

# Mechanistic Modeling of In Vitro Transcription

Nathan Merica Stover, Krystian Ganko, Richard D. Braatz\*

stover@mit.edu, kkganko@mit.edu, braatz@mit.edu

\* Corresponding author

Phone: 617.253.3112

Fax: 617.258.5042

E19-551, 77 Massachusetts Avenue, Massachusetts Institute of Technology,  
Cambridge, MA 02139

August 24, 2023

**Funding Statement:** Funding Statement: This research was supported by the U.S. Food and Drug Administration under the FDA BAA-22-00123 program, Award Number 75F40122C00200. This research was in addition supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Department of Energy Computational Science Graduate Fellowship under Award Number DE-SC0022158.

**Data and Code Availability:** All software and data will be made available upon publication of the manuscript.

**Conflict of Interest Disclosure:** 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.

## Abstract

The in vitro transcription (IVT) reaction used in the production of mRNA vaccines and therapies remains poorly quantitatively understood. Mechanistic modeling of IVT could inform reaction design, scale up, optimization, and control. In this work, we develop a mechanistic model of IVT to include nucleation and growth of magnesium pyrophosphate crystals and subsequent agglomeration of crystals and DNA. A novel quantitative description is included for the rate of transcription as a function of target sequence length, DNA concentration, and T7 polymerase concentration. The model explains previously unexplained trends in IVT data and quantitatively predicts the effect of adding the pyrophosphatase enzyme to the reaction system. The model is validated on additional literature data showing an ability to predict transcription rates as a function of RNA sequence length.

**Keywords:** Cell-free Synthesis, mRNA Production, In Vitro Transcription, Mechanistic Modeling, Pyrophosphatase

## 1 Introduction

In recent years, mRNA-based vaccines and immunotherapies have shown clinical efficacy for COVID-19, seasonal influenza, Epstein-Barr virus, HIV, and some forms of cancer (Barbier et al., 2022). In addition, mRNA is a promising method of delivering therapeutic proteins, with several mRNA therapies in the process of early-stage clinical trials (Rohner et al., 2022). Producing mRNA vaccines at the scale needed for quickly immunizing populations, however, remains a challenge (Kis et al., 2021). Also, while mRNA therapies are targeted toward small population groups, 50–1000 times greater dosages are required than for mRNA vaccines, which adds to manufacturing costs (Barbier et al., 2022). Due to the broad reach of the mRNA platform, even modest advances in efficiency and quality control of mRNA production would have a significant impact on the availability of a wide variety of therapies. Consumption of reagents for the in vitro transcription (IVT) reaction used for RNA synthesis is a key source of cost of goods (Rosa et al., 2021). Mechanistic modeling of this biomanufacturing process can be useful to organise existing data, understand the dynamics of key processes, and design novel reaction schemes and reactors (Destro & Barolo, 2022; Hong et al., 2018).

The quantitative effect of pyrophosphatase, an enzyme that degrades the pyrophosphate (PPi) byproduct of IVT and is heuristically included in most IVT schemes, remains unclear. While the mechanism of action and rate that pyrophosphatase catalyses the degradation of pyrophosphate is well studied, its effect on the IVT system is poorly understood, and it remains unclear why PPi needs to be removed from the IVT system. Owing to this lack of mechanistic understanding, there remains disagreement as to whether pyrophosphatase is even useful for increasing IVT yields at all. Previous researchers have reported pyrophosphatase to both be one of the most important components in their reaction schemes (Akama et al., 2012; Rosa et al., 2022), or to have no effect on yields (Kanwal et al., 2018; Samnuan et al., 2021). One phenomenon of interest to mechanistic modeling of IVT is the crystallization of the pyrophosphate byproduct in the form of magnesium pyrophosphate ( $\text{Mg}_2\text{PPi}$ ). This process has been associated with decreased yields, but no mechanistic model has been published that can describe major trends in these data (Akama et al., 2012; Young et al., 1997).

In parallel to research on manufacturing of RNA, recent work in the design of biological nanostructures for potential therapeutic and diagnostic uses have incidentally increased understanding of IVT. It has become well understood that  $\text{Mg}_2\text{PPi}$  crystallization in the presence of DNA forms  $\text{Mg}_2\text{PPi}$ -DNA composites that remove DNA from solution (Kim et al., 2019; Kim et al., 2017). In addition,  $\text{Mg}_2\text{PPi}$ -RNA composites have been observed and studied in conditions where crystallization of  $\text{Mg}_2\text{PPi}$  occurs in the presence of RNA (Baker et al., 2018;

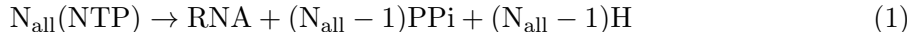
Shopsowitz et al., 2014). These DNA and RNA nanostructures have primarily been investigated with the goal of identifying synthesis-structure relationships for engineering of delivery or diagnostic platforms. However, the knowledge that  $\text{Mg}_2\text{PPi}$  solid formation has this effect on both an essential reagent and the product of the reaction has profound implications for engineering and optimization of IVT. This phenomenon has never been quantitatively modeled as part of a larger system for understanding the kinetics of IVT.

An additional unmet challenge in mechanistic modeling of IVT is to develop generalizable models for arbitrary target RNA sequences. While predicting complex phenomena such as RNA secondary structure remains a grand challenge, incorporating simple characteristics like sequence length is a straightforward first step in developing generally applicable and easily translatable IVT models. Most previous work in applying mechanistic models to IVT data has ignored the effect of sequence length in both model development and data collection (Akama et al., 2012; van de Berg et al., 2021; Young et al., 1997), and the most complete past work in incorporating sequence length into mechanistic models of IVT was restricted by the limited data and fundamental understanding of the transcription process available at the time (Arnold et al., 2001).

In this work, a mechanistic model is developed for IVT that incorporates new quantitative descriptions of crystallization of magnesium pyrophosphate, the sequestration of DNA due to crystal formation, and the degradation of pyrophosphate by pyrophosphatase. In order to generalize this model across multiple target RNA sequences of different lengths, our transcription rate law incorporates descriptions of both initiation and elongation steps. This mechanistic model is fit to a literature dataset that is unique in the published literature in measuring the dynamics of pyrophosphate concentration (Akama et al., 2012). The inclusion of these new phenomena into the model enables it to capture of important trends of this dataset. In addition, the model quantitatively predicts the effect of adding pyrophosphatase to the IVT system. Our model for the IVT rate accurately predicts the effect of sequence length on the IVT rate measured for an independent set of experiments (Rosa et al., 2022).

## 2 Review of Past Models of IVT Reactions

The primary process of IVT is the polymerization of RNA from nucleoside triphosphate (NTP) monomers, which has the overall stoichiometry



where  $\text{N}_{\text{all}}$  is the number of NTPs incorporated into each RNA chain. The reaction forms pyrophosphate (PPi) and proton (H) byproducts. A typical IVT scheme requires a linearized template DNA of the target sequence, nucleoside triphosphates, T7 RNA Polymerase (T7 RNAP), and a magnesium salt in an aqueous buffered reaction at 37°C and a pH around 7.5–8 (Beckert & Masquida, 2011). In addition, many IVT reaction schemes include pyrophosphatase, surfactants, spermadine, and dithiothreitol. However, past mechanistic models of IVT have only focused on modeling the concentrations and effects of NTPs, T7RNAP, and Mg, and there is little to no public data describing the effect on IVT by the latter set of components.

In addition to the transcription reaction, past mechanistic models for IVT have included a number of secondary processes based on experimental observations and first principles. First, a network of equilibrium reactions between free species concentrations and complexes such as  $\text{MgNTP}$  are described using a series of algebraic relations (Kern & Davis, 1997). In addition, past mechanistic models have included additional kinetic phenomena, including  $\text{Mg}_2\text{PPi}$  crystallization, RNA degradation, and T7RNAP degradation (Akama et al., 2012; van de Berg

et al., 2021; Young et al., 1997). While the latter two of these phenomena were introduced to help conform IVT models to individual datasets and are not directly observed in the context of IVT,  $\text{Mg}_2\text{PPi}$  crystallization is a confirmed phenomenon that is easily reproduced owing to the visibility of solid formation.

Past mechanistic models have focused on isolated operating regimes and design spaces of the IVT reaction due to a diversity of goals involved. The first mechanistic model for describing trajectories of solution concentrations in the IVT reaction was primarily focused on empirically modeling experimental data (Young et al., 1997). This work uniquely focused on modeling the presence of aborts, which are short transcription sequences that do not match the desired full sequence. A later work (Arnold et al., 2001) developed a mechanistic model of IVT with the goal of deriving rate expressions from first principles of the known biochemistry of IVT. This work is unique in quantitatively including initiation, elongation, and termination of the RNA polymerization process into an IVT mechanistic model, and in including quantitative descriptions of the effect of DNA concentration on IVT rate. Another study (Akama et al., 2012) developed a mechanistic model to describe IVT in tandem with  $\text{Mg}_2\text{PPi}$  crystallization. A recent study (van de Berg et al., 2021) built on two past models (Akama et al., 2012; Young et al., 1997) to fit a dataset collected for a range of operating conditions.

The most comprehensive published dataset on  $\text{Mg}_2\text{PPi}$  crystallization in IVT is by Akama and coworkers (Akama et al., 2012), which is referred to as the *Akama dataset* in this work. This dataset includes the temporal evolution of both RNA and PPi concentration, which is unique among published datasets. Despite the high quality and relevance of the Akama dataset, no publications (not even Akama et al.) have fit these temporal reaction trajectories to a mechanistic model. Our mechanistic model, which is described in the next section, is fit to the Akama dataset.

### 3 Materials and Methods

#### 3.1 Mechanistic Model Formulation

Our mechanistic model uses a set of differential equations

$$\frac{d[\text{DNA}]}{dt} = -V_{\text{sequestration}} \quad (2)$$

$$\frac{d[\text{RNA}]}{dt} = V_{\text{tr}} \quad (3)$$

$$\frac{d[\text{PPi}]_{\text{tot}}}{dt} = (N_{\text{all}} - 1)V_{\text{tr}} - V_{\text{solid}} - V_{\text{PPiase}} \quad (4)$$

$$\frac{d[\text{NTP}]_{\text{tot}}}{dt} = -N_{\text{all}}V_{\text{tr}} \quad (5)$$

$$\frac{d[\text{H}]_{\text{tot}}}{dt} = (N_{\text{all}} - 1)V_{\text{tr}} \quad (6)$$

$$\frac{d[\text{Mg}]_{\text{tot}}}{dt} = -2V_{\text{solid}} \quad (7)$$

$$\frac{d[\text{Nuc}]}{dt} = V_{\text{nuc}} \quad (8)$$

$$\frac{d[\text{Pi}]_{\text{tot}}}{dt} = 2V_{\text{PPiase}} \quad (9)$$

which track the temporal evolution of eight state variables representing the total concentration of species in the reaction: DNA, RNA, pyrophosphate (PPi), NTP, H, Mg, phosphate (Pi), and  $\text{Mg}_2\text{PPi}$  nuclei (Figure 1). The model contains five kinetic processes: the transcription reaction

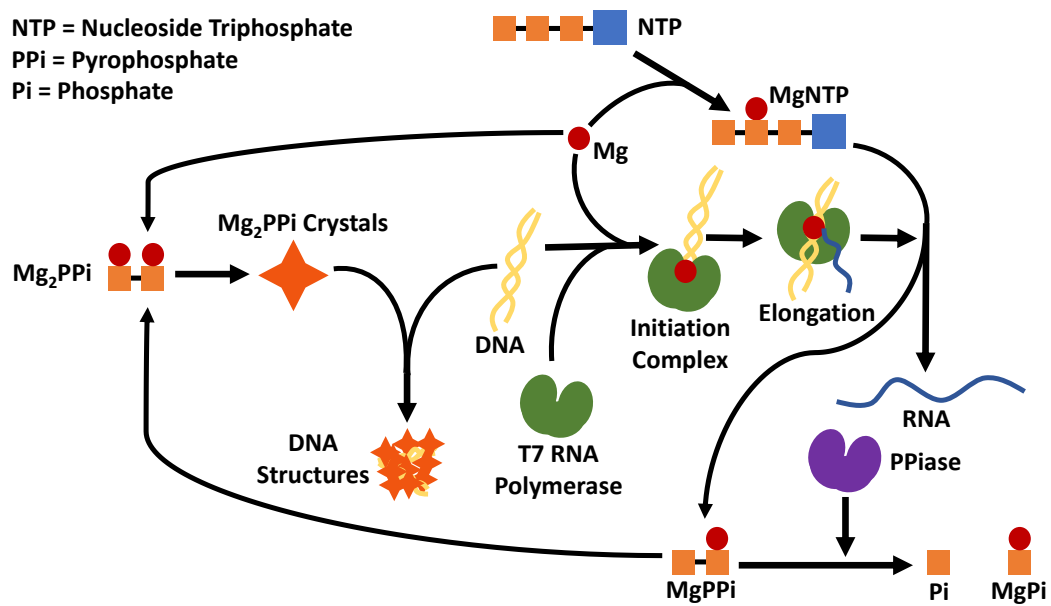


Figure 1: Major species in the IVT reaction model. The elongation of mRNA chains produces pyrophosphate byproduct. This byproduct can complex with magnesium to form solid crystals, which sequester DNA, inhibiting transcription. The pyrophosphatase enzyme inhibits the formation of crystals by decreasing the free concentration of pyrophosphate.

( $V_{\text{tr}}$ ), nucleation and growth of  $\text{Mg}_2\text{PPi}$  crystals ( $V_{\text{nuc}}$ ,  $V_{\text{solid}}$ ), agglomeration of  $\text{Mg}_2\text{PPi}$  nuclei and DNA ( $V_{\text{sequestration}}$ ), and the degradation of  $\text{PPi}$  by pyrophosphatase ( $V_{\text{PPiase}}$ ).

The transcription rate was modeled as a process of reversible binding of T7 RNA polymerase and DNA promoter, coupled with an irreversible initiation step and an elongation step dependent on sequence length (SI Section 1),

$$V_{\text{tr}} = k_i[\text{T7RNAP} \cdot \text{DNA}] \left( \frac{[\text{MgNTP}]}{K_1(1 + \frac{[\text{MgPPi}]}{K_{i,\text{PPi}}}) + [\text{MgNTP}]} \right) \left( \frac{[\text{Mg}]}{K_2 + [\text{Mg}]} \right), \quad (10)$$

where the concentration of polymerase-DNA complex is given by

$$[\text{T7RNAP} \cdot \text{DNA}] = \frac{[\text{T7RNAP}] + \alpha[\text{DNA}] + K_{\text{MD}} - \sqrt{([\text{T7RNAP}] + \alpha[\text{DNA}] + K_{\text{MD}})^2 - 4\alpha[\text{T7RNAP}][\text{DNA}]}}{2\alpha}, \quad (11)$$

$$\alpha = 1 + \frac{k_i N_{\text{all}}}{k_e}, \quad K_{\text{MD}} = \frac{k_i \left( \frac{[\text{MgNTP}]}{K_1(1 + \frac{[\text{MgPPi}]}{K_{i,\text{PPi}}}) + [\text{MgNTP}]} \right) \left( \frac{[\text{Mg}]}{K_2 + [\text{Mg}]} \right) + k_{\text{off}}}{k_{\text{on}}}, \quad (12)$$

and  $k_{\text{on}}$ ,  $k_{\text{off}}$ ,  $k_i$ , and  $k_e$  are rate constants for T7 RNA polymerase and the DNA promoter binding, unbinding, initiating transcription, and elongating a RNA transcript, respectively (SI Section 2). Whereas Akama et al. (2012) modeled the rate of  $\text{Mg}_2\text{PPi}$  solid formation using empirical induction time models, we model the crystallization using common equations from classical nucleation theory (Bergwerff & van Paassen, 2021)

$$\hat{V}_{\text{nuc}} = \begin{cases} \hat{k}_{\text{nuc}} \exp\left(\frac{-\hat{B}}{\ln^2 S}\right), & \text{for } S > 1 \\ 0, & \text{for } S \leq 1 \end{cases} \quad (13)$$

where  $S$  is the supersaturation,

$$S = \frac{[\text{Mg}_2\text{PPi}]}{[\text{Mg}_2\text{PPi}]_{\text{eq}}}, \quad (14)$$

and  $\hat{B}$  and  $\hat{k}_{\text{nuc}}$  are the dimensionless free energy barrier to nucleation and the nucleation rate constant, respectively. The total rate of solid formation of  $\text{PPi}$  (in mol/L-hr) is

$$\hat{V}_{\text{solid}} = \begin{cases} \hat{k}_g [\widehat{\text{Nuc}}] (S - 1), & \text{for } S > 1 \\ 0, & \text{for } S \leq 1 \end{cases} \quad (15)$$

where  $\hat{k}_g$  is a rate constant governing the growth of nuclei as a function of nuclei concentration,  $[\widehat{\text{Nuc}}]$ . The number of model parameters can be reduced, by some algebraic manipulation (see SI Section 3), to give

$$V_{\text{nuc}} = \begin{cases} \exp\left(\frac{-B}{\ln^2 S}\right), & \text{for } S > 1 \\ 0, & \text{for } S \leq 1 \end{cases} \quad (16)$$

$$V_{\text{solid}} = \begin{cases} k_{\text{growth}} [\text{Nuc}] (S - 1), & \text{for } S > 1 \\ 0, & \text{for } S \leq 1 \end{cases} \quad (17)$$

where  $k_{\text{growth}}$  and  $B$  are fitted parameters and  $[\text{Nuc}]$  is the rate-normalized concentration of nuclei. In addition, based on qualitative work demonstrating the agglomeration of  $\text{Mg}_2\text{PPi}$  crystals and DNA (Kim et al., 2019), a term was included to describe the rate of DNA sequestration in the solid phase,

$$V_{\text{sequestration}} = k_d[\text{DNA}][\text{Nuc}], \quad (18)$$

which hypothesizes that the rate is first order in both DNA concentration and rate-normalized nuclei concentration with a rate constant  $k_d$ . Past work has qualitatively shown that a similar phenomenon takes place in sequestering RNA (Shopsowitz et al., 2014). However, the experimental procedure used by Akama et al. (2012) re-dissolved any solid before measuring RNA concentrations, meaning that any RNA sequestration cannot be observed from the Akama dataset.

Enzymatic degradation of PPi is modeled by

$$V_{\text{PPiase}} = k_{\text{PPiase}}[\text{PPiase}] \frac{[\text{MgPPi}]}{[\text{MgPPi}] + K_{\text{M,PPiase}}}, \quad (19)$$

with the rate law and parameters from a kinetic study of pyrophosphatase (Chao et al., 2006), where  $[\text{PPiase}]$  is in units of volume-based enzyme activity ( $\text{U}/\mu\text{L}$ ).

The above rates are dependent on the concentration of complexes like  $\text{MgNTP}$  and  $\text{Mg}_2\text{PPi}$ . The concentrations of these complexes over time are modeled by a set of algebraic equations that describe known equilibrium relations and material balances of the system. The material balances for the ionic species are

$$[\text{Mg}]_{\text{tot}} = [\text{Mg}] + [\text{MgPPi}] + [\text{HMgPPi}] + [\text{MgNTP}] + 2[\text{Mg}_2\text{NTP}] + 2[\text{Mg}_2\text{PPi}] + [\text{HMgNTP}] + [\text{H}_2\text{MgPPi}] + [\text{MgPi}] \quad (20)$$

$$[\text{NTP}]_{\text{tot}} = [\text{NTP}] + [\text{HNTP}] + [\text{HMgNTP}] + [\text{MgNTP}] + [\text{Mg}_2\text{NTP}] \quad (21)$$

$$[\text{H}]_{\text{tot}} = [\text{H}] + [\text{HNTP}] + [\text{HMgNTP}] + [\text{HPPi}] + [\text{HMgPPi}] + 2[\text{H}_2\text{PPi}] + 2[\text{H}_2\text{MgPPi}] + [\text{HBuffer}] \quad (22)$$

$$[\text{PPi}]_{\text{tot}} = [\text{PPi}] + [\text{MgPPi}] + [\text{Mg}_2\text{PPi}] + [\text{HPPi}] + [\text{HMgPPi}] + [\text{H}_2\text{PPi}] + [\text{H}_2\text{MgPPi}] \quad (23)$$

$$[\text{Buffer}]_{\text{tot}} = [\text{Buffer}] + [\text{HBuffer}] \quad (24)$$

$$[\text{Pi}]_{\text{tot}} = [\text{Pi}] + [\text{MgPi}] \quad (25)$$

where the complex concentrations are defined by the equilibrium relations

$$[\text{HNTP}] = [\text{H}][\text{NTP}]K_{\text{HNTP}} \quad (26)$$

$$[\text{HMgNTP}] = [\text{HNTP}][\text{Mg}]K_{\text{HMgNTP}} \quad (27)$$

$$[\text{HPPi}] = [\text{H}][\text{PPi}]K_{\text{HPPi}} \quad (28)$$

$$[\text{HMgPPi}] = [\text{HPPi}][\text{Mg}]K_{\text{HMgPPi}} \quad (29)$$

$$[\text{H}_2\text{PPi}] = [\text{HPPi}][\text{H}]K_{\text{H}_2\text{PPi}} \quad (30)$$

$$[\text{H}_2\text{MgPPi}] = [\text{H}_2\text{PPi}][\text{Mg}]K_{\text{H}_2\text{MgPPi}} \quad (31)$$

$$[\text{HBuffer}] = [\text{H}][\text{Buffer}]K_{\text{Buffer}} \quad (32)$$

$$[\text{MgNTP}] = [\text{Mg}][\text{NTP}]K_{\text{MgNTP}} \quad (33)$$

$$[\text{Mg}_2\text{NTP}] = [\text{MgNTP}][\text{Mg}]K_{\text{Mg}_2\text{NTP}} \quad (34)$$

$$[\text{MgPPi}] = [\text{Mg}][\text{PPi}]K_{\text{MgPPi}} \quad (35)$$

$$[\text{Mg}_2\text{PPi}] = [\text{MgPPi}][\text{Mg}]K_{\text{Mg}_2\text{PPi}} \quad (36)$$

$$[\text{MgPi}] = [\text{Mg}][\text{Pi}]K_{\text{MgPi}} \quad (37)$$

As for all first-principles models of complex reactions, some assumptions and simplifications are made: (1) The transcription rate law (10) ignores the effect of non-coding and coding sequence identity. (2) While past work has modeled each of the four NTPs as separate components (Arnold et al., 2001; Young et al., 1997), this level of detail is outside the scope of this work due to being not relevant to modeling the results of Akama et al. (2012), which used equal initial concentrations of each NTP and an RNA sequence with roughly equal amounts of each nucleotide. (3) Product quality variables, such as the presence of aborts and double-stranded RNA, are not considered in this work, as the literature data describing these byproducts are sparse. (4) The pH of the system was modeled by a proton balance (22) rather than a charge balance due to incomplete knowledge about the full population of ionic species in the reaction. The proton balance can be used because effects of pH are a minor portion of this model. (5) Degradation of RNA and T7 RNA polymerase are not considered as those effects were not essential for capturing the dynamics of the Akama dataset. (6) The Michaelis-Menten description of pyrophosphatase action is a simplification of a more sophisticated network of reversible and irreversible reactions (Halonen et al., 2002; Tammenkoski et al., 2007). In addition, the rate law used in this work has only been shown for pyrophosphatase from *Helicobacter pylori*, which is not commonly used in IVT (Chao et al., 2006). However, as the PPi concentrations in the Akama data are relatively high, the most important part of this rate law is the maximum rate, which is quantitatively well understood and captured by such a simple model. (7) The nucleation-growth model does not take into account the effects of size heterogeneity of  $\text{Mg}_2\text{PPi}$  crystals and the contribution to total solid formation of the nuclei formation step. (8) Crystallization of magnesium hydrogen phosphate ( $\text{MgHPI}$ ), which has been postulated as an additional process in the IVT system (Kim et al., 2019), is not considered.

### 3.2 Computational Methods

Model evaluation and parameter optimization are performed in the Julia language. The set of equations in the preceding section is combined into a system of differential algebraic equations (DAEs) which is solved forward in time using the high-order integrators available in the DifferentialEquations.jl package. Experimental measurements are assumed to have additive, uncorrelated measurement error with a normal distribution of zero mean and diagonal measurement error covariance matrix  $V_y$ . Parameter estimation is performed in  $\log_{10}$  space, to search the large numerical space and to best represent the prior distribution of parameters, where the vector  $k$  represents the  $\log_{10}$  of the parameters. The prior distribution for  $k$  is assumed to follow a normal distribution with mean  $\mu$  and covariance  $V_\mu$ , which is equivalent to assuming a log-normal distribution of parameters.

MAP estimation of the vector  $k$  was carried out,

$$\min_k (y - u(k))^T V_y^{-1} (y - u(k)) + (k - \mu)^T V_\mu^{-1} (k - \mu), \quad (38)$$

where  $y$  is the vector containing all of the experimental data used for estimating parameters and  $u(k)$  is the vector of corresponding model outputs as a function of the  $\log_{10}$  parameter vector  $k$ . The error covariance matrix  $V_{k^*}$  of the best-fit estimate  $k^*$  is approximated by (Beck & Arnold, 1977)

$$\text{cov}(k^* - k_{\text{true}}) = V_{k^*} \approx (S^T V_y^{-1} S + V_\mu)^{-1}, \quad (39)$$

where  $k_{\text{true}}$  denotes the true  $\log_{10}$  of the parameters and  $S$  is the sensitivity of the model outputs with respect to the vector  $k$ . Additional details on the parameter estimation strategy can be found in SI Section 5.

Local gradient-based optimization is carried out with L-BFGS optimization using the ForwardDiff.jl and NLOpt.jl packages in Julia to compute model output sensitivities to parameters



and emded those sensitivities into the gradient-based optimizers, respectively (Liu & Nocedal, 1989). Multistart optimization using 4000 random starting points is performed to search for a global optimum (Martí, 2003). Best-fit parameter estimates in  $k^*$  are given in Table 1, and the parameter error covariance matrix  $V_{k^*}$  is given in SI Section 10.

Parameter	Units	Process	Prior Value	Value after Fitting
$k_i$	$h^{-1}$	transcription initiation prior	$10^{2.97 \pm 0.6}$ (Koh et al., 2018)	$10^{3.93 \pm 0.09}$
$k_e$	$h^{-1}$	transcription elongation prior	$10^{5.72 \pm 0.3}$ (Tang et al., 2011)	$10^{5.13 \pm 0.05}$
$k_{off}$	$h^{-1}$	T7RNAP-DNA binding prior	$10^{3.64 \pm 0.5}$ (Koh et al., 2018)	$10^{3.71 \pm 0.45}$
$k_{on}$	$h^{-1}$	T7RNAP-DNA binding prior	$10^{2.31 \pm 0.1}$ (Koh et al., 2018)	$10^{2.30 \pm 0.10}$
$K_1$	M	transcription MgNTP dependence	—	$10^{-3.09 \pm 0.16}$
$K_2$	M	transcription Mg dependence	—	$10^{-3.80 \pm 0.18}$
$K_{i,PPi}$	M	transcription PPi inhibition prior	$10^{-3.70 \pm 0.6}$ (Arnold et al., 2001)	$10^{-4.50 \pm 0.17}$
$k_{growth}$	moles/h	Mg <sub>2</sub> PPi solid growth	—	$10^{-0.26 \pm 0.38}$
$B$	arb. unit	Mg <sub>2</sub> PPi solid growth prior	$10^{1.13 \pm 0.2}$ (Akama et al., 2012)	$10^{1.65 \pm 0.10}$
$k_d$	$h^{-1}M^{-1}$	DNA-Mg <sub>2</sub> PPi agglomeration	—	$10^{4.47 \pm 0.40}$
$K_{HNTP}$	$M^{-1}$	ion equilibrium prior	$10^{6.91 \pm 1.16}$	$10^{6.91 \pm 0.04}$
$K_{HMgNTP}$	$M^{-1}$	ion equilibrium prior	$10^{2.08 \pm 1.16}$	$10^{2.08 \pm 0.10}$
$K_{HPPi}$	$M^{-1}$	ion equilibrium prior	$10^{9.02 \pm 1.16}$	$10^{9.03 \pm 0.10}$
$K_{HMgPPi}$	$M^{-1}$	ion equilibrium prior	$10^{3.32 \pm 1.16}$	$10^{3.30 \pm 0.10}$
$K_{H_2PPi}$	$M^{-1}$	ion equilibrium prior	$10^{6.26 \pm 1.16}$	$10^{6.91 \pm 0.20}$
$K_{H_2MgPPi}$	$M^{-1}$	ion equilibrium prior	$10^{2.11 \pm 1.16}$	$10^{2.08 \pm 0.20}$
$K_{MgNTP}$	$M^{-1}$	ion equilibrium prior	$10^{4.54 \pm 1.16}$	$10^{4.22 \pm 0.14}$
$K_{Mg_2NTP}$	$M^{-1}$	ion equilibrium prior	$10^{1.77 \pm 1.16}$	$10^{1.84 \pm 0.29}$
$K_{MgPPi}$	$M^{-1}$	ion equilibrium prior	$10^{4.80 \pm 1.16}$	$10^{5.23 \pm 0.15}$
$K_{Mg_2PPi}$	$M^{-1}$	ion equilibrium prior	$10^{2.57 \pm 1.16}$	$10^{3.99 \pm 0.12}$
$K_{MgPi}$	$M^{-1}$	ion equilibrium	$10^{1.88}$	$10^{1.88}$
$Mg_2PPi_{eq}$	M	ion equilibrium prior	$10^{-4.85 \pm 2}$ (Akama et al., 2012)	$10^{-3.95 \pm 0.09}$
$k_{PPiase}$	$\frac{M}{h^{-1}(U/\mu L)^{-1}}$	degradation of PPi	60 (SI Section 7)	60
$K_{M,PPiase}$	M	degradation of PPi	$2.14 \times 10^{-4}$ (Chao et al., 2006)	$2.14 \times 10^{-4}$

Table 1: Model Parameters. Strategy and sources for generating prior values of equilibrium constants are discussed in SI Section 6. Error on parameter priors represents 95% confidence interval using standard deviation estimated from literature. Error on parameter posteriors represents the 95% pointwise confidence intervals as approximated by drawing samples from the probability distribution defined by the parameter error covariance matrix. The unit U represents enzyme activity units, discussed in SI section 7.

## 4 Results

### 4.1 Fitting Model to Training Data

The batch IVT reaction model is fit to the dataset of Akama et al. (2012). The model features 24 parameters, which are fixed based on literature data, estimated from the calibration data using a Bayesian prior from the literature, or estimated from the calibration data without a prior (SI Section 5). The model demonstrates an ability to describe the trends of the calibration dataset. Pyrophosphate (PPi), a byproduct of transcription, initially grows rapidly as the transcription reaction progresses, but reaches a peak in concentration after which the PPi concentration decays rapidly due to competing  $\text{Mg}_2\text{PPi}$  solid formation (Figure 2D).  $\text{Mg}_2\text{PPi}$  solid formation is prevented in cases of high NTP and low Mg concentration (Figure 2EF). In cases of solid formation, the reaction halts before reaching full conversion of NTPs (Figure 2A-C). At low NTP concentrations, increasing NTP concentration increases initial transcription rates. However, at higher NTP concentrations, this effect reverses as addition of NTP decreases free Mg concentrations (Figure 2H). When small amounts of PPi (less than one equivalent of Mg) are added to aqueous Mg in the absence of the transcription reaction,  $\text{Mg}_2\text{PPi}$  solid formation decreases the concentration of Mg in solution after 24 hours (Figure 2G). The magnitude of this effect is decreased upon addition of NTP.

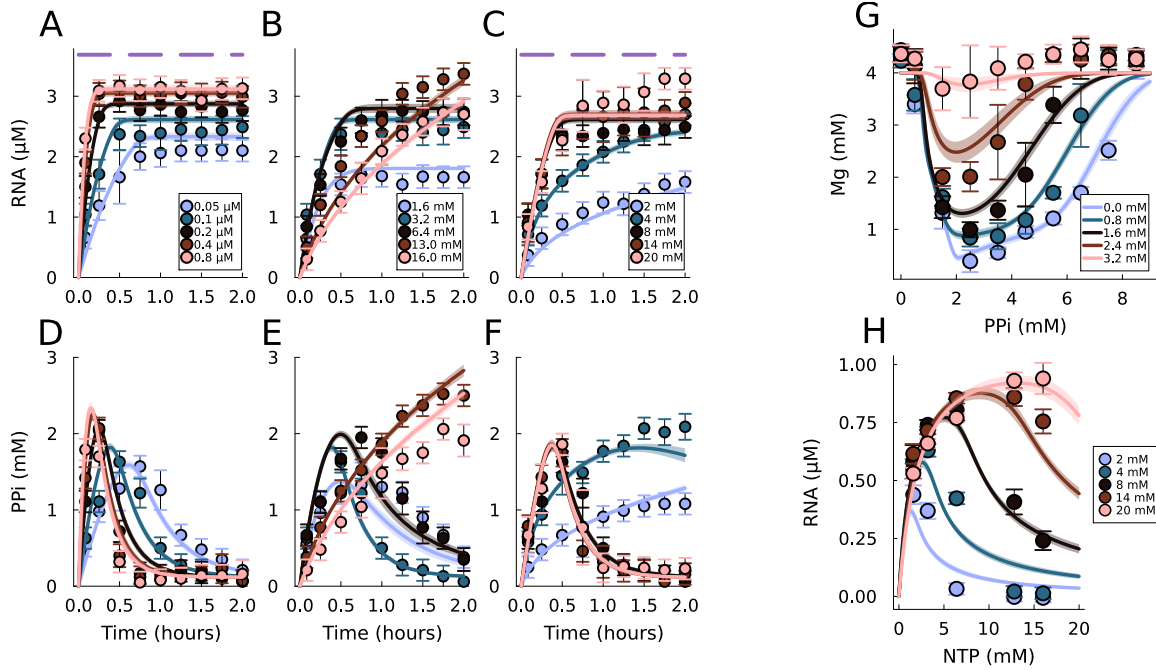


Figure 2: Model fitting results compared to fitting dataset. Temporal trajectories of RNA (A,B,C) and PPi (D,E,F) concentration as a function of changing T7 RNA Polymerase (A,D), NTP (B,E), and Mg (C,F) input concentrations. Magnesium (initially 4 mM) remaining in solution after 24 hours as a function of input PPi and NTP input concentrations in the absence of the transcription reaction (G). Initial transcription rates as a function of NTP and Mg input concentrations represented by RNA yields after 5 minutes of reaction (H). Purple dashed lines represent maximum possible mRNA yield based on stoichiometry. Shaded areas about each model prediction is the 95% prediction interval (SI Section 4).

## 4.2 Predicting Effect of Pyrophosphatase on the IVT System

The above training data are for experiments that did not include pyrophosphatase. In addition to these data, Akama et al. (2012) generated a small dataset on the effect of pyrophosphatase on reaction yields, which is used for model validation in this work (Figure 3). Akama and coworkers did not report the quantities of pyrophosphatase used but showed that, when pyrophosphatase was used, the concentration of PPi was indistinguishable from zero. Pyrophosphatase addition was set to an excess value of 1 U/ $\mu$ L in the model to predict these results. Our model predictions for the effect of pyrophosphatase on the IVT reaction are within the experimental error bars (Figure 3).

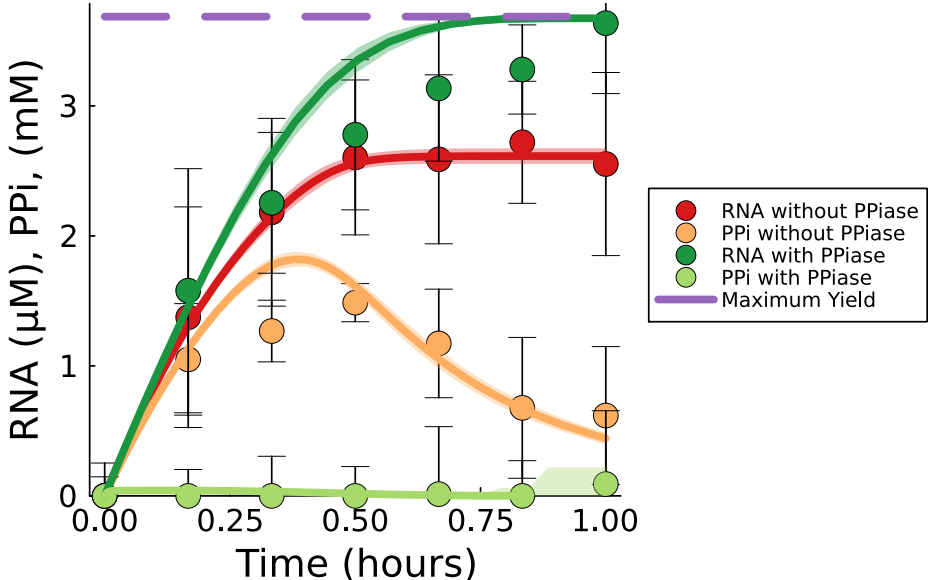


Figure 3: Model Validation: Effect of Pyrophosphatase. Model predictions are compared with experimental results of Akama et al. (2012) showing the yields of otherwise identical reaction conditions with and without pyrophosphatase (PPIase). Shaded areas about the model predictions are 95% prediction interval (SI Section 4). The error bar on each experimental data point is the 95% confidence interval based on the t-distribution of points taken in triplicate.

## 4.3 Predicting Effect of Sequence Length

Rosa et al. (2022) collected trajectories of IVT yields for a set of three DNA constructs varying in length between 846 and 4950 nucleotides (Figure 4), with excess (1 U/ $\mu$ L) pyrophosphate in the reaction. Our model is able to predict the dependency of the transcription rate on sequence length. This dataset was not used in fitting the model parameters.

## 5 Discussion

The primary result of this work is that adding a nucleation-growth model for  $\text{Mg}_2\text{PPi}$  crystallization, as well as a quantitative term describing the first-order agglomeration of DNA and  $\text{Mg}_2\text{PPi}$  nuclei, is a sufficient addition to past models to describe trends in experimental data. Our modeling of the additional phenomena is consistent with the qualitatively understood physics of the IVT system. This mechanism can additionally predict the effect of adding the pyrophosphatase enzyme on IVT yields, and is the first mechanistic model to do so. Validating

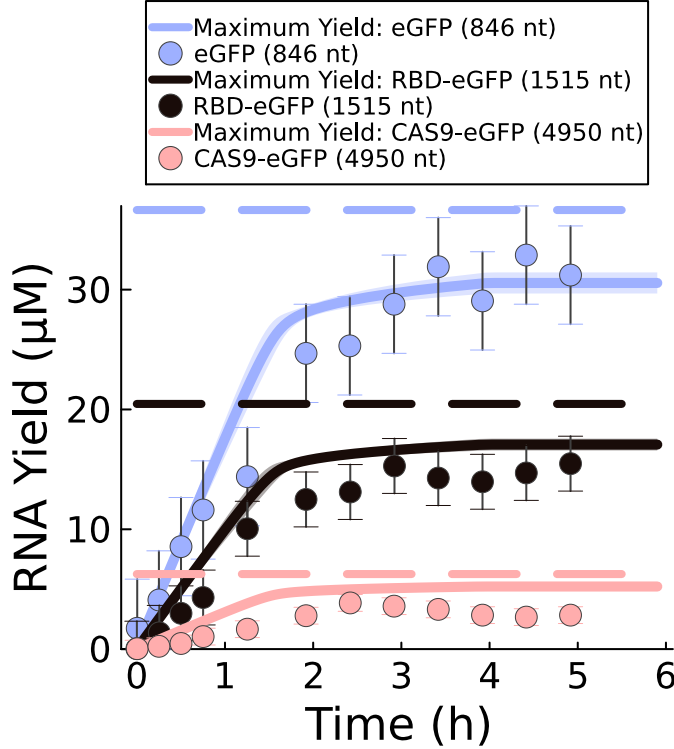


Figure 4: Model Validation: IVT yields of multiple DNA sequences of varying lengths (Rosa et al., 2022). Shaded areas about the model predictions are 95% prediction intervals (SI Section 4). The error bars on the data points are 95% confidence intervals based on the standard deviation estimated from the entire dataset of Rosa et al. (2022).

this model on recent data demonstrates an ability to predict IVT rates and yields across a range of input conditions and sequence lengths.

This work identifies  $\text{Mg}_2\text{PPi}$  solid formation as an important failure mode that is highly nonlinear, and the value of our model is its ability to predict these nonlinear effects (Figure 2). As such, our model is a suitable foundation for the development of model-based optimal design and control strategies. In addition, the incorporation of sequence length into model predictions is a first step for the development of models that are easily adaptable to arbitrary RNA sequences.

### 5.1 $\text{Mg}_2\text{PPi}$ Solid Formation is a Crucial Element for Describing IVT

Past experimental studies have shown that DNA agglomerates with  $\text{Mg}_2\text{PPi}$  nuclei during IVT, removing DNA from solution (Wang et al., 2019). The inclusion in the model of terms describing this sequestration of DNA (18) is able to describe trends in the calibration data, especially the early halting of reactions that do not go to full conversion (Figure 2). The reaction halting is roughly associated with a concurrent decrease in  $\text{PPi}$  concentration (Figure 2A-F).

When paired with ion equilibrium laws describing the known thermodynamics of the system, this model is able to describe the conditions at which solid formation, and therefore reaction halting, occur. At high NTP concentrations, NTP competes with  $\text{PPi}$  for the  $\text{Mg}$  ion, which leads to decreased  $\text{Mg}_2\text{PPi}$  solution concentrations and decreased solid formation (Figure 2G). This in turn prevents early halting of the reaction (Figure 2BE). By the same mechanism, solid formation and reaction halting are prevented at low  $\text{Mg}$  concentrations (Fig-

ure 2CF). The model also has the ability to describe competition between the irreversible kinetic processes of transcription and DNA sequestration. Increased T7 RNAP concentration increases the initial rate of reaction (Figure 2A); while this causes solid formation to initiate earlier (Figure 2D), it ultimately leads to higher yields.

This model clears up some misconceptions and mis-prescriptions in the academic literature. Past studies in the mechanistic modeling of IVT have noticed that the onset of crystallization is associated with a decrease or complete stoppage in transcription rates (Akama et al., 2012; Kern & Davis, 1997), but have attributed this stoppage to the onset of  $\text{Mg}_2\text{PPi}$  crystallization causing a decrease in magnesium concentration in the solution. As noted by Akama et al. (2012) and in this work, this description cannot explain trends in experimental data (SI Section 9). Our model uses an entirely different explanatory pathway to describe this phenomenon.

While this development may seem like an academic distinction, the true cause of early halting is highly relevant for IVT reaction optimization and control. The general message from past work that has been transmitted to practitioners is that, because  $\text{Mg}_2\text{PPi}$  solid formation decreases the solution concentration of Mg, reactions should be designed with high concentration of Mg to preempt this effect. The literature contains many reports in the last three years of academic researchers justifying IVT reaction schemes and explaining results based on this idea (Pregeljc et al., 2023; Rosa et al., 2022; Samnuan et al., 2021). While the higher order effects of magnesium on the IVT system remain poorly understood, one conclusion from this work is the decrease in free Mg concentration due to  $\text{Mg}_2\text{PPi}$  solid formation cannot quantitatively describe the early stopping of IVT as measured by Akama et al. (2012).

### 5.1.1 Model Describes Mechanism of Action of Pyrophosphatase

The pyrophosphatase enzyme, which degrades PPi, is commonly added to IVT reaction schemes on a heuristic basis. The dataset used to fit our model (Figure 2) did not include the use of the pyrophosphatase enzyme. Data from Akama et al. (2012) describing the effect of adding pyrophosphatase on IVT yields was used for model validation (Figure 3). When pyrophosphatase was added to the reaction, pyrophosphate was degraded to phosphate, preventing solid formation and sequestration of DNA as well as competitive inhibition of the transcription process (10). The quantitative predictions of the model that pyrophosphatase extends the length of the reaction without changing initial rates is consistent with the observed data (Figure 3).

As described in Section 4.2, the addition of pyrophosphatase (PPiase) to an IVT reaction system can lead to increases in reaction yield, depending on system inputs. However, despite the widespread adoption of PPiase based on heuristic observations, researchers provide conflicting explanations for the importance of both  $\text{Mg}_2\text{PPi}$  crystallization and PPiase to the IVT system, and many “rational” attempts at IVT optimization start by removing PPiase (Kanwal et al., 2018; Samnuan et al., 2021). The model developed in this work represents a first step towards uniting regimes sensitive and insensitive to PPiase and is suitable for guiding future optimization.

## 5.2 Mechanistic Model Predicts Effect of Sequence Length on Transcription Rates

In addition to the small Akama dataset describing the effect of pyrophosphatase (Figure 3), our model was validated on data from Rosa et al. (2022), which modulated sequence length as an independent variable. This dataset is outside of the input and output range of the calibration dataset used in this work (Table 2). Owing to the high concentrations of pyrophosphatase used, the model predicts that early stopping due to  $\text{Mg}_2\text{PPi}$  crystallization did not occur, which is consistent with observed data (Figure 4). Considering that it is heuristically understood

amongst practitioners that the parameters of the IVT reaction are sequence dependent, we do not argue that our model is correct by virtue of correctly predicting the reaction rates of these experiments. These results should primarily be viewed as an evaluation of the model’s ability to predict (the lack of) early reaction halting and the general trend of the effect of sequence length on transcription rates.

These predictions may seem trivial, in the sense that they predict that sequence length has roughly no effect on the mass-based transcription rate of the IVT system. However, most previous models of IVT rely on the assumption of initiation-limitation and would predict that the initial rates of the three curves in Figure 4 should be identical (Akama et al., 2012; van de Berg et al., 2021; Young et al., 1997). The only past work on developing expressions for the transcription rate that attempts to model the effect of sequence length was carried out using a relative paucity of data and concluded that the IVT reaction is primarily initiation-limited (Arnold et al., 2001). As such, the formulation presented in this work and demonstrated in Figure 4 is a break with past modeling conventions and a framework for future development of IVT models.

Author	T7 RNAP (nM)	DNA (nM)	NTP (mM)	Mg (mM)	PPiase (U/ $\mu$ L)	Sequence Length (nt)	RNA output ( $\mu$ M)	RNA output (g/L)
Akama et al., 2012	50–100	7.4	3.2–12.8	8–20	0	868	2–3.5	0.65–1.3
Rosa et al., 2022	124	90	31	50	1	846–4950	24	8–14

Table 2: Ranges of inputs and outputs in data explored in this work. The Akama dataset was used for model fitting (Figure 2), with exception of PPiase addition data which was used for model validation (Figure 3). The Rosa data (Figure 4) are used for model validation.

### 5.3 Limitations and Directions for Model Improvement

The mechanism in our model that is most poorly understood in the literature is the physics by which  $\text{Mg}_2\text{PPi}$  solid formation inhibits forward progress of the reaction. The hypothesis presented in this work – that  $\text{Mg}_2\text{PPi}$  nuclei agglomerate with DNA and decrease solution DNA concentration – is the best available explanatory mechanism based on findings of past qualitative work (Kim et al., 2019) and the consistency of quantitative predictions with data. However, the mechanistic understanding does not currently exist to rule out interactions of  $\text{Mg}_2\text{PPi}$  nuclei and other biomolecules such as T7 RNA polymerase as an alternative cause. In addition, a more developed understanding of how solution conditions affect the inhibition process is needed to accurately extend these results to different regimes. In the Akama dataset, increasing magnesium concentrations at already high Mg concentrations (from 8 mM to 20 mM) has a non-negligible effect on final yields without affecting the initial rate (Figure 2C). The model does not have the ability to describe this phenomenon, which is possibly due to Mg modulating the rate of sequestration. While this behavior is a limited component of the calibration data used in this work, capturing that effect would be needed for describing regimes of high Mg concentration that experience  $\text{Mg}_2\text{PPi}$  solid formation.



## 6 Conclusion

Modern design, scale-up, and optimization of the IVT reaction continues to rely on the limited capabilities of heuristics-based design-of-experiments and data-driven modelling methods. Mechanistic models for IVT stand to provide rational and interpretable predictions of RNA yields outside of previously tested design spaces. In this work, we synthesized the first mechanistic model to feature an interpretable description of IVT alongside magnesium pyrophosphate crystallization, DNA-Mg<sub>2</sub>PPi agglomeration, and pyrophosphatase enzyme activity. This model successfully describes trends observed in IVT experimental data, many of which lead to critically low RNA yields for previously unexplained reasons. Given that the IVT reaction is a foundational component for the manufacturing of a diverse and growing set of modern therapeutics, this model has the potential to provide insights for a variety of biomanufacturing systems.

## 7 Author Contributions

**Nathan Merica Stover:** Conceptualization, methodology, Software, Validation, Verification, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization.

**Krystian Ganko:** Methodology, Writing - review & editing.

**Richard D. Braatz:** Conceptualization, Methodology, Resources, Writing - review & editing, Supervision, Project administration, Funding acquisition.

## 8 Acknowledgements

This research was supported by the U.S. Food and Drug Administration under the FDA BAA-22-00123 program, Award Number 75F40122C00200. K.G was supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Department of Energy Computational Science Graduate Fellowship under Award # DE-SC0022158. The authors would like to thank Drs. Vivekananda Bal, Francesco Destro, and Christopher Canova for advice on methodology.

## References

- Akama, S., Yamamura, M., & Kigawa, T. (2012). A multiphysics model of in vitro transcription coupling enzymatic reaction and precipitation formation. *Biophysical Journal*, 102(2), 221–230. <https://doi.org/10.1016/j.bpj.2011.12.014>
- Arnold, S., Siemann, M., Scharnweber, K., Werner, M., Baumann, S., & Reuss, M. (2001). Kinetic modeling and simulation of in vitro transcription by phage t7 RNA polymerase. *Biotechnology and Bioengineering*, 72(5).
- Baker, Y. R., Chen, J., Brown, J., El-Sagheer, A. H., Wiseman, P., Johnson, E., Goddard, P., & Brown, T. (2018). Preparation and characterization of manganese, cobalt and zinc DNA nanoflowers with tuneable morphology, DNA content and size. *Nucleic Acids Research*, 46(15), 7495–7505. <https://doi.org/10.1093/nar/gky630>
- Barbier, A. J., Jiang, A. Y., Zhang, P., Wooster, R., & Anderson, D. G. (2022). The clinical progress of mRNA vaccines and immunotherapies. *Nature Biotechnology*, 40(6), 840–854. <https://doi.org/10.1038/s41587-022-01294-2>
- Beck, J. V., & Arnold, K. J. (1977). *Parameter Estimation in Engineering and Science*, [page 378]. John Wiley & Sons, New York.

- Beckert, B., & Masquida, B. (2011). Synthesis of RNA by in vitro transcription. In H. Nielsen (Ed.), *RNA: Methods and protocols* (p. 378). [https://doi.org/10.1007/978-1-59745-248-9\\_3](https://doi.org/10.1007/978-1-59745-248-9_3)
- Bergwerff, L., & van Paassen, L. A. (2021). Review and recalculation of growth and nucleation kinetics for calcite, vaterite and amorphous calcium carbonate. *Crystals*, 11(11), 1318. <https://doi.org/10.3390/cryst11111318>
- Chao, T.-C., Huang, H., Tsai, J.-Y., Huang, C.-Y., & Sun, Y.-J. (2006). Kinetic and structural properties of inorganic pyrophosphatase from the pathogenic bacterium helicobacter pylori. *Proteins: Structure, Function, and Bioinformatics*, 65(3), 670–680. <https://doi.org/10.1002/prot.21093>
- Destro, F., & Barolo, M. (2022). A review on the modernization of pharmaceutical development and manufacturing – trends, perspectives, and the role of mathematical modeling. *International Journal of Pharmaceutics*, 620, 121715. <https://doi.org/10.1016/j.ijpharm.2022.121715>
- Halonen, P., Baykov, A. A., Goldman, A., Lahti, R., & Cooperman, B. S. (2002). Single-turnover kinetics of saccharomyces cerevisiae inorganic pyrophosphatase. *Biochemistry*, 41(40), 12025–12031. <https://doi.org/10.1021/bi026018z>
- Hong, M. S., Severson, K. A., Jiang, M., Lu, A. E., Love, J. C., & Braatz, R. D. (2018). Challenges and opportunities in biopharmaceutical manufacturing control. *Computers & Chemical Engineering*, 110, 106–114. <https://doi.org/10.1016/j.compchemeng.2017.12.007>
- Kanwal, F., Chen, T., Zhang, Y., Simair, A., Rujie, C., Sadaf Zaidi, N. u. S., Guo, X., Wei, X., Siegel, G., & Lu, C. (2018). Large-scale in vitro transcription, RNA purification and chemical probing analysis. *Cellular Physiology and Biochemistry*, 48(5), 1915–1927. <https://doi.org/10.1159/000492512>
- Kern, J. A., & Davis, R. H. (1997). Application of solution equilibrium analysis to in vitro RNA transcription. *Biotechnology Progress*, 13(6), 747–756. <https://doi.org/10.1021/bp970094p>
- Kim, E., Agarwal, S., Kim, N., Hage, F. S., Leonardo, V., Gelmi, A., & Stevens, M. M. (2019). Bioinspired fabrication of DNA–inorganic hybrid composites using synthetic DNA. *ACS Nano*, 13(3), 2888–2900. <https://doi.org/10.1021/acsnano.8b06492>
- Kim, E., Zwi-Dantsis, L., Reznikov, N., Hansel, C. S., Agarwal, S., & Stevens, M. M. (2017). One-pot synthesis of multiple protein-encapsulated DNA flowers and their application in intracellular protein delivery. *Advanced Materials*, 29(26), 1701086. <https://doi.org/10.1002/adma.201701086>
- Kis, Z., Kontoravdi, C., Shattock, R., & Shah, N. (2021). Resources, production scales and time required for producing RNA vaccines for the global pandemic demand. *Vaccines*, 9(1), 3. <https://doi.org/10.3390/vaccines9010003>
- Koh, H. R., Roy, R., Sorokina, M., Tang, G.-Q., Nandakumar, D., Patel, S. S., & Ha, T. (2018). Correlating transcription initiation and conformational changes by a single-subunit RNA polymerase with near base-pair resolution. *Molecular Cell*, 70(4), 695–706.e5. <https://doi.org/10.1016/j.molcel.2018.04.018>
- Liu, D. C., & Nocedal, J. (1989). On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, 45(1), 503–528. <https://doi.org/10.1007/BF01589116>
- Martí, R. (2003). Multi-start methods. In F. Glover & G. A. Kochenberger (Eds.), *Handbook of Metaheuristics* (pp. 355–368). Springer US. [https://doi.org/10.1007/0-306-48056-5\\_12](https://doi.org/10.1007/0-306-48056-5_12)
- Pregeljc, D., Skok, J., Vodopivec, T., Mencin, N., Krušič, A., Ličen, J., Nemec, K. Š., Štrancar, A., & Sekirnik, R. (2023). Increasing yield of in vitro transcription reaction with at-line



- high pressure liquid chromatography monitoring. *Biotechnology and Bioengineering*, 120(3), 737–747. <https://doi.org/10.1002/bit.28299>
- Rohner, E., Yang, R., Foo, K. S., Goedel, A., & Chien, K. R. (2022). Unlocking the promise of mRNA therapeutics. *Nature Biotechnology*, 40(11), 1586–1600. <https://doi.org/10.1038/s41587-022-01491-z>
- Rosa, S. S., Nunes, D., Antunes, L., Prazeres, D. M. F., Marques, M. P. C., & Azevedo, A. M. (2022). Maximizing mRNA vaccine production with bayesian optimization. *Biotechnology and Bioengineering*, 119(11), 3127–3139. <https://doi.org/10.1002/bit.28216>
- Rosa, S. S., Prazeres, D. M. F., Azevedo, A. M., & Marques, M. P. C. (2021). mRNA vaccines manufacturing: Challenges and bottlenecks. *Vaccine*, 39(16), 2190–2200. <https://doi.org/10.1016/j.vaccine.2021.03.038>
- Samnuan, K., Blakney, A. K., McKay, P. F., & Shattock, R. J. (2021). Design-of-experiments in vitro transcription yield optimization of self-amplifying RNA [was published in F1000 Research]. *BioRxiv*. <https://doi.org/10.1101/2021.01.08.425833>
- Shopsowitz, K. E., Roh, Y. H., Deng, Z. J., Morton, S. W., & Hammond, P. T. (2014). Composite RNAi-microsponges form through self-assembly of the organic and inorganic products of transcription. *Small*, 10(8), 1623–1633. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4031615/>
- Tammenkoski, M., Moiseev, V., Lahti, M., Ugochukwu, E., Brondijk, T. H. C., White, S. A., Lahti, R., & Baykov, A. A. (2007). Kinetic and mutational analyses of the major cytosolic exopolyphosphatase from *saccharomyces cerevisiae*. *Journal of Biological Chemistry*, 282(13), 9302–9311. <https://doi.org/10.1074/jbc.M609423200>
- Tang, G.-Q., Anand, V. S., & Patel, S. S. (2011). Fluorescence-based assay to measure the real-time kinetics of nucleotide incorporation during transcription elongation. *Journal of Molecular Biology*, 405(3), 666–678. <https://doi.org/10.1016/j.jmb.2010.10.020>
- van de Berg, D., Kis, Z., Behmer, C. F., Samnuan, K., Blakney, A. K., Kontoravdi, C., Shattock, R., & Shah, N. (2021). Quality by design modelling to support rapid RNA vaccine production against emerging infectious diseases. *npj Vaccines*, 6(1), 1–10. <https://doi.org/10.1038/s41541-021-00322-7>
- Wang, Y., Kim, E., Lin, Y., Kim, N., Kit-Anan, W., Gopal, S., Agarwal, S., Howes, P. D., & Stevens, M. M. (2019). Rolling circle transcription-amplified hierarchically structured organic–inorganic hybrid RNA flowers for enzyme immobilization. *ACS Applied Materials & Interfaces*, 11(26), 22932–22940. <https://doi.org/10.1021/acsami.9b04663>
- Young, J. S., Ramirez, W. F., & Davis, R. H. (1997). Modeling and optimization of a batch process for in vitro RNA production. *Biotechnology and Bioengineering*, 56(2), 210–220. [https://doi.org/10.1002/\(SICI\)1097-0290\(19971020\)56:2<210::AID-BIT10>3.0.CO;2-K](https://doi.org/10.1002/(SICI)1097-0290(19971020)56:2<210::AID-BIT10>3.0.CO;2-K)