text{crit,novibro}} = \frac{LQ_f^n}{K+Q_f^n}, \tag{12}$$

where $Q_f$ is the feed flow. The parameters $A$, $n$, $L$, and $K$ were fit to match experimental data. If a user-requested conversion exceeds $X_{\text{crit}} = J_{\text{crit}}A/q_0$, the model caps the run at $X_{\text{crit}}$.

*Countercurrent diafiltration (CCDF).* CCDF is represented by $n$ well-mixed retentate stages plus a final permeate stage (Fig. 2). With retentate flowrate $R_0$, buffer flow $P$, and stage volume $V_{\text{TFF}}$,

$$\frac{dC_{i,1}}{dt} = \frac{C_{i,0}R_0 + s_iP\,C_{i,2} - (R_0 + s_iP)C_{i,1}}{V_{\text{TFF}}}, \tag{13}$$

$$\frac{dC_{i,j}}{dt} = \frac{C_{i,j-1}R_0 + s_iP\,C_{i,j+1} - (R_0 + s_iP)C_{i,j}}{V_{\text{TFF}}}, \qquad 2 \le j \le n, \tag{14}$$

$$\frac{dC_{i,n+1}}{dt} = \frac{C_{i,n}R_0 - (R_0 + s_iP)C_{i,n+1}}{V_{\text{TFF}}}, \tag{15}$$

for $i = 1, 2, 3$. Eqs. (8)–(15) are integrated with MATLAB's `ode15s` solver.

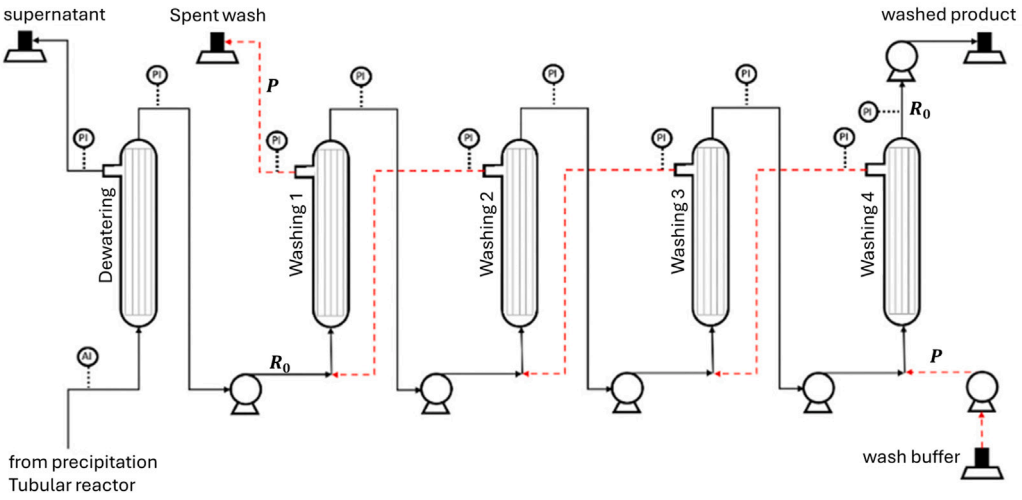


**Fig. 2.** Schematic diagram of the concentration (dewatering) and 4-stage countercurrent washing. $R_0$ and $P$ correspond to the retentate flow rate and the buffer flow rate through the system.
*Source:* Adapted from Minervini et al. (2024).

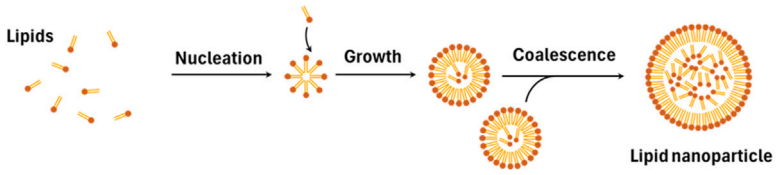


**Fig. 3.** Schematic diagram of LNP formation used for model development in this work.

### 3.4. LNP formulation

The LNP model was developed to calculate the PSD using population balance equations (PBE). To implement PBEs, LNP formation is modeled as an antisolvent crystallization driven by lipid supersaturation (as shown in Fig. 3). Readers are referred to Shin et al. (2025) for a detailed discussion of the model development.

In this work, nucleation, growth, and coalescence are assumed to be the main driving kinetics in LNP formation. Accordingly, PBEs are formulated to describe the time-evolution of the number density function as expressed by

$$\frac{\partial n(L)}{\partial t} = B_n + B_c + D_c - \frac{\partial (Gn)}{\partial L} \tag{16}$$

where $n$ represents the number density of LNPs, $B_n$ represents the birth term due to nucleation, and $B_c$ and $D_c$ represent the birth and death terms due to coalescence, respectively. The last partial term represents the contribution of growth.

To facilitate precise control over the LNP product, the LNP model considers the flow rate ratio (FRR, the ratio of the aqueous stream to the organic stream), lipid concentration, and electrolyte concentration as input parameters, as these factors have a profound impact on LNP size. To numerically integrate the PBEs, the equations are discretized into 200 bins and integrated using the `ode15s` solver in MATLAB.

### 3.5. Lyophilization

We used a mechanistic model of the state-of-the-art continuous lyophilization technology from our earlier works (Srisuma et al., 2025), in which a number of vials containing biologics are suspended and move continuously. The mechanistic model consists of three main parts corresponding to the key steps of lyophilization, namely: (1) freezing, (2) primary drying, and (3) secondary drying (see Fig. 4). The model captures all key phenomena and accurately predicts the critical process parameters for the entire process. Detailed description of the model, including modeling strategies, model equations, parameters, and relevant numerical methods, can be found in Fig. 4.

During the freezing step, the product and liquid solvent (usually water) in a vial are cooled such that most of the liquid (free water) is frozen, with the remaining fraction (bound water) retaining its liquid state adsorbed to the solid material between the ice crystals. The model is derived based on the energy balance associated with cooling, nucleation, and solidification, resulting in a system of ordinary differential equations (ODEs) describing the evolution of the product temperature $T(t)$ and fraction of ice/water $m_w(t)$. The stochastic nature of ice nucleation follows a Poisson process, that is,

$$P = 1 - \exp\left(\int_t^{t+\Delta t} -\lambda dt\right), \tag{17}$$

where $P$ is the probability of first nucleation, $t$ is time, and $\lambda$ is the rate constant of the Poison process. This probability can be incorporated into the heat and mass balances for freezing.

In primary drying, the product is heated under vacuum (below the triple point), and so the free water (in the form of ice crystals) is removed via sublimation. Our model comprises a system of partial differential equations (PDEs) that describe the simultaneous heat and mass transfer associated with sublimation of ice crystals, which can predict the evolution of the product temperature $T(z,t)$ and sublimation front $S(t)$. The energy balance equation is

$$\rho C_p \frac{\partial T}{\partial t} = k\frac{\partial^2 T}{\partial z^2} + Q_V, \tag{18}$$

where $T(z,t)$ is the temperature, $z$ is the vertical direction, $t$ is time, $k$ is the thermal conductivity, $\rho$ is the density, $C_p$ is the heat capacity, $Q_V$ is the volumetric heat source (e.g., thermal radiation, energy associated with phase transition). During primary drying (sublimation), the energy balance at the sublimation front is

$$N_w \Delta H_{sub} = k\frac{\partial T}{\partial z} + Q_{top}, \tag{19}$$

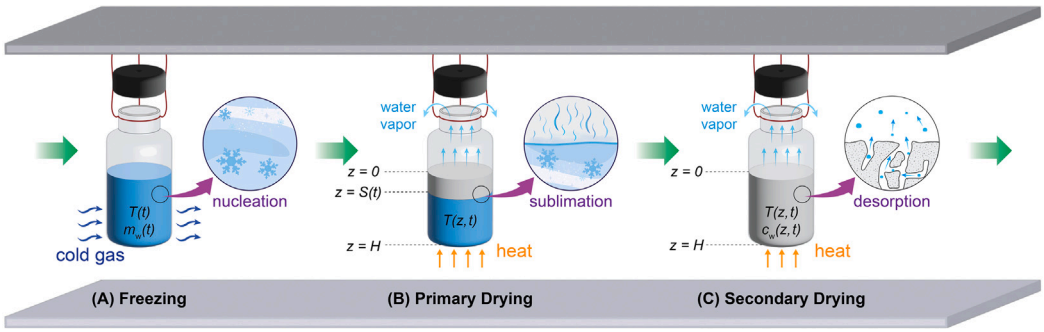


**Fig. 4.** Freeze-dryer configuration used for the lyophilization model.
*Source:* Adapted from Srisuma et al. (2025).

where $N_w$ is the sublimation flux, $\Delta H_{sub}$ is the heat of sublimation, and $Q_{top}$ represents the contribution of heat transfer from the top surface. The mass balance of water vapor at the sublimation front is

$$\frac{dS}{dt} = \frac{N_w}{\rho - \rho_e}, \tag{20}$$

where $S(t)$ is the sublimation front and $\rho_e$ is the effective density of the dried region above the sublimation front. The driving force for mass transfer at the sublimation interface is

$$N_w = \frac{p_{w,sat} - p_{w,c}}{R_p}, \tag{21}$$

where $p_{w,sat}$ is the saturation/equilibrium pressure of water, $p_{w,c}$ is the partial pressure of water in the chamber (environment), and $R_p$ is the mass transfer resistance.

Finally, in secondary drying, the product is heated further such that the remaining bound water is removed mainly via desorption. The model is based on a system of PDEs that characterize the coupled heat and mass transfer during desorption of bound water. The evolution of the concentration of bound water $c_w(z,t)$ as well as the product temperature $T(z,t)$ can be predicted. The desorption kinetics of bound water is described by

$$\frac{\partial c_w}{\partial t} = k_d(c_w^* - c_w), \tag{22}$$

where $c_w$ is the concentration of bound water, $c_w^*$ is the equilibrium concentration, and $k_d$ is the rate constant for desorption that exhibits Arrhenius temperature dependence

$$k_d = f_a e^{-E_a/RT}, \tag{23}$$

where $f_a$ is the frequency factor (aka collision frequency), $E_a$ is the activation energy, and $R$ is the gas constant.

The lyophilization simulation was done in MATLAB using the `ode15s` solver. The PDE models were discretized spatially to produce a system of ODEs prior to integration with `ode15s`.

A process-wide summary of CPPs and CQAs used for software demonstration is provided in Table 1. Users can extend these lists with minimal code changes.

## 4. Software architecture and features

We developed the mRNA manufacturing digital twin software to unify multiple languages, specialized domain code, and a user-friendly frontend into a cohesive system for end-to-end continuous mRNA manufacturing. The decision to embed process models in MATLAB or Julia while orchestrating everything through Python and FastAPI was motivated by the availability of well-tested models in those languages for some unit operations and by the readability and flexibility of Python for bridging. A React frontend was added to maximize accessibility, so that users can configure and run simulations without writing code.

Fig. 5 illustrates how each software component interacts from user request to data storage. The user first accesses the web-based graphical interface (React), which sends requests (JSON) to the FastAPI server. The server orchestrates process calculations by invoking mechanistic models (e.g., Julia or MATLAB) and then stores or retrieves results in a central SQL database. Historical data, such as previous runs with unique Run IDs, can be easily compared by issuing subsequent queries. This architecture enables rapid scenario iteration and robust traceability of simulation inputs, outputs, and process parameters.

We view the plug-and-play approach as the defining principle: users can arrange different unit operations in any order, with each unit operation having the freedom to be replaced, removed, or duplicated as needed. In doing so, the software demonstrates a high degree of flexibility that we find essential, especially in the mRNA vaccine manufacturing field where modeling approaches often emerge at a rapid pace. Next, we discuss each of the software components in detail.

### 4.1. Python server-side backend

Our backend uses a Python + FastAPI setup that exposes a RESTful API for the frontend. We adopted Python because it is easy to maintain and allows straightforward interfacing with MATLAB and Julia. FastAPI was selected for its speed, clear structure, and built-in support for Pydantic data validation. When the server starts, it creates a FastAPI app and configures CORS to let the React frontend communicate. The codebase maintains an in-memory dictionary to store intermediate and final simulation outputs, keyed by unique identifiers. The main logic resides in `main.py` and defines two main endpoints: `/run_chain` and `/get_unit_result`.

- `/run_chain` accepts a JSON body describing the sequence of unit operations that the user wishes to run. Each unit operation is identified by an ID and an inputs dictionary that follows a Pydantic schema, `MembraneInput`). If the JSON format is invalid (e.g., negative flow rate), FastAPI rejects the request. Upon acceptance, the server calls a bridging function specific to each operation (e.g., `julia_interface.py` for IVT). Simulation outputs are stored in the dictionary and passed to the next operation in the sequence. The endpoint returns a JSON list of results.
- `/get_unit_result` returns the final output for a specific unit operation identified by its unique ID. It enables users to inspect individual results or test units in isolation.

These two endpoints, together with the bridging scripts that call MATLAB or Julia code, form the backend foundation of the digital twin software.

### 4.2. Frontend and graphical user interface

We built a React frontend that organizes the entire continuous manufacturing process and lets users add, remove, and reorder unit

**Table 1**
CPPs and CQAs/monitored outputs used in this paper. Users can add additional CPPs/CQAs with minimal code edits.

| Unit operation | CPPs shown in GUI | CQAs/outputs shown in GUI (used here) |
|---|---|---|
| IVT | F101 (flow rate); Reactor volume; T7RNAP; ATP; CTP; GTP; UTP; Mg; DNA; (simulation step & final time) | mRNA titer (TotalRNA); ATP/UTP/CTP/GTP (residuals); Phosphate; pH; Total Mg |
| CCTC | mRNA inlet concentration; Resin loading; Additional parameters include adsorption rate constant, bed void fraction, and adsorption rate constant | Unbound mRNA (eluate); Bound mRNA (on-resin trajectory) |
| TFF | F103 (feed flow); mRNA In; Protein In; NTPs In; Conversion setpoint; Stages (diafiltration); Buffer flow (D); Filter type; Additional parameters also provided | Actual conversion ($X_{actual}$); Critical flux ($J_{crit}$); Stage-wise concentrations (mRNA/protein/NTPs) over time/position |
| LNP formulation | Residence time; Flow-rate ratio (FRR); pH; Ionic strength; Total flow rate; Lipid In.; mRNA In. | Encapsulation efficiency (EE); Mass fraction Solids; Z-average diameter; PSD curve; PDI; PSD percentiles D10, D25, D50, D75, D90) |
| Lyophilization (freeze-drying) | Initial temperatures (freeze/primary/secondary); Shelf temperatures (primary/secondary); Cold-gas (freezing) temperature; Chamber pressure; Fluid volume; Solids mass fraction | Mass of ice; Bound water; Product temperature; Operating pressure; Operating temperature |

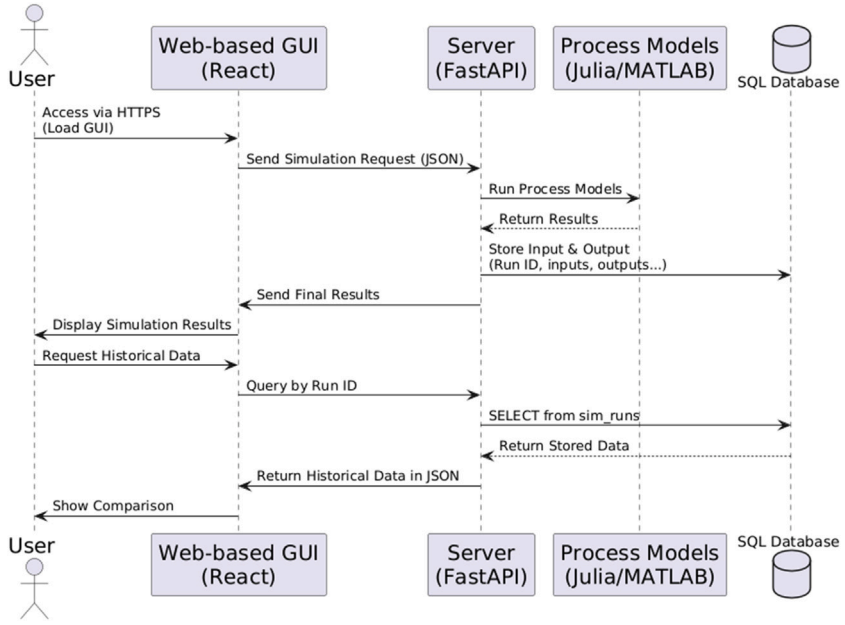


**Fig. 5.** Schematic diagram of the software architecture and data flow of the mRNA manufacturing digital twin.

operations in a visual workspace (Fig. 6). The "Available Units" bar includes buttons like "IVT Unit", "TFF Unit", "CCTC Unit", "LNP Unit", and "Freeze-drying Unit". Users drag units into a "Current Flow" row, assigning each a unique identifier. Multiple instances (e.g., two TFFs) are also supported (Fig. 6).

Each unit has an Edit button that opens a modal for entering parameters aligned with the backend's Pydantic schema. When users click "Run Plant", the frontend sends a JSON payload describing the simulation sequence to `/run_chain`.

Each unit also has an individual GUI accessible via endpoints such as `/ivt?uniqueId=ivt_1234`. These interfaces allow focused runs with time series outputs and adjustable parameters. We also introduced a tag-based variable selection system: users click variable tags (e.g., mRNA concentration) to visualize trends without overwhelming plots. Switching variables does not re-run the simulation; it only updates the graph.

### 4.3. SQL-based data management and retrieval

We integrated an SQL-based data management system using SQLite for comprehensive data analytics and historical comparison. Each simulation run generates an entry into the database structured as follows:

- **Timestamp**: A UTC-based ISO 8601 timestamp to record the initiation moment.
- **ChainRequest**: A JSON serialized record of user-defined inputs for each unit operation.
- **ChainResults**: JSON serialized results capturing final and intermediate outputs from each unit operation.

Data storage utilizes Python's SQLite3 interface that ensures efficient serialization and retrieval. This implementation provides robust data integrity and retrieval performance. Users leverage SQL queries to:

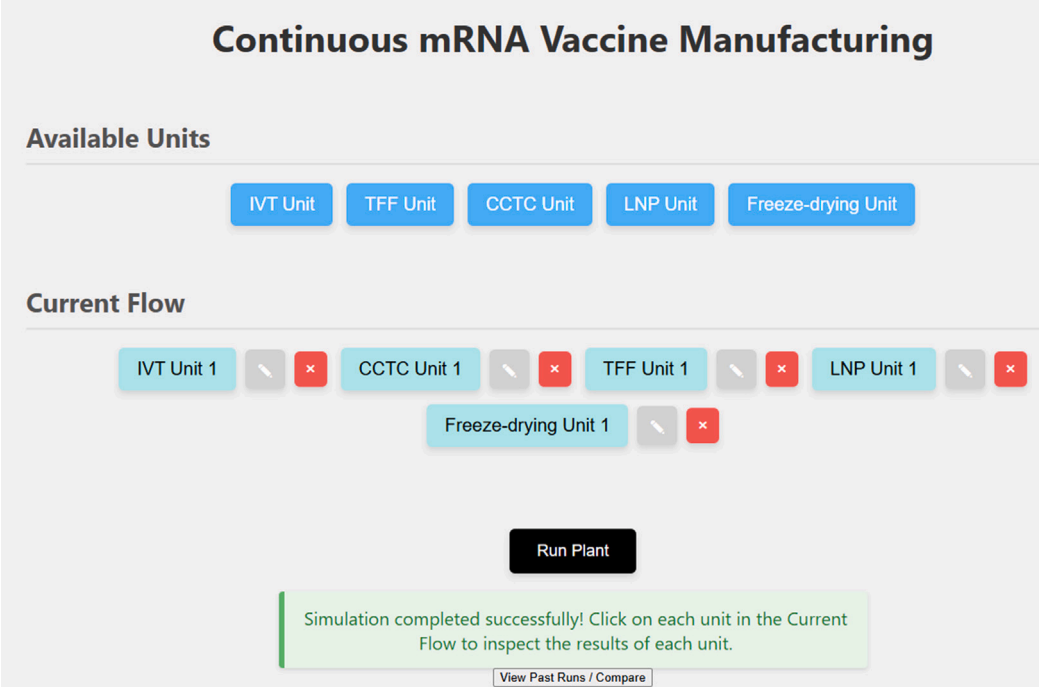


**Fig. 6.** Main user interface of the software that allows users to add and rearrange unit operations for the simulation, specify parameters, and initiate an end-to-end run.

1. Retrieve historical run data indexed by timestamps and unique identifiers for informed decision-making during process design by referencing previous experimental setups and outcomes.
2. Execute direct comparisons across multiple simulation runs for specific output variables (e.g., IVT yields or CCTC mRNA binding efficiency).
3. Access prior configurations and instantly reopen associated unit-specific GUIs based on stored unique identifiers for detailed re-inspection and targeted improvements to individual unit operations.
4. Leverage aggregated run data to conduct design-of-experiments (DoE) studies and scenario analyses for early-stage process design and scale-up.

### 4.4. Data and state management

We handle state both in the backend and frontend. The backend manages transient simulation data in memory during active sessions, whereas persistent historical records are securely maintained in the SQLite database. On the frontend, React local state and context manage pipeline layout and simulation results. This ensures consistent display unless the user manually changes the configuration.

Unique IDs prevent data overwrites when multiple instances of the same operation are run. This enables partial re-runs and preserves outputs for all units. Together, these choices reinforce the plug-and-play concept: each unit can function standalone or be connected in a continuous simulation pipeline.

### 4.5. Unit testing

To help maintain reliability and flexibility in our digital twin, we introduced unit testing for each major step in the pipeline (e.g., IVT, Membrane, CCTC, LNP, and Lyophilization). Whenever there is a change in a unit operation, automated tests verify that it still accepts valid inputs (matching the Pydantic schemas), processes them as intended, and returns the expected outputs.

By isolating each module's logic in this way, we can catch issues quickly if new modifications inadvertently break existing functionality. For this, we have used Python's built-in `unittest` framework to run these tests on every commit, ensuring that the digital twin behaves predictably even as we extend individual unit operation models.

## 5. Illustrative example

To illustrate the capability of the proposed software for end-to-end simulation of mRNA vaccine manufacturing, we configured five sequential unit operations – IVT, TFF, CCTC, LNP formulation, and Freeze-Drying – and ran a complete simulation under representative biopharmaceutical conditions. This section demonstrates how the system couples *mechanistic* models with an *interactive* user interface that yields quantitative, multi-dimensional outputs that mimic a full-scale, continuous mRNA manufacturing process.

Fig. 6 (the "Main Configuration Window") shows how each unit operation is dragged from an "Available Units" bar into a "Current Flow" workspace. By clicking a parameters button for each operation, the user can edit input parameter of each unit operation (see Table 2 for inputs used in this illustrative example). Once all units are arranged in the sequence IVT → CCTC → TFF → LNP → Freeze-Drying, pressing the "Run Plant" button prompts the backend to execute each unit operation in turn, automatically passing outputs from one unit to the next. The IVT unit simulates a CSTR transcription reaction using T7 RNA polymerase. As shown in Fig. 7, the IVT unit's interface provides multiple real-time displays to help users interpret the transcription process. In the upper-right portion of the interface, key *Measured Variables* (e.g., ATP, CTP, UTP, GTP, $Mg^{2+}$) are reported numerically, reflecting their current concentration and activity based on the ongoing simulation. Each variable label functions as a toggle: clicking it dynamically generates and removes corresponding time-series curves in the lower plot area, allowing the user to selectively view trends of interest without overcrowding the graph. For instance, users can track how ATP levels deplete from their initial setpoint while mRNA concentration increases. This layout enables users to trace at a glance how feed streams evolve into product streams and which

**Table 2**
Input parameters for each unit operation in the illustrative example.

| | |
|---|---|
| **IVT** | |
| Q (Feed flow rate) | 1.0 L/h |
| V (Reactor volume) | 2.0 L |
| T7RNAP | $1 \times 10^{-7}$ M |
| ATP, CTP, GTP, UTP | 0.0032 mM |
| Mg (Magnesium) | 0.008 mM |
| DNA | 7.4 nM |
| finaltime (Reaction) | 2.0 h |
| **CCTC** | |
| F103 (Flow rate) | 1.0 L/h |
| mRNA (Inlet concentration) | From IVT |
| resin (Resin loading) | 0.2 |
| **TFF** | |
| qF (Feed flow rate) | 1.0 mL/min |
| c0_mRNA (mRNA inlet concentration) | From CCTC |
| c0_protein (Protein concentration) | 0.5 mg/mL |
| c0_ntps (NTPs concentration) | 0.5 mg/mL |
| X (Conversion) | 0.9 |
| n_stages (Diafiltration stages) | 3 |
| D (Buffer flow) | 4.0 mL/min |
| filterType | VIBRO |
| **LNP** | |
| Residential_time | 1.0 s |
| FRR (Flow-rate ratio) | 3.0 |
| pH | 5.5 |
| Ion (Ionic strength) | 0.1 M |
| TF (Additional flow rate) | 0.0 mL/min |
| C_lipid | 10.0 mg/mL |
| mRNA_in | 0.05 mg/mL |
| **Freeze-Drying** | |
| fluidVolume | $3 \times 10^{-6}$ L |
| massFractionSolids | From LNP |
| InitfreezingTemperature | 298.15 K |
| InitprimaryDryingTemperature | 228 K |
| InitsecondaryDryingTemperature | 273 K |
| TempColdGasfreezing | 268 K |
| TempShelfprimaryDrying | 270 K |
| TempShelfsecondaryDrying | 295 K |
| Pressure | 10 kPa |

reaction components govern process yields. The final simulation result indicated an mRNA titer of approximately 2.31 μM at the end of the transcription reaction.

Output from the IVT unit is next passed to a CCTC unit for purification. Fig. 8 illustrates the CCTC unit's interface, which simulates the bind-and-elute purification of mRNA under continuous flow conditions. The software plots several *time-series* graphs (Fig. 8). For instance, the middle graph panel in Fig. 8 monitors the *unbound* mRNA fraction over a 3-hour run, revealing how binding capacity is progressively utilized. Meanwhile, the right-hand graph tracks the *bound* mRNA concentration, reflecting the resin's adsorption efficiency. In this example, the unit reaches a near-steady state with unbound mRNA at 0.71 g $L^{-1}$ and bound mRNA at 0.97 g $L^{-1}$ after approximately 2 hours, as observed in Fig. 8. By observing how the unbound fraction decreases and the bound fraction increases, users can readily identify resin saturation points and infer whether additional wash steps might be needed.

Following chromatographic purification, the partially purified mRNA stream enters the TFF unit, which performs both concentration and diafiltration. As shown in Fig. 9, the interface is divided into several key panels. The *Concentration Step* and *Diafiltration Step* panels show the concentrations of mRNA, protein, and NTPs. A drop-down menu in the *Diafiltration* panel enables the user to switch among individual TFF stages (Stage 1, Stage 2, Stage 3, etc.), which allows inspection of composition changes across multiple stages of diafiltration in the middle graph panel. A "3D Plots" toggle provides a surface map, for example, of protein concentration as a function of axial position and filtration time. In the future, we plan to replace the bulk concentrations with the concentrations at the membrane surface as a function of axial position, which will help identify operating regimes where fouling may occur. The third graph panel shows time-series plots for the chosen solute (e.g., mRNA) across the specified diafiltration stage. In the illustrative example discussed in this paper, the retentate concentration of mRNA rises from 0.96 g $L^{-1}$ to 8.23 g $L^{-1}$ through concentration and stages of diafiltration.

After TFF, the concentrated mRNA solution proceeds to the LNP formulation unit. The LNP unit receives the upstream mRNA concentration in mg/mL; in this work we standardize this input to 0.05 mg $mL^{-1}$ (assuming dilution), consistent with reported formulation ranges of 0.001–0.05 mg $mL^{-1}$ (Petersen et al., 2023). As illustrated in Fig. 10, the user can set parameters (when initializing the sequence in the main window) such as residence time (defaulting to 1 s), FRR, pH, ionic strength, and lipid concentration. The schematic on the right visually depicts the converging flow streams in a microfluidic or similar mixing device, where the aqueous mRNA phase merges with the lipid phase. In Fig. 10, multiple plots provide key insights into particle formation. The *Z-average Diameter* plot (left) tracks how the mean particle size evolves over the specified residence period. By default, this diameter may begin near zero (immediately after mixing) and then increase as lipids self-assemble around the mRNA cargo. The (PSD) plot (right) displays the intensity-based distribution of particle sizes at the end of the run, often revealing a monomodal peak if formulation conditions are well-optimized. In addition to the Z-average, the user interface reports the polydispersity index (PDI) and D-percentiles (D10, D25, D50, D75, D90).

Beyond these size-related metrics, the LNP unit interface also reports *Encapsulation Efficiency (EE)* – the fraction of total mRNA encapsulated within lipid nanoparticles – and *Mass Fraction*, which indicates the ratio of lipid to mRNA in the final formulation (Fig. 10). An *EE* of around 95% is consistent with small-scale literature data and suggests minimal mRNA loss. From the LNP unit, the solids mass fraction passed to LYO is

$$x_{\text{solids}} = \frac{\text{EE} + C_{\text{lipid}}\left(1 - \text{FRR}\right)}{\rho_{\text{mix}}}, \tag{24}$$

where $\rho_{\text{mix}}$ is the bulk mixture density at the mixer outlet (mg $mL^{-1}$), EE is the encapsulation efficiency, $m_{\text{RNA,in}}$ and $C_{\text{lipid}}$ are in mg $mL^{-1}$, and FRR is the flow rate ratio.

Finally, the formulated LNP suspension undergoes freeze-drying to stabilize the product for storage and transport. The interface shows the progression of the drying cycle, plotting *time–temperature* profiles to illustrate how product temperature moves through critical phases – cooling below the freezing point, sublimation under reduced pressure, and final desorption of bound water (Fig. 11). 2D line graphs (first two graph panels in Fig. 11) show the product's temperature and pressure across the entire process duration. Additionally, a "3D Plot" toggle lets users visualize how temperature varies along the *axial height* of the vial that reveals potential spatial gradients. This level of detail helps process engineers identify optimal pressure–temperature conditions that avoid glass transition temperatures for LNP stability. Moreover, the GUI also reports key drying results, such as final *ice mass* and *bound water* content. In this demonstration, the model predicts a complete sublimation of ice (final ice mass ≈ 0 kg) and a bound-water content of 0.01 kg water $kg^{-1}$ solid, confirming effective removal of residual moisture. Such data provide insights into product homogeneity and dryness, both critical for preserving LNP integrity.

Upon completion, the software automatically serializes all parameter inputs, intermediate outputs, and final results into an SQL database under a unique Run ID (Fig. 12). Each run is listed with its corresponding timestamp and a randomly generated UUID, which ensures traceability and uniqueness across potentially hundreds of trials. Users can click *View Details* to inspect the specific unit operations used in

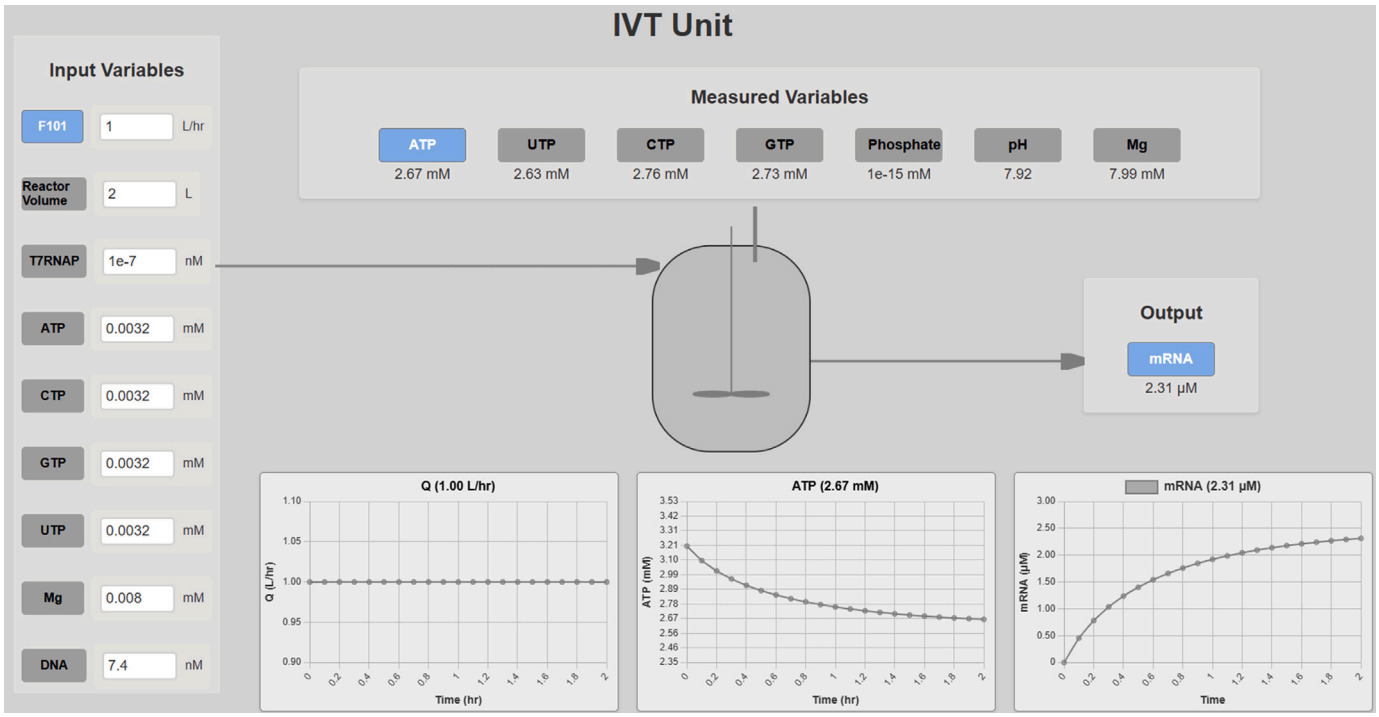


**Fig. 7.** User interface of the IVT unit, displaying time-series results of mRNA and NTP concentrations.

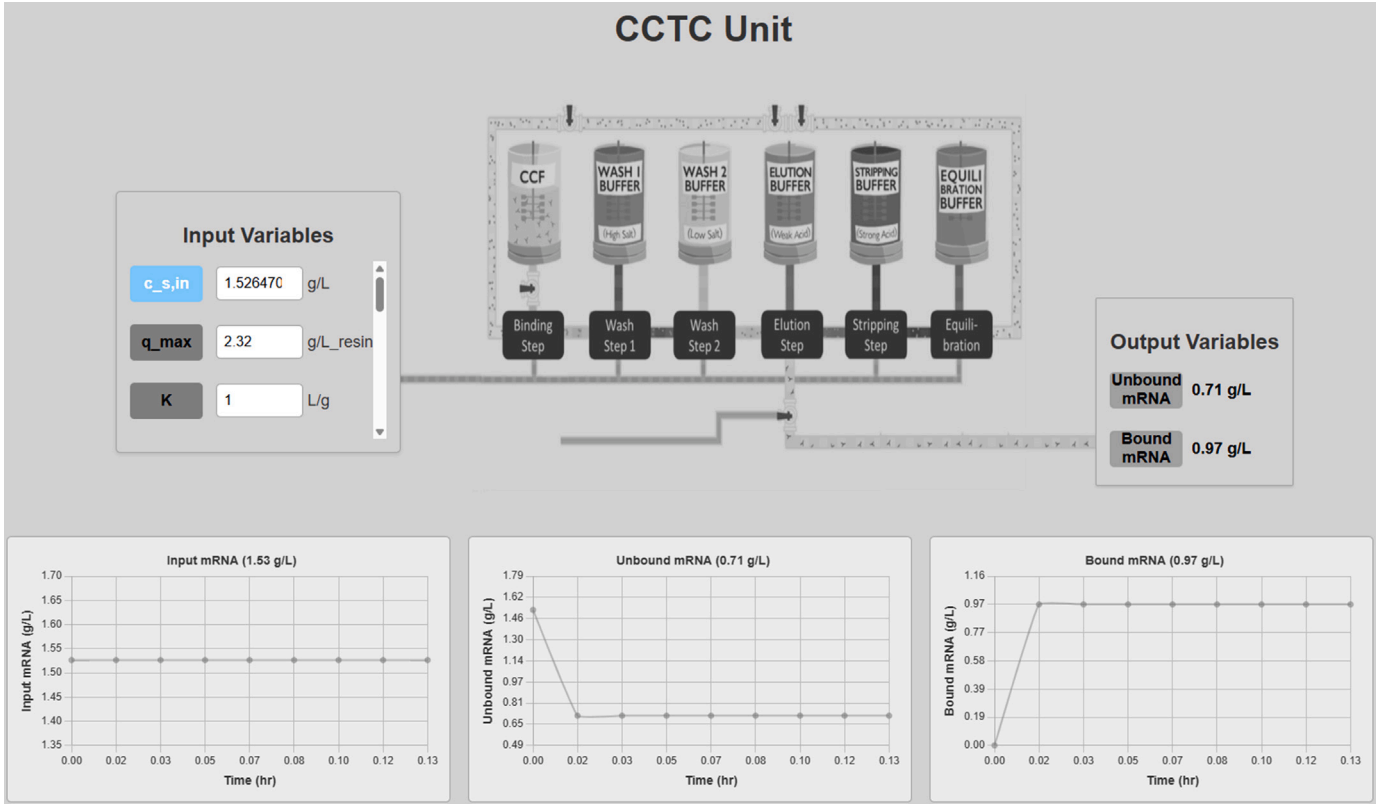


**Fig. 8.** User interface of the CCTC unit, illustrating the bound and unbound mRNA profile.

that run along with their final outputs. Additionally, the interface offers a *Select* checkbox for each run that allows multiple runs to be compared side-by-side on key process variables. For instance, choosing two runs and comparing the `membrane:TFF_mRNA` variable reveals how different input parameters impacted mRNA concentration. This database-centric approach accelerates process tuning and maintains an *audit trail* that supports QbD workflows. Because each Run ID encapsulates the entire simulation chain, any future analyst can reproduce and refine a given setup with minimal effort. This allows rapid retrieval, side-by-side comparisons, and direct integration into

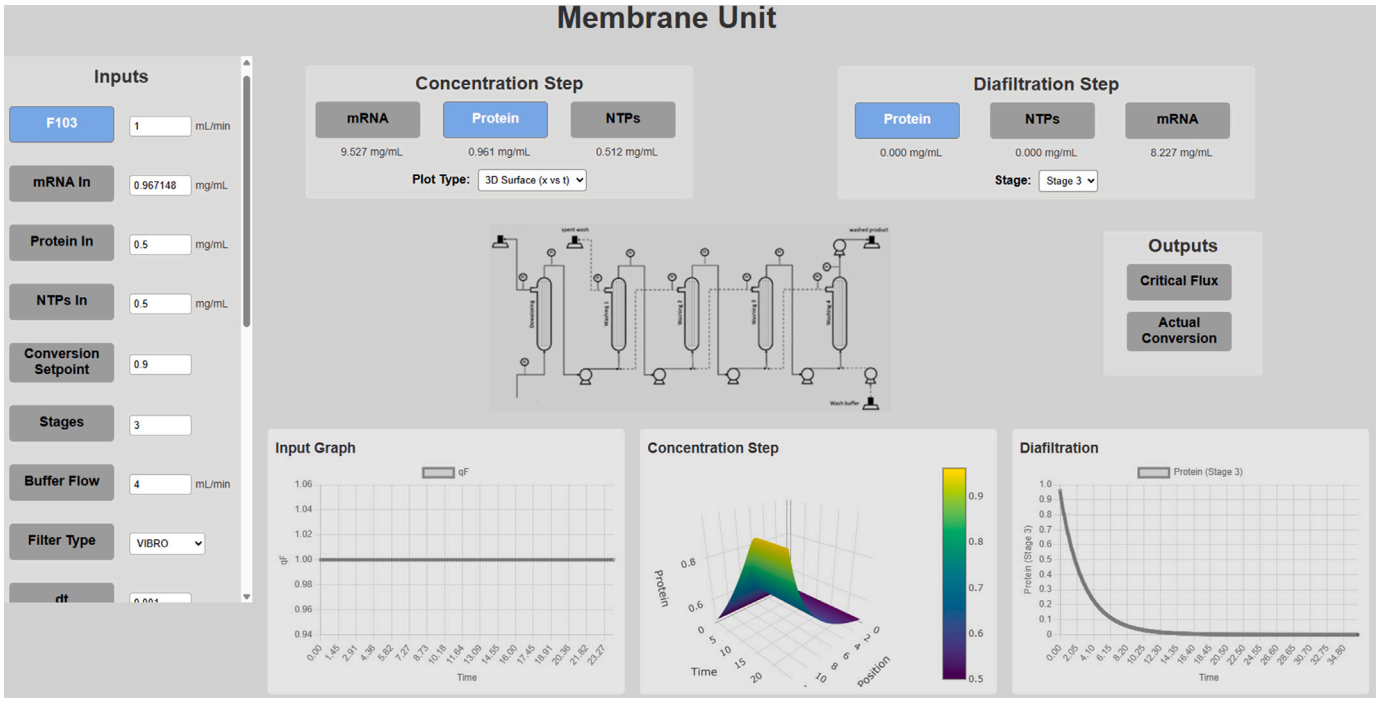


**Fig. 9.** User interface of the TFF unit displaying how mRNA concentration increases over time during concentration and stages of diafiltration. Switching to 3D view allows us to see the axial dependence of the three solutes.

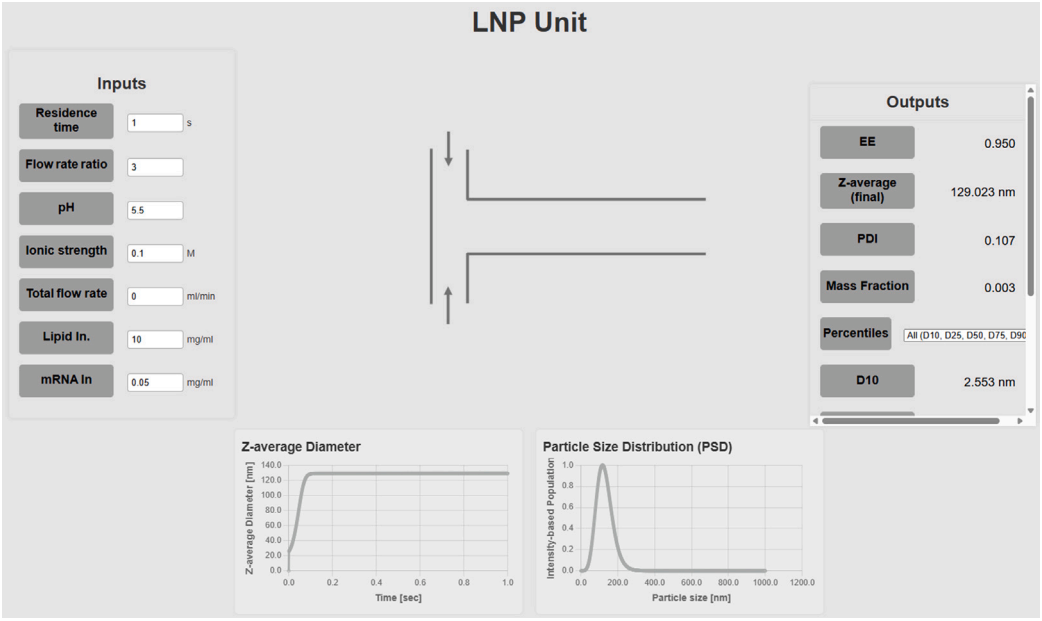


**Fig. 10.** User interface of the LNP unit. The interface allows visualization of the particle size distribution (PSD), Z-average diameter, encapsulation efficiency, and D-percentiles.

DoE workflows. This is critical for accelerating process optimization and enabling advanced QbD strategies.

## 6. Global sensitivity analysis

The software includes the Morris *elementary effects* method (Morris, 1991) to screen the impact of key process parameters on each unit operation. For a model $y = f(x_1, \ldots, x_k)$, the *elementary effect* of parameter $x_i$ at point $\mathbf{x}$ is defined as

$$\mathrm{EE}_i(\mathbf{x}) = \frac{f(x_1, \ldots, x_i + \Delta, \ldots, x_k) - f(\mathbf{x})}{\Delta},$$

where $\Delta$ is a fixed step size in the normalized input space. Over $r$ trajectories, the mean of absolute effects and their standard deviation are computed as

$$\mu_i^* = \frac{1}{r}\sum_{j=1}^{r}\left|\mathrm{EE}_i(\mathbf{x}_j)\right|, \quad \sigma_i = \sqrt{\frac{1}{r-1}\sum_{j=1}^{r}\left(\mathrm{EE}_i(\mathbf{x}_j) - \mu_i\right)^2},$$

where $\mu_i^*$ quantifies the overall sensitivity of $x_i$ on the model output, while $\sigma_i$ quantifies nonlinear and interaction effects. Applying this method, several key insights emerged across the different unit operations.

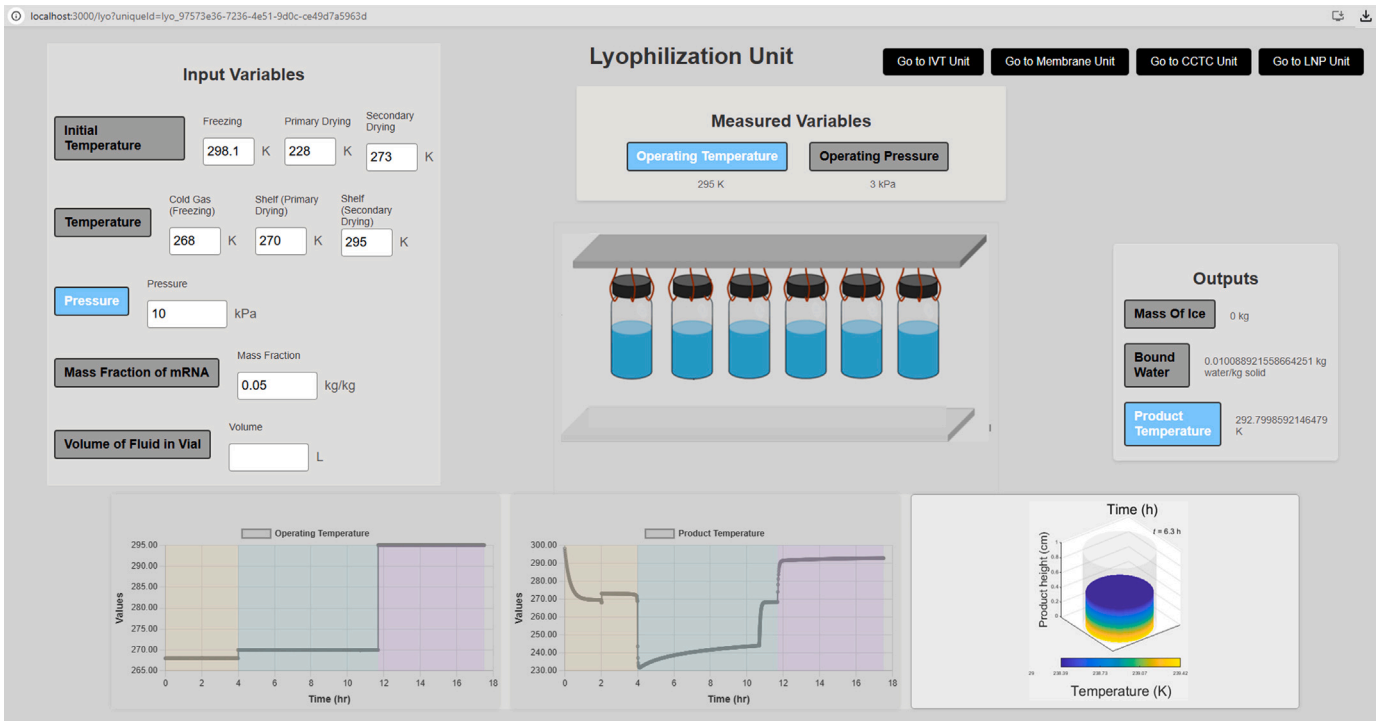


**Fig. 11.** User interface of the Freeze-Drying unit with a time–temperature plot of the product undergoing primary and secondary drying. A 3D view captures the evolution of temperature at different vial depths.

## All Simulation Runs

| Select | Run ID | Timestamp | Actions |
|---|---|---|---|
| ☐ | d2541111-ce2c-4c37-afce-83870152adde | 2025-10-02T21:11:10.910740 | View Details |
| ☐ | c03e58e2-4b8a-4642-974c-138f7e232272 | 2025-10-02T15:40:52.111722 | View Details |
| ☐ | 18a037fc-8710-4ef9-95ea-79c18a283c13 | 2025-10-02T15:20:04.638561 | View Details |
| ☐ | 61c402c0-8ad6-43f6-b860-7b8754026d96 | 2025-10-02T15:19:01.532989 | View Details |
| ☐ | aae1b161-106e-44ff-a8e6-f971aff3e910 | 2025-10-02T15:18:17.505538 | View Details |
| ☐ | e2e5181d-a695-440c-9a6a-c19fc33cfd66 | 2025-10-02T15:18:08.873872 | View Details |
| ☐ | b2466985-2407-444c-bccb-40ed0b811e45 | 2025-10-01T23:54:07.242298 | View Details |
| ☐ | 54186ee4-1db0-4c6d-a281-c8f527d56b20 | 2025-10-01T23:53:53.973450 | View Details |
| ☐ | 19fdf02c-6029-4b4c-92ef-52e77e7dec58 | 2025-10-01T14:04:54.620316 | View Details |
| ☐ | 95f51b43-44f4-4858-8c99-f675545f854d | 2025-10-01T14:04:29.475928 | View Details |
| ☐ | 42b8c65e-913f-442d-a62e-91bc4921cada | 2025-10-01T14:04:12.831101 | View Details |
| ☐ | ad907190-44da-4caf-ad37-dc1cc00b4ebd | 2025-10-01T13:19:23.172149 | View Details |
| ☐ | b8a1c9e7-1694-4be6-b79f-6aa94c75ed67 | 2025-10-01T05:34:49.078570 | View Details |
| ☐ | 7180e961-e829-436a-a9c2-4389280237ab | 2025-10-01T05:08:55.291775 | View Details |
| ☐ | d3293e0d-2218-4772-9f96-24632025a260 | 2025-10-01T04:58:06.868480 | View Details |

**Fig. 12.** Database interface showing a unique Run ID for each end-to-end simulation. All unit operation inputs, intermediate states, and final outputs are automatically logged, facilitating traceability and iterative process refinement.

For the IVT unit, magnesium ion concentration is the single most influential factor controlling final RNA titer, exhibiting a Morris index of approximately $\mu^* \approx 1.71 \pm 0.51$ (Fig. 13A). This finding concurs with extensive biochemical literature that identifies $Mg^{2+}$ as a crucial cofactor for T7 RNA polymerase function and affects both catalytic activity and reaction fidelity (Milligan et al., 1987). The loading of T7

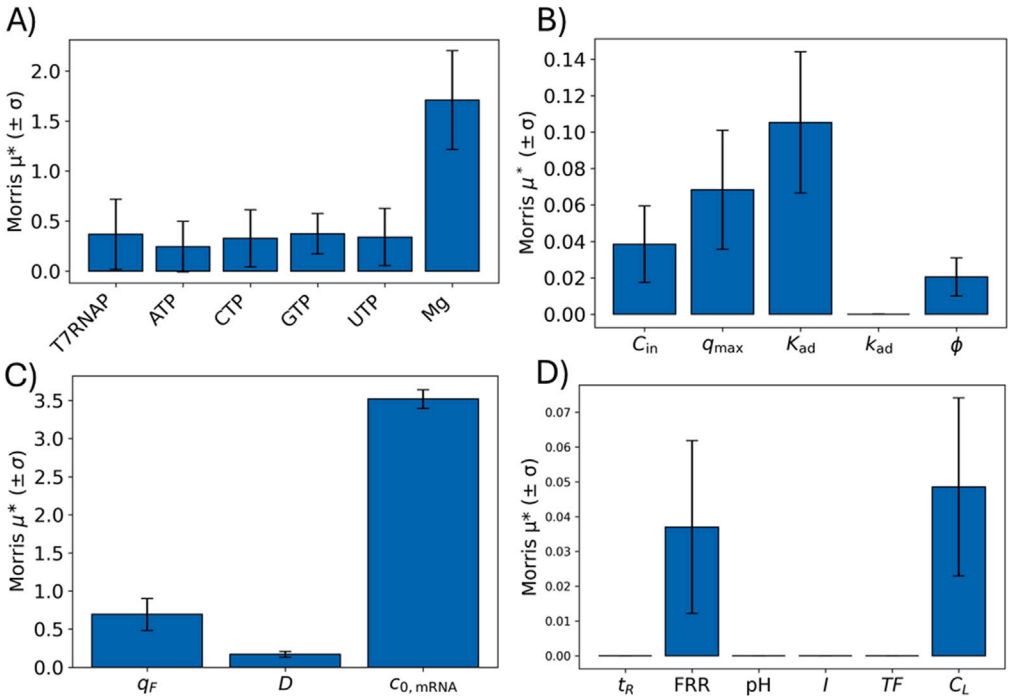


**Fig. 13.** Morris sensitivity screening of the: (A) IVT unit (final RNA titer), (B) CCTC unit (final bound mRNA concentration), (C) TFF unit (mRNA concentration in the diafiltration retentate of the last stage), (D) LNP formulation unit (polydispersity index, PDI). Bars denote the mean absolute elementary effects ($\mu^*$) for each input parameter, with error bars representing the standard deviation ($\sigma$). T7RNAP = T7 RNA polymerase; ATP, CTP, GTP, UTP = nucleotide triphosphates; $C_{in}$ = inlet mRNA concentration; $q_{max}$ = resin adsorption capacity; $K_{ad}$ = adsorption equilibrium constant; $k_{ad}$ = adsorption rate constant; $\phi$ = packing fraction; $q_F$ = input flow rate; $D$ = diafiltration flow rate; $c_{0,mRNA}$ = initial mRNA concentration; $t_R$ = residence time; FRR = flow-rate ratio; TF = total flow rate; $C_L$ = lipid concentration; Ion = ionic strength.

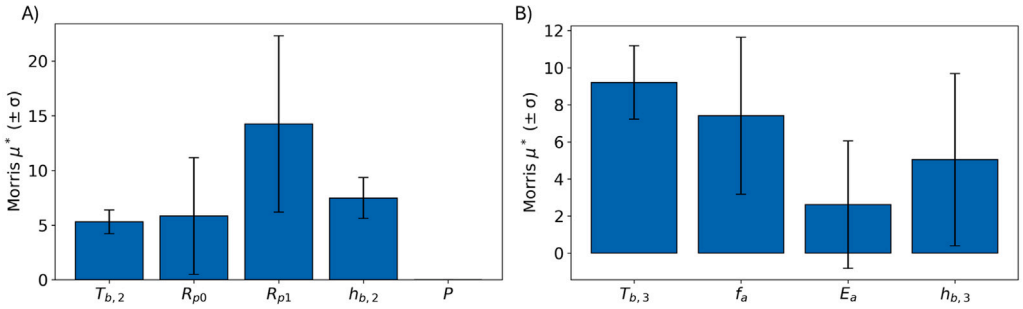


**Fig. 14.** Morris sensitivity of the lyophilization unit: mean product temperature by phase. (A) Primary drying; (B) Secondary drying.
$T_{b,1}$ = freezing shelf temperature; $T_{b,2}$ = primary drying shelf temperature; $T_{b,3}$ = secondary drying shelf temperature; $P$ = chamber pressure; $R_{p0}, R_{p1}, R_{p2}$ = cake-resistance coefficients; $h_{b,2}, h_{b,3}$ = heat-transfer coefficients; $f_a$ = desorption pre-exponential; $E_a$ = desorption activation energy; $T_{prod}$ = product temperature; $\mu^*$ = mean absolute elementary effect; $\sigma$ = standard deviation of elementary effects.

RNA polymerase also significantly impacts output, with $\mu^* \approx 0.59 \pm 0.42$, consistent with enzyme kinetics where polymerase concentration governs transcription rates up to saturation (Henderson et al., 2019). By contrast, the four nucleotide triphosphates (ATP, CTP, GTP, UTP) show moderate and comparable sensitivity levels ($\mu^* \approx 0.25$–$0.35$, $\sigma \approx 0.19$–$0.30$), which indicates their secondary yet essential roles within typical stoichiometric ranges. The relatively large $\sigma$ values for magnesium and polymerase reflect nonlinear and interaction-driven behaviors characteristic of enzymatic regulation and substrate saturation (Kern and Davis, 1997).

As shown in Fig. 13B, the sensitivity screening of the CCTC unit highlights the adsorption equilibrium constant $K_{ad}$ as the primary control knob ($\mu^* \approx 0.11 \pm 0.03$). Resin capacity $q_{max}$ and inlet mRNA concentration $C_{in}$ follow in importance ($\mu^* \approx 0.06 \pm 0.02$ and $0.05 \pm 0.02$, respectively), indicating that both the maximum loading capacity and the feed concentration critically determine final payload. The packing fraction $\phi$ moderately influences flow and residence time ($\mu^* \approx 0.03 \pm 0.02$) through bed morphology effects (Guiochon et al., 2006), while the adsorption rate constant $k_{ad}$ has negligible effect ($\mu^* \approx 5\times10^{-6}$), indicating adsorption operates near equilibrium (Carta and Jungbauer, 2020). The marked separation between $\mu^*_{K_{ad}}$ and the others underscores that thermodynamic binding affinity dominates over kinetic and volumetric factors in defining CCTC performance.

In TFF unit, as shown in Fig. 13C, the analysis reveals inlet mRNA concentration as the dominant determinant of final retentate mRNA levels (with $\mu^* \approx 3.48 \pm 0.12$), whereas the cross-flow feed rate exerts a moderate effect ($\mu^* \approx 0.70 \pm 0.20$) and diafiltration wash rate has negligible influence ($\mu^* \approx 0.15 \pm 0.05$). These findings align with classical mass transfer models of membrane filtration, where concentration polarization and solute back-diffusion, which scale directly with feed concentration, govern retention behavior (Van Reis and Zydney, 2001). Meanwhile, increase in cross-flow feed beyond moderate shear

rates provide diminishing returns. Thus, effective control of feed mRNA concentration is paramount for optimizing TFF recovery, with flow rate adjustments serving as fine-tuning parameters.

As part of our LNP unit sensitivity screening, we applied the Morris method to assess how six process parameters influence the final polydispersity index (PDI). Fig. 13D shows that the FRR and lipid concentration ($C_L$) exert the strongest effects ($\mu^*$ on the order of $10^{-2}$ with large $\sigma$), consistent with FRR determining mixing intensity (Ripoll et al., 2022) and lipid loading controlling the availability of material for particle assembly (Maeki et al., 2017). Buffer pH, ionic strength, and residence time contribute negligibly to PDI in this short-residence window ($\mu^* \approx 0$) (Koitabashi et al., 2021; Nakamura et al., 2022), and total flow rate likewise shows essentially no influence ($\mu^* \approx 0$), indicating that microfluidic mixing geometry rather than bulk flow governs PDI (Ripoll et al., 2022; Maeki et al., 2017).

For the lyophilization unit, we applied Morris sensitivity phase-wise (freezing, primary drying, secondary drying) to identify which inputs control the mean product temperature in each phase. In the freezing phase, varying the gas/shelf temperature $T_{b,1}$ produces a strong, monotonic increase in the mean product temperature during freezing $\mu^* = 10.575 \pm 0.000$. In primary drying, Morris ranks the inputs as $R_{p1} \gg h_{b,2} \gtrsim R_{p0} > T_{b,2} \gg P$ (Fig. 14A); the large $\sigma$ for $R_{p1}$ indicates nonlinearity and interaction, consistent with higher cake resistance suppressing sublimation (less evaporative cooling) and thereby raising $T_{\mathrm{prod}}$, whereas chamber pressure mainly changes the sublimation rate (and thus drying time) with little effect on the product temperature. In secondary drying, sensitivity is dominated by $T_{b,3}$ and the desorption pre-exponential $f_a$, with moderate influence of $h_{b,3}$ and a weaker effect of $E_a$; the elevated $\sigma$ for $f_a$ reflects nonlinearity or interaction in the desorption kinetics over the tested range (Fig. 14B). Overall, shelf setpoints determine the asymptotic product temperature in each phase, whereas resistance and kinetic parameters modulate departures via mass-transfer limits and endothermic moisture removal.

## 7. Conclusions

This work presents a novel digital twin software for integrated, end-to-end manufacturing of mRNA-based therapeutics. The platform unifies key unit operations – IVT, TFF, CCTC, LNP formulation, and freeze-drying – within a modular, plug-and-play architecture. Industry practitioners can add or remove unit operations, incorporate updated mechanistic models, and investigate a wide range of process scenarios through an intuitive user interface. The illustrative example confirmed that the proposed digital twin software provides detailed numerical outputs and multidimensional visualizations that captures the complexity of continuous mRNA production lines in a single automated workflow. Sensitivity results, consistent with the underlying transport and thermodynamics, ranked parameters by influence and provide a basis for future control variable selection and quantification of operating regions of robust operation. By reducing reliance on trial-and-error experiments and consolidating critical process parameters within one interface, the software accelerates, de-risks, and scales up mRNA vaccine manufacturing. It also offers an interactive environment for operator training, thus preparing personnel for continuous operation and enabling advanced QbD workflows.

Future work involves incorporating uncertainty analysis, closed-loop control, and real-time optimization using model predictive control (MPC). We also plan to establish real-time data exchange between the physical process and the digital twin via standardized communication protocols (e.g., OPC-UA). This step will transform the prototype into a fully realized digital twin ecosystem, capable of live data integration and autonomous decision-making. In doing so, the software will serve as an essential foundation for robust, data-driven, and scalable mRNA manufacturing solutions that meet global health demands.

## CRediT authorship contribution statement

**Mohammed Aatif Shahab:** Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Nathan Merica Stover:** Software, Methodology, Conceptualization. **Hasan Al-Mahayni:** Software, Methodology, Conceptualization. **Prakitr Srisuma:** Software, Methodology, Conceptualization. **Jie Wang:** Software, Methodology, Conceptualization. **Sunkyu Shin:** Software, Methodology, Conceptualization. **Anish Dighe:** Software, Methodology, Conceptualization. **Bernhardt L. Trout:** Writing – review & editing, Funding acquisition, Conceptualization. **Allan S. Myerson:** Writing – review & editing, Funding acquisition, Conceptualization. **Richard D. Braatz:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

This research was supported by the U.S. Food and Drug Administration, United States under the FDA BAA-22-00123 program, Award Number 75F40122C00200.

## Data availability

The models, graphical user interface (GUI), and sample results are available at https://github.com/aatifshahab/mRNAdigitalTwin-public.

## References

Agrawal, P., Wilkstein, K., Guinn, E., Mason, M., Serrano Martinez, C.I., Saylae, J., 2023. A review of tangential flow filtration: Process development and applications in the pharmaceutical industry. Org. Process. Res. Dev. 27 (4), 571–591.

Akama, S., Yamamura, M., Kigawa, T., 2012. A multiphysics model of in vitro transcription coupling enzymatic reaction and precipitation formation. Biophys. J. 102 (2), 221–230.

Bae, S.-H., Choi, H., Lee, J., Kang, M.-H., Ahn, S.-H., Lee, Y.-S., Choi, H., Jo, S., Lee, Y., Park, H.-J., 2025. Rational design of lipid nanoparticles for enhanced mRNA vaccine delivery via machine learning. Small 21 (8), 2405618.

Behboudi, A., Javidanbardan, A., Wang, Z., Zhou, L., Zydney, A.L., 2025. Vibration-assisted ultrafiltration dramatically improves mRNA purification. Sep. Purif. Technol. 363, 132179.

Byun, M.J., Lim, J., Kim, S.-N., Park, D.-H., Kim, T.-H., Park, W., Park, C.G., 2022. Advances in nanoparticles for effective delivery of RNA therapeutics. Biochip J. 16 (2), 128–145.

Carta, G., Jungbauer, A., 2020. Protein Chromatography: Process Development and Scale-Up. John Wiley & Sons.

Challener, C.A., 2018. Making the move to continuous chromatography. BioPharm Int. 31 (4), 14–18.

Chaubal, A.S., Zydney, A.L., 2023. Single-pass tangential flow filtration (SPTFF) of nanoparticles: Achieving sustainable operation with dilute colloidal suspensions for gene therapy applications. Membranes 13 (4), 433.

ChromaTan, I., 2024. ChromaTan receives a $2m phase IIB NIH grant to develop a fully integrated continuous purification of AAV via BioRMB™ – a column-free steady-state continuous elution purification platform. https://www.businesswire.com/news/home/20240530827437/en/ChromaTan-Receives-a-%242M-Phase-IIB-NIH-Grant-to-Develop-a-Fully-Integrated-Continuous-Purification-of-AAV-via-BioRMB-a-Column-Free-Steady-State-Continuous-Elution-Purification-Platform, accessed: (Accessed 16 April 2025).

Coffman, J., Brower, M., Connell-Crowley, L., Deldari, S., Farid, S.S., Horowski, B., Patil, U., Pollard, D., Qadan, M., Rose, S., et al., 2021. A common framework for integrated and continuous biomanufacturing. Biotechnol. Bioeng. 118 (4), 1735–1749.

Devos, C., Mukherjee, S., Inguva, P., Singh, S., Wei, Y., Mondal, S., Yu, H., Barbastathis, G., Stelzer, T., Braatz, R.D., et al., 2025. Impinging jet mixers: A review of their mixing characteristics, performance considerations, and applications. AIChE J. 71 (1), e18595.

Dutta, A.K., Fedorenko, D., Tan, J., Costanzo, J.A., Kahn, D.S., Zydney, A.L., Shinkazh, O., 2017. Continuous countercurrent tangential chromatography for mixed mode post-capture operations in monoclonal antibody purification. J. Chromatogr. A 1511, 37–44.

Dutta, A.K., Tan, J., Napadensky, B., Zydney, A.L., Shinkazh, O., 2016. Performance optimization of continuous countercurrent tangential chromatography for antibody capture. Biotechnol. Prog. 32 (2), 430–439.

Dutta, A.K., Tran, T., Napadensky, B., Teella, A., Brookhart, G., Ropp, P.A., Zhang, A.W., Tustian, A.D., Zydney, A.L., Shinkazh, O., 2015. Purification of monoclonal antibodies from clarified cell culture fluid using protein a capture continuous countercurrent tangential chromatography. J. Biotech. 213, 54–64.

Fedorenko, D., Dutta, A.K., Tan, J., Walko, J., Brower, M., Pinto, N.D., Zydney, A.L., Shinkazh, O., 2020. Improved protein a resin for antibody capture in a continuous countercurrent tangential chromatography system. Biotechnol. Bioeng. 117 (3), 646–653.

Fisher, A.C., Lee, S.L., Harris, D.P., Buhse, L., Kozlowski, S., Yu, L., Kopcha, M., Woodcock, J., 2016. Advancing pharmaceutical quality: An overview of science and research in the US FDA's Office of Pharmaceutical Quality. Int. J. Pharm. 515 (1–2), 390–402.

Flook, K., 2021. Supporting development of mRNA-based therapies by addressing large-scale purification challenges. Cell Gene Ther. Insights 7 (5), 489–502. http://dx.doi.org/10.18609/cgti.2021.073.

Fuchs, M., Bhawnani, R., Sripada, S.A., Molek, J., Ghodbane, M., 2023. Predictive modeling of single pass tangential flow filtration for continuous biomanufacturing. Biotechnol. Prog. 39 (5), e3353.

Grieves, M., 2014. Digital twin: Manufacturing excellence through virtual factory replication. White Pap. 1 (2014), 1–7.

Gu, T., 2015. Mathematical Modeling and Scale-Up of Liquid Chromatography: With Application Examples. Springer.

Guiochon, G., 2006. Fundamentals of Preparative and Nonlinear Chromatography, second ed. Academic Press, Amsterdam.

Guiochon, G., Shirazi, D.G., Felinger, A., 2006. Fundamentals of Preparative and Nonlinear Chromatography. Academic Press.

Hahn, T., Geng, N., Petrushevska-Seebach, K., Dolan, M.E., Scheindel, M., Graf, P., Takenaka, K., Izumida, K., Li, L., Ma, Z., Schuelke, N., 2023. Mechanistic modeling, simulation, and optimization of mixed-mode chromatography for an antibody polishing step. Biotechnol. Prog. 39 (2), e3316. http://dx.doi.org/10.1002/btpr.3316.

International Conference on Harmonisation, 2009. Pharmaceutical development Q8(R2). https://database.ich.org/sites/default/files/Q8_R2_Guideline.pdf.

Helgers, H., Hengelbrock, A., Schmidt, A., Strube, J., 2021. Digital twins for continuous mRNA production. Processes 9 (11), 1967.

Henderson, K.L., Evensen, C.E., Molzahn, C.M., Felth, L.C., Dyke, S., Liao, G., Shkel, I.A., Record, Jr., M.T., 2019. RNA polymerase: Step-by-step kinetics and mechanism of transcription initiation. Biochemistry 58 (18), 2339–2352.

Hengelbrock, A., Probst, F., Baukmann, S., Uhl, A., Tschorn, N., Stitz, J., Schmidt, A., Strube, J., 2024. Digital twin for continuous production of virus-like particles toward autonomous operation. ACS Omega 9 (32), 34990–35013.

Inguva, P.K., Mukherjee, S., Walker, P.J., Kanso, M.A., Wang, J., Wu, Y., Tenberg, V., Santra, S., Singh, S., Kim, S.H., Trout, B.L., Bazant, M.Z., Myerson, A.S., Braatz, R.D., 2025. Mechanistic modeling of lipid nanoparticle formation for the delivery of nucleic acid therapeutics. Biotechnol. Adv. 84, 108643.

Javidanbardan, A., Wang, Z., Kostic, A., Behboudi, A., Zydney, A.L., 2025. Single-pass tangential flow filtration (SPTFF) for continuous mRNA concentration and purification. J. Membr. Sci. 719, 123730.

Jesubalan, N.G., Thakur, G., Rathore, A.S., 2023. Deep neural network for prediction and control of permeability decline in single pass tangential flow ultrafiltration in continuous processing of monoclonal antibodies. Front. Chem. Eng. 5, 1182817.

Kaiser, J., Krarup, J., Hansen, E.B., Pinelo, M., Krühne, U., 2022. Defining the optimal operating conditions and configuration of a single-pass tangential flow filtration (SPTFF) system via CFD modelling. Sep. Purif. Technol. 290, 120776.

Kern, J.A., Davis, R.H., 1997. Application of solution equilibrium analysis to in vitro RNA transcription. Biotechnol. Prog. 13 (6), 747–756.

Kimura, N., Maeki, M., Sasaki, K., Sato, Y., Ishida, A., Tani, H., Harashima, H., Tokeshi, M., 2021. Three-dimensional, symmetrically assembled microfluidic device for lipid nanoparticle production. RSC Adv. 11 (3), 1430–1439.

Koitabashi, K., Nagumo, H., Nakao, M., Machida, T., Yoshida, K., Sakai-Kato, K., 2021. Acidic pH-induced changes in lipid nanoparticle membrane packing. Biochim. et Biophys. Acta (BBA)-Biomembranes 1863 (8), 183627.

Kulkarni, J.A., Cullis, P.R., Van Der Meel, R., 2018. Lipid nanoparticles enabling gene therapies: From concepts to clinical utility. Nucleic Acid Ther. 28 (3), 146–157.

Kulkarni, J.A., Witzigmann, D., Thomson, S.B., Chen, S., Leavitt, B.R., Cullis, P.R., van der Meel, R., 2021. The current landscape of nucleic acid therapeutics. Nature Nanotechnology 16 (6), 630–643.

Lamoot, A., Lammens, J., De Lombaerde, E., Zhong, Z., Gontsarik, M., Chen, Y., De Beer, T.R., De Geest, B.G., 2023. Successful batch and continuous lyophilization of mRNA LNP formulations depend on cryoprotectants and ionizable lipids. Biomater. Sci. 11 (12), 4327–4334.

Leng, J., Wang, D., Shen, W., Li, X., Liu, Q., Chen, X., 2021. Digital twins-based smart manufacturing system design in Industry 4.0: A review. J. Manuf. Syst. 60, 119–137.

Leweke, S., von Lieres, E., 2018. Chromatography analysis and design toolkit (CADET). Comput. Chem. Eng. 113, 274–294.

Madsen, E., Kaiser, J., Krühne, U., Pinelo, M., 2022. Single pass tangential flow filtration: Critical operational variables. Fouling, Main Curr. Appl. Sep. Purif. Technol. 291, 120949.

Maeki, M., Fujishima, Y., Sato, Y., Yasui, T., Kaji, N., Ishida, A., Tani, H., Baba, Y., Harashima, H., Tokeshi, M., 2017. Understanding the formation mechanism of lipid nanoparticles in microfluidic devices with chaotic micromixers. PLoS One 12 (11), e0187962.

Maeki, M., Uno, S., Niwa, A., Okada, Y., Tokeshi, M., 2022. Microfluidic technologies and devices for lipid nanoparticle-based RNA delivery. J. Control. Release 344, 80–96.

Massei, A., Fissore, D., 2023. A new model-based approach for the development of freeze-drying cycles using a small-scale freeze-dryer. J. Pharm. Sci. 112 (8), 2176–2189.

Messerian, K.O., Zverev, A., Kramarczyk, J.F., Zydney, A.L., 2022. Pressure-dependent fouling behavior during sterile filtration of mRNA-containing lipid nanoparticles. Biotechnol. Bioeng. 119 (11), 3221–3229.

Meulewaeter, S., Nuytten, G., Cheng, M.H., De Smedt, S.C., Cullis, P.R., De Beer, T., Lentacker, I., Verbeke, R., 2023. Continuous freeze-drying of messenger rna lipid nanoparticles enables storage at higher temperatures. J. Control. Release 357, 149–160.

Milligan, J.F., Groebe, D.R., Witherell, G.W., Uhlenbeck, O.C., 1987. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. Nucleic Acids Res. 15 (21), 8783–8798.

Minervini, M., Mergy, M., Zhu, Y., Gutierrez Diaz, M.A., Pointer, C., Shinkazh, O., Oppenheim, S.F., Cramer, S.M., Przybycien, T.M., Zydney, A.L., 2024. Continuous precipitation-filtration process for initial capture of a monoclonal antibody product using a four-stage countercurrent hollow fiber membrane washing step. Biotechnol. Bioeng. 121 (8), 2258–2268.

Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. Technometrics 33 (2), 161–174.

Muramatsu, H., Lam, K., Bajusz, C., Laczkó, D., Karikó, K., Schreiner, P., Martin, A., Lutwyche, P., Heyes, J., Pardi, N., 2022. Lyophilization provides long-term stability for a lipid nanoparticle-formulated, nucleoside-modified mRNA vaccine. Mol. Ther. 30 (5), 1941–1951.

Nakamura, K., Aihara, K., Ishida, T., 2022. Importance of process parameters influencing the mean diameters of siRNA-containing lipid nanoparticles (LNPs) on the in vitro activity of prepared LNPs. Biol. Pharm. Bull. 45 (4), 497–507.

Napadensky, B., Shinkazh, O., Teella, A., Zydney, A.L., 2013. Continuous countercurrent tangential chromatography for monoclonal antibody purification. Sep. Sci. Technol. 48 (9), 1289–1297.

Nourafkan, E., Kenyon, C., Nair, A., Loveday, K.A., Welbourne, E.N., Tao, M., Ahmed, M., Middleton, J., Dickman, M.J., Brown, S.F., et al., 2024. An experimental and modeling approach to study tangential flow filtration performance for mRNA drug substance purification. Biotechnol. J. 19 (11), e202400473.

Park, S.-Y., Park, C.-H., Choi, D.-H., Hong, J.K., Lee, D.-Y., 2021. Bioprocess digital twins of mammalian cell culture for advanced biomanufacturing. Curr. Opin. Chem. Eng. 33, 100702.

Petersen, D.M.S., Chaudhary, N., Arral, M.L., Weiss, R.M., Whitehead, K.A., 2023. The mixing method used to formulate lipid nanoparticles affects mRNA delivery efficacy and organ tropism. Eur. J. Pharmaceut. Biopharmaceut. 192, 126–135.

Qu, J., Nair, A., Muir, G.W., Loveday, K.A., Yang, Z., Nourafkan, E., Welbourne, E.N., Maamra, M., Dickman, M.J., Kis, Z., 2024. Quality by design for mRNA platform purification based on continuous oligo-dt chromatography. Mol. Ther. Nucleic Acids 35 (4), 102333.

Ramkrishna, D., 2000. Population Balances: Theory and Applications To Particulate Systems in Engineering. Elsevier.

Ripoll, M., Martin, E., Enot, M., Robbe, O., Rapisarda, C., Nicolai, M.-C., Deliot, A., Tabeling, P., Authelin, J.-R., Nakach, M., Wils, P., 2022. Optimal self-assembly of lipid nanoparticles (LNP) in a ring micromixer. Sci. Rep. 12 (1), 9483.

Rosa, S.S., Prazeres, D.M., Azevedo, A.M., Marques, M.P., 2021. MRNA vaccines manufacturing: Challenges and bottlenecks. Vaccine 39 (16), 2190–2200.

Schmidt, A., Helgers, H., Vetter, F.L., Juckers, A., Strube, J., 2021. Digital twin of mRNA-based SARS-COVID-19 vaccine manufacturing towards autonomous operation for improvements in speed, scale, robustness, flexibility and real-time release testing. Process. 9 (5), 748.

Schwaminger, S.P., Zimmermann, I., Berensmeier, S., 2022. Current research approaches in downstream processing of pharmaceutically relevant proteins. Curr. Opin. Biotechnol. 77, 102768.

Scorza, F.B., Wen, Y., Geall, A., Porter, F., 2021. RNA Purification Methods. US Patent US20210214388A1, July 15.

Shin, S., Devos, C., Udepurkar, A.P., Inguva, P.K., Myerson, A.S., Braatz, R.D., 2025. Mechanistic modeling of lipid nanoparticle (LNP) precipitation via population balance equations (PBEs). Chem. Eng. J. 523, 167786.

Srisuma, P., Barbastathis, G., Braatz, R.D., 2024. Mechanistic modeling and analysis of thermal radiation in conventional, microwave-assisted, and hybrid freeze drying for biopharmaceutical manufacturing. Int. J. Heat Mass Transfer 221, 125023.

Srisuma, P., Chen, G., Braatz, R.D., 2025. Mechanistic modeling of continuous lyophilization for pharmaceutical manufacturing. Adv. Sci. e11693.

Stover, N.M., Ganko, K., Braatz, R.D., 2024. Mechanistic modeling of in vitro transcription incorporating effects of magnesium pyrophosphate crystallization. Biotechnol. Bioeng. 121 (9), 2636–2647.

Van Reis, R., Zydney, A., 2001. Membrane separations in biotechnology. Curr. Opin. Biotechnol. 12 (2), 208–211.

Velardi, S.A., Barresi, A.A., 2008. Development of simplified models for the freeze-drying process and investigation of the optimal operating conditions. Chem. Eng. Res. Des. 86 (1), 9–22.

Von Der Mülbe, F., Reidel, L., Ketterer, T., Gontcharova, L., Bauer, S., Pascolo, S., Probst, J., Schmid, A., 2018. Method for Producing RNA. US Patent US10017826B2, July 10.

Wang, R., Zhang, Y., Zhong, H., Zang, J., Wang, W., Cheng, H., Chen, Y., Ouyang, D., 2025. Understanding the self-assembly and molecular structure of mRNA lipid nanoparticles at real size: Insights from the ultra-large-scale simulation. Int. J. Pharm. 670, 125114.

Whitley, J., Zwolinski, C., Denis, C., Maughan, M., Hayles, L., Clarke, D., Snare, M., Liao, H., Chiou, S., Marmura, T., et al., 2022. Development of mRNA manufacturing for vaccines and therapeutics: mRNA platform requirements and development of a scalable production process to support early phase clinical trials. Transl. Res. 242, 38–55.

Wu, W., Oliveira, L.T., Jain, A., Karpov, Y., Olsen, K., Wu, Y., Panicker, R.K.G., 2025. Process development of tangential flow filtration and sterile filtration for manufacturing of mRNA-lipid nanoparticles: A study on membrane performance and filtration modeling. Int. J. Pharm. 675, 125520.

Ziegenhals, T., Frieling, R., Wolf, P., Göbel, K., Koch, S., Lohmann, M., Baiersdörfer, M., Fesser, S., Sahin, U., Kuhn, A.N., 2023. Formation of dsRNA by-products during in vitro transcription can be reduced by using low steady-state levels of UTP. Front. Mol. Biosci. 10, 1291045.

Zydney, A.L., 2016. Continuous downstream processing for high value biological products: A review. Biotechnol. Bioeng. 113 (3), 465–475.