Received: 26 February 2024 | Revised: 10 May 2024 | Accepted: 3 July 2024

DOI: 10.1002/bit.28806

ARTICLE

BIOTECHNOLOGY AND BIOENGINEERING WILEY

Quality by design approach to improve quality and decrease cost of in vitro transcription of mRNA using design of experiments

Jimmy Boman[1] | Tjaša Marušič[2,3] | Tina Vodopivec Seravalli[2] | Janja Skok[2] | Fredrik Pettersson[1] | Kristina Šprinzar Nemec[2] | Henrik Widmark[1] | Rok Sekirnik[2]

[1]Sartorius Stedim Data Analytics AB, Umeå, Sweden

[2]Sartorius BIA Separations d.o.o., Sartorius Company, Ajdovščina, Slovenia

[3]Biotechnical Faculty, University of Ljubljana, Ljubljana, Slovenia

Correspondence

Rok Sekirnik, Sartorius BIA Separations d.o.o., Sartorius Company, 5270 Ajdovščina, Slovenia.
Email: rok.sekirnik@biaseparations.com

Henrik Widmark, Sartorius Stedim Data Analytics AB, Umeå, 903 33, Sweden.
Email: henrik.widmark@sartorius.com

Abstract

In vitro transcription (IVT) reaction is an RNA polymerase-catalyzed production of messenger RNA (mRNA) from DNA template, and the unit operation with highest cost of goods in the mRNA drug substance production process. To decrease the cost of mRNA production, reagents should be optimally utilized. Due to the catalytic, multicomponent nature of the IVT reaction, optimization is a multi-factorial problem, ideally suited to design-of-experiment approach for optimization and identification of design space. We derived a data-driven model of the IVT reaction and explored factors that drive process yield (in g/L), including impact of nucleoside triphosphate (NTP) concentration and Mg:NTP ratio on reaction yield and how to optimize the main cost drivers RNA polymerase and DNA template, while minimizing dsRNA formation, a critical quality attribute in mRNA products. We report a methodological approach to derive an optimum reaction design, with which cost efficiency of the reaction was improved by 44%. We demonstrate the validity of the model on mRNA construct of different lengths. Finally, we maximized the yield of the IVT reaction to 24.9 ± 1.5 g/L in batch, thus doubling the highest ever reported IVT yield.

KEYWORDS

DOE, dsRNA, HPLC, in vitro transcription, mRNA, QbD

1 | INTRODUCTION

The SARS-Cov-2 pandemic has contributed to accelerated research of messenger RNA (mRNA) based vaccines. In the battle against SARS-Cov-2, mRNA vaccines were the first to be authorized in both Europe and the United States (Barbier et al., 2022). Increased interest in mRNA was further driven by promising therapeutic applications such as cancer immunotherapies, protein replacement therapies, regenerative medicine and cellular reprogramming (Sahin et al., 2014). The process of mRNA production by in vitro transcription (IVT) is relatively simple and allows to produce large amounts of mRNA in a short time (3–5 days) compared to weeks required to produce traditional vaccines (Rosa et al., 2021). Reaction yields of 2–5 g/L can be achieved within a few hours and up to 12 g/L has been reported (Hengelbrock et al., 2023; Rosa et al., 2022). Once a viable operating range of IVT reaction parameters, also called the

Jimmy Boman and Tjaša Marušič contributed equally to this study.

knowledge space, is understood, different constructs of mRNA (and other RNA) therapeutics could be produced using the same reaction components, equipment, analytical methods and even same IVT reaction model (Daniel et al., 2022). Coupling effective process analytical technology (PAT) and statistical methods to establish a uniform platform with a high predictive power (e.g., digital twin) and a robust design space (DSp) where optimal outcomes are guaranteed not only in terms of cost and efficiency but also product safety, will therefore be of significant importance for this new class of medicines (Kis et al., 2022).

The basis of IVT reaction is a DNA template (typically linearized plasmid) with a promoter region, selectively recognized by bacteriophage RNA polymerase (RNAP) which synthesizes mRNA complementary to DNA template sequence. The most commonly used RNA polymerase is T7, a magnesium-dependent enzyme, consisting of a single, approximately 100 kDa subunit highly specific for a 23 base pair (bp) promoter sequence (Martin & Coleman, 1987). RNAP is sensitive to oxidation and to stabilize the enzyme and maintain its activity, dithiothreitol (DTT) is commonly used as an antioxidant at concentrations of 1–10 mM (Chamberlin & Ring, 1973; Rosa et al., 2022). Mg^{2+} ions are critical for binding of RNAP to DNA template. Moreover, each nucleoside triphosphate (NTP) forms a complex with one Mg^{2+} ion before entering the active site of the enzyme to be incorporated into the growing mRNA chain. It has been reported that due to the low dissociation constant, Mg^{2+} ions exist in solution primarily as $MgNTP^{2-}$ and that the elongation phase of T7 RNAP-based transcription requires both $MgNTP^{2-}$ and free Mg^{2+} ions (Akama et al., 2012; Thomen et al., 2008). Minimally 5 mM and optimally 10–20 mM of free Mg^{2+} was reported to be needed for effective transcription (Kern & Davis, 1997). Excessive concentrations of Mg^{2+} are unfavorable, due to increased ionic power and inhibitory effects of counterions on RNAP, which reduce transcription rate and efficiency of NTP incorporation into mRNA chain (Kern & Davis, 1997). Mg^{2+} salts in the form of chloride (Cl^-) or acetate (OAc^-) are usually used in IVT reaction; Cl^- was shown to be a stronger inhibitor of RNAP than OAc^-; inhibitory effect of Cl^- ions is most notable above 100 mM and OAc^- above 200 mM (Kern & Davis, 1997; Maslak & Martin, 1994; Young et al., 1997). In addition, high concentrations of free Mg^{2+} contribute to mRNA hydrolysis (Guth-Metzler et al., 2023), which is exacerbated by prolonged reaction duration (Gößringer et al., 2014). The concentration of Mg^{2+} must therefore be adjusted to initial NTP concentration (Young et al., 1997). A wide range of optimal concentrations of Mg^{2+} and NTP have been reported (20–75 mM for Mg^{2+}, 4–12 mM for NTP, see Table 1), suggesting that understanding of factors governing IVT reaction has thus far been incomplete. We postulate that Mg:NTP ratio should be studied rather than their concentrations separately.

The main cost-driver of the reaction (assuming post-transcriptional capping and wild-type NTPs) is RNAP, followed by plasmid DNA (Kis et al., 2020). These two factors were demonstrated to greatly impact the yield of the IVT reaction, alongside NTP and Mg^{2+} (Yin & Carter, 1996). There have been multiple attempts to optimize the IVT reaction with one-factor-at-a-time (OFAT), design of experiments (DOE) or machine learning approaches from either production or cost perspective, but none of them studied concurrent effects on product quality. A major undesired by-product of the IVT reaction is double-stranded RNA (dsRNA) which is a potent stimulant of the innate immune system that can elicit undesired and

TABLE 1 Previously published concentrations of Mg^{2+}, NTP, Mg^{2+} counterion and Mg:NTP ratio reported as optimal for IVT reaction.

Source	Approach	RNA type	Mg^{2+} (mM)	NTP (mM)	Mg:NTP	Time (h)	Quantification	Yield (g/L)	Relative yield (%)
Yin and Carter (1996)	DOE	tRNA	56.75 (Cl^-)	11.6	1.22[a]	Over-night	PAGE	2.45	17[b]
Young et al. (1997)	Model	dodecamer	54 (OAc^-)	12	1.5[a]	4	D-PAGE	3.2[c]	33
Kanwal et al. (2018)	OFAT	mRNA	45 (n.a.)	8	1.41[a]	9	D-PAGE	n.a.	n.a.
Kartje et al. (2021)	OFAT	aptamer	30 (Cl^-)	5	1.5[a]	0.5	qPCR, D-PAGE	0.13[c]	3
Samnuan et al. (2021)	DOE	saRNA	75 (OAc^-)	10	1.875	6	Qubit (FLD)	2.8	22[b]
Rosa et al. (2022)	Bayesian optimization	mRNA	49.3 (OAc^-)	7.75	1.59[a]	2	RP-HPLC	12.2	123[b]
Hengelbrock et al. (2023)	DOE	mRNA	50 (OAc^-)	10	1.25[a]	n.a.	HPLC	12	94[b]
Pregeljc et al. (2023)	OFAT	mRNA	20 (Cl^-)	4	1.25[a]	2	HPLC	4.2	94[b]

Note: Listed NTP concentrations refer to the concentrations of individual NTPs.

Abbreviations: Cl^-, chloride counterion; DOE, design of experiments; D-PAGE, denaturing PAGE; HPLC, high-pressure liquid chromatography; OFAT, one-factor-at-a-time; OAc^-, acetate counterion; PAGE, polyacrylamide gel electrophoresis; qPCR, quantitative PCR; RP-HPLC, reverse phase HPLC; saRNA, self-amplifying RNA; tRNA, transfer RNA.

[a]Mg:NTP ratio was calculated by the authors and was not specified in the original source.

[b]Relative yields were calculated with the assumption that same amount of each nucleoside is present in mRNA sequence.

[c]Yields were reported in molar concentration and converted to g/L by the authors.

uncontrolled immune responses in cells (Mu & Hur, 2021). Therefore, minimizing dsRNA formation during the IVT reaction and/or its removal after the IVT process is one of the key goals during mRNA upstream and downstream processing (Cho et al., 2023; Nagaraj et al., 2022). Many strategies for decreasing dsRNA formation in the IVT reaction have been reported, including optimization of mRNA sequence (e.g. modification of the 3' end of the template (Mu et al., 2018; Nacheva & Berzal-Herranz, 2003)), use of modified nucleotides (Karikó et al., 2008; Mu et al., 2018), lower concentration of UTP (Nelson et al., 2020; Ziegenhals et al., 2023), lower levels of magnesium ions (Mu et al., 2018), the addition of competing oligonucleotides (Gholamalipour et al., 2019), performing IVT reaction at higher temperature with use of thermostable T7 RNAP (Wu et al., 2020) or with the addition of chaotropic agents such as urea (Piao et al., 2022). Point mutations in RNAP sequence were also shown to decrease dsRNA production in IVT (Dousis et al., 2023). However, many of these techniques can pose economic, technical, or regulatory challenges during the scale-up of the mRNA production process (Cho et al., 2023). Lowering the concentration of UTP or Mg^{2+} appears to be a cost-effective and technically feasible approach to reduce dsRNA formation but could lead to decreased mRNA yield and should be carefully optimized. To our knowledge, reports that describe factors influencing dsRNA levels did not also quantify mRNA yield.

mRNA synthesis is thus a multi-parameter process with multiple potentially interacting factors. By defining a design space for the mRNA manufacturing process to meet the Quality Target Product Profile according to guidelines such as ICH Q8, consistent quality of commercial lots, and cost efficiency can be achieved. Furthermore, future process optimization is possible within the design space without triggering a change request to the regulatory authorities. The European Medicines Agency states that "any multivariate interactions between the DSp parameters need to be studied. In particular, when the acceptable range of one parameter in DSp is dependent on any other parameter, this should be thoroughly investigated, including consideration of scale" (EMA/604040/2016EMA/CHMP/CVMP/QWP/354895/2017,4040/2016EMA/CHMP/CVMP/QWP/354895/2017, n.d.) and traditional OFAT approaches to process optimization cannot be efficiently applied to a multi-variate problem such as IVT (Eriksson, 2008; Yin & Carter, 1996).

Our study aimed to identify a design space for a generic IVT reaction, to maximize cost efficiency while satisfying a set quality attribute (dsRNA). We also strived to unify the apparently disparate prior reports of factors governing mRNA yield. Our approach was to couple at-line analysis, able to quantify NTPs and mRNA in a time-resolved manner, with DOE tools, to derive a data-driven model of the IVT reaction that could be used for prediction of mRNA production outcomes regardless of sequence and length of mRNA construct. We evaluated the effects of individual and combined parameters on mRNA production, yield, cost-effectiveness of the reaction, and dsRNA formation, thus establishing a design space where all set requirements can be met, on a model mRNA construct encoding the eGFP sequence (995 nucleotides, including a 45-nucleotide poly(A) tail). Our study included dsRNA content as the sole critical quality attribute (CQA), but the model could in principle be extended to other CQAs, such as mRNA integrity, 5′ capping efficiency and others.

2 | MATERIALS AND METHODS

2.1 | IVT reaction

RNAse inhibitor (40 U/μL), pyrophosphatase (100 U/mL), T7 RNA polymerase (50 U/μL), ATP, UTP, CTP, and GTP (100 and 200 mM stocks) were from Mebep Bioscience. $MgCl_2$ (1 M) was from Invitrogen, 10 × IVT buffer (400 mM Tris, 20 mM spermidine, 10 mM DTT, pH 7.9) was prepared in-house. All IVT reagents listed except enzymes were preheated to 37°C, mixed in a 1.5 mL plastic tube in Thermomixer™C (Eppendorf) and, after the addition of enzymes, incubated at 37°C with shaking at 300 rpm. For sampling, 4 μL aliquots were quenched with 4 μL of 100 mM EDTA pH 8.0 for HPLC analytics every 15 min for the first 60 min, thereafter every 30 min until quenching the reaction at 210 min as described below. A reference protocol was used as the basis for optimization: 30 mM $MgCl_2$, 8 mM of each NTP, 50 ng/μL DNA, 10 U/μL T7 RNAP, 1 U/μL RNase inhibitor, 1 U/mL pyrophosphatase and 1 × IVT buffer (40 mM Tris, 2 mM spermidine, 1 mM DTT, pH 7.9) (Skok et al., 2022).

2.2 | Yield calculation

For a given mRNA construct of length n, the number of each NTP in the sequence is n_A, n_C, n_G, and n_U. For any NTP mixture at least one of the NTPs (referred to as NTP_{LIM}) will limit the production of mRNA, and a maximum of $1/n_{LIM}$ M of mRNA can be produced from 1 M of NTP_{LIM}. The maximum theoretical yield can then be calculated as

$$Yield_{MAX}\left(\frac{g}{L}\right) = M_{mRNA}\left(\frac{g}{mol}\right)\frac{c_{NTP_{LIM}}(M)}{n_{LIM}}, \quad (1)$$

where M_{mRNA} is the molecular weight of mRNA:

$$M_{mRNA}\left(\frac{g}{mol}\right) = 329.2\,\frac{g}{mol}n_A + 306.2\,\frac{g}{mol}n_U + 305.2\,\frac{g}{mol}n_C + 345.2\,\frac{g}{mol}n_G. \quad (2)$$

Relative process yields were calculated as

$$\text{Relative yield (\%)} = 100 \times \frac{\text{Yield (g/L)}}{\text{Yield}_{MAX}\,(g/L)}. \quad (3)$$

2.3 | Design of experiments

All DOEs were created and analyzed in MODDE® 13 (Sartorius Stedim Data Analytics AB). Responses were mRNA yield (g/L), dsRNA content, relative yield as a percentage of the maximum possible

TABLE 2 Factors spanning prior knowledge space, with their high and low settings for the DOE.

Factor	Unit	Low setting	High setting	Yield (g/L)	dsRNA	Relative yield (%)	Cost (€/mg)
Mg:NTP (total)	Ratio	0.8	1.6	∩	++	∩	0
NTP (individual)	mM	6	12	+	0	-	+
DNA template	ng/µL	10	100	+/0	0	+	++
T7 RNAP	U/µL	2	15	+	0	+	+++

Note: The four rightmost columns represent the expected effect of increasing each factor on the different responses. "+" indicates that a response is expected to increase with an increase of the factor, and the opposite is true for effects marked with a "-", "0" indicates that a response is not expected to change with an increase of the factor. For effects marked with "∩", an optimum is expected within the knowledge space.

TABLE 3 List of full quadratic model terms.

Main effects	Mg:NTP, NTP, DNA, T7
Interaction terms	**Mg:NTP*NTP**, Mg:NTP*DNA, Mg:NTP*T7, NTP*DNA, NTP*T7, **DNA*T7**
Square terms	**Mg:NTP*Mg:NTP, NTP*NTP,** DNA*DNA, T7*T7

Note: Bold indicates terms/effects considered as most impactful in setting up DOE.

theoretical yield, and cost efficiency (µg/€) calculated as the mass of mRNA produced divided by the cost of the reagents (Supporting Information S1: Table S1). Relative yield response was not directly linked to an objective but was instead used for evaluating model validity. We have arbitrarily set a required limit for dsRNA response below 3 as determined by J2 dot blot (scale 0–10 defined below). Selected DOE factors are listed in Table 2, along with their low and high settings and estimated effects. Ranges were identified based on a literature review (Supporting Information S1: Figure S1).

Based on the selected list of factors, the 6 interaction terms, four square terms and four main effects were ranked on their estimated relative impact on the process output to select the eight most impactful terms to include in the model (bolded terms, Table 3).

The four main effects, interaction terms between Mg:NTP ratio and NTP, and between T7 RNAP and DNA, as well as square terms for NTP and Mg:NTP ratio were ranked as most impactful (see Discussion for selection criteria) and selected for the first iteration of the model. An initial iteration based on a D-optimal design of 12 experiments (Supporting Information S1: Table S2) was generated to verify factor ranges and identify potential optima or plateaus within the knowledge space. The DOE was complemented with two additional iterations (Supporting Information S1: Tables S3 and S4) to support the full model of all 14 model terms.

A multiple linear regression (MLR) was fitted and evaluated based on heredity, parsimony, lack of fit and model validity using MODDE's Analysis Wizard. Model terms were selected to give the highest fit to data (R^2) and prediction ability (Q^2). The final model was used to identify a design space and the most cost-efficient settings, which was verified in a final set of experiments (Supporting Information S1: Table S5).

2.4 | CIMac PrimaS analysis for determination of mRNA concentration and NTP consumption

HPLC analysis with CIMac PrimaS™ (Sartorius BIA Separations) for mRNA quantification and determination of NTP consumption was performed as previously reported (Skok et al., 2022). PATfix 2.0 software (Sartorius BIA Separations) was used for instrument control, data acquisition and data analysis. Linearity of signal responses (mRNA, UTP/CTP, GTP, ATP) is shown in Supporting Information S1: Figure S2, samples for HPLC analysis were diluted accordingly.

2.5 | Agarose gel electrophoresis

Agarose gel electrophoresis was used for visualization of mRNA and confirming mRNA integrity. Final IVT samples were diluted to 4 ng/µL in ddH_2O, 18 µL of samples were then mixed with 2 µL of TriTrack loading dye (Thermo Fischer Scientific) and loaded on 1% agarose gel. Electrophoresis was performed as previously described (Pregeljc et al., 2023). Visualization was performed with iBright (Thermo Fischer Scientific).

2.6 | Dot blot for detection of dsRNA

J2 dot blot was used for detection of double-stranded RNA in 1 µg of final RNA samples and was performed as previously described (Skok et al., 2022). After final wash step, membrane was incubated for 2 min in 5 mL Westernbright ECL components mixed in 1:1 ratio (Advansta), protected from light. Excess solution was drained off and signals were visualized with chemiluminescence with iBright (Thermo Fisher Scientific). dsRNA content was estimated on a scale of 0–10, 0 corresponding to intensity of negative control (ddH_2O) and 10 to intensity of Magi dsRNA standard (10 ng). Linearity of the

chemiluminescence signal is shown in Supporting Information S1: Figure S3.

2.7 | CIMmultus Oligo dT affinity chromatography

Chromatographic purification of mRNA was performed on ÄKTA Pure 150 (Cytiva) FPLC system composed of two pumps and a multi-wavelength UV-Vis detector (2 mm flow cell path length). Unicorn software (Cytiva) was used for instrument control and data acquisition. The IVT mixture was diluted 10-fold in loading buffer (50 mM sodium phosphate, 0.5 M NaCl, pH 7.4) and loaded onto 1 mL CIMmultus Oligo dT18 (2 mm channel size) column (Sartorius BIA Separations). After unbound IVT components eluted in flow-through and the UV 260 nm signal was stabilized, a wash step was performed with 50 mM sodium phosphate, pH 7.4, followed by step elution of polyadenylated mRNA with ddH_2O.

3 | RESULTS

3.1 | Iteration 1

The objective of the first modeling iteration was to evaluate the high and low settings of each selected factor (Table 2), to verify that the knowledge space contains an optimum for NTP concentration and Mg:NTP ratio, as well as to investigate how lowering DNA template and T7 RNAP concentrations impacts the reaction yield.

The response surface (Figure 1a) indicated that the highest process yield (12 g/L) was achieved with DNA template and T7 RNAP at their highest concentrations (100 and 15 U/μL, respectively), Mg:NTP ratio at its lowest setting (0.8) and NTP concentration between 8 and 9 mM. Fit to data was high ($R^2 = 0.90$), but predictive capability was lower ($Q^2 = 0.64$), indicating that model was not stable and required more model terms. Pure error between two replicate experiments was ±0.7 g/L (Supporting Information S1: Figure S4).

The response surface for the relative process yield (Figure 1b) shows that for a significant portion of knowledge space, the model predicted negative relative yields, as well as relative yields >160%, further suggesting that more model terms were required to describe IVT reaction accurately.

The model predicted further improvements to process yield by increasing T7 RNAP and DNA above the 'high' setting of Iteration 1; however, this was considered unlikely to be cost-efficient, given that T7 RNAP and DNA are the main cost drivers and increasing either factor from 'low' to 'high'—a 7.5-fold increase in T7 RNAP concentration (from 2 to 15 U/μL)—resulted in only a 50% increase in process output (from 8 to 12 g/L; Figure 1a). Therefore, the 'high' and 'low' settings for these factors were maintained in subsequent iterations. The model also pointed toward further improvements to process output with lower Mg:NTP. However, this was deemed unlikely based on a previous report that Mg:NTP < 0.8 resulted in very low mRNA yields (Kartje et al., 2021), and instead taken as an indication that the model should be expanded with more terms.

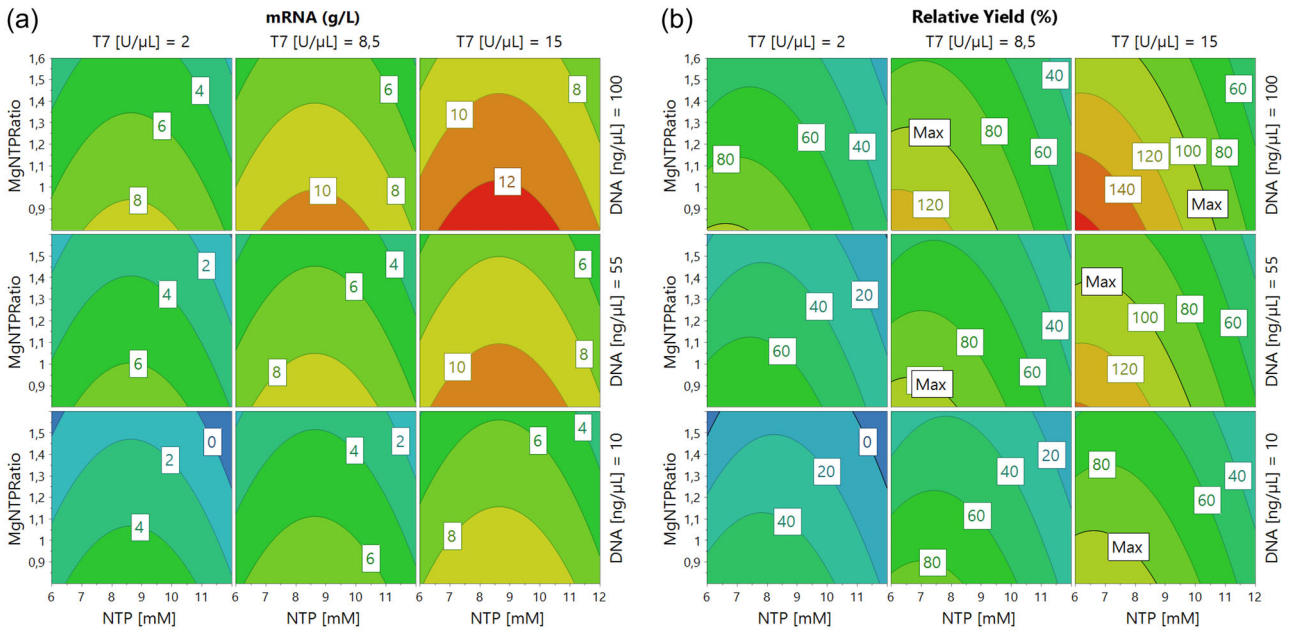

FIGURE 1 Response surfaces of DOE Iteration 1 model. (a) mRNA yield response. The model predicted that optimal process yield is achieved with NTP concentration of 8.7 mM, low Mg:NTP, high T7 RNAP and high DNA. (b) Relative yield response. Relative yield in area of the knowledge space with highest mRNA yield exceeded 100%, suggesting that the model had low validity and that more model terms were required.

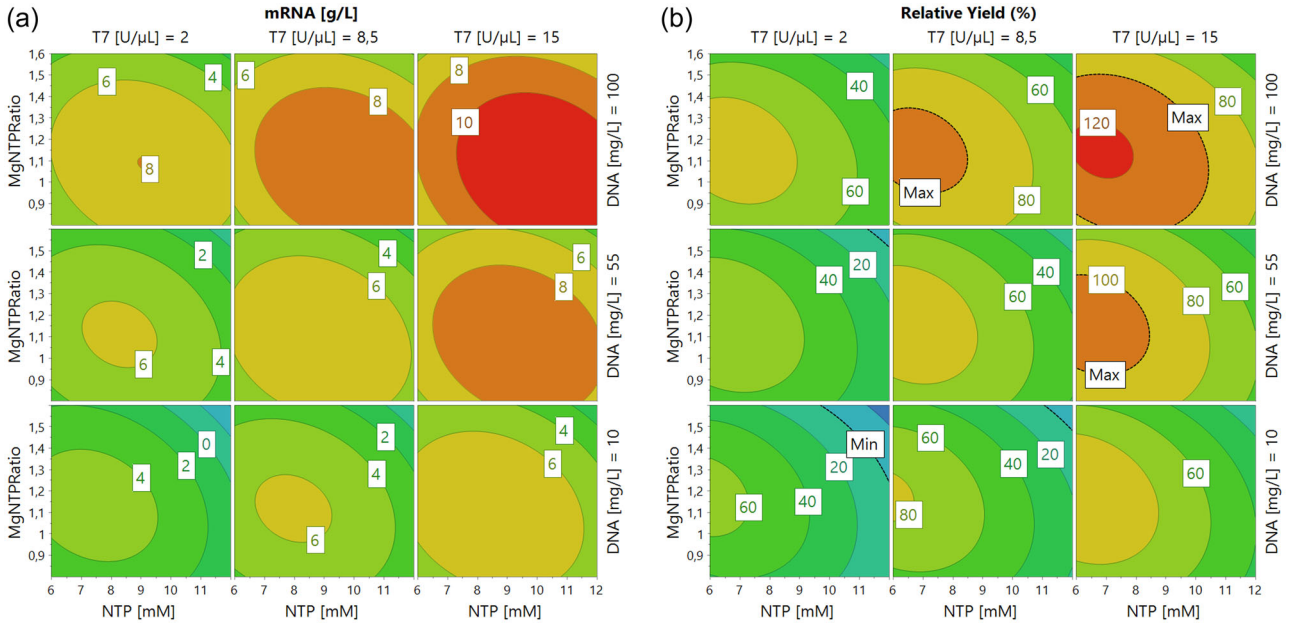

FIGURE 2 Modeling results after Iteration 2. (a) mRNA yield. (b) Relative mRNA yield. There was still an area of the knowledge space where the model predicted unrealistic yields, but the overestimation was lower compared to Iteration 1.

3.2 | Iteration 2

New factor combinations, not previously explored in Iteration 1, were selected by D-optimal function to study the interaction term between NTP and T7 RNAP (IVT098–099; Supporting Information S1: Table S3). The predictive power of the model was investigated in an area of the knowledge space where the predicted yield was realistic, that is, below 100%; selected Iteration 1 experiments were repeated as controls.

Model fit improved, now consisting of nine significant model terms compared to five in the Iteration 1 (Supporting Information S1: Figure S5). Notably, the square term for Mg:NTP ratio was statistically significant, with an optimum at 1.1 (Figure 2a); fit to data and prediction capability values improved considerably ($R^2 = 0.97$, $Q^2 = 0.85$). Maximum predicted yield was 120% at high T7 RNAP and DNA (Figure 2b), and the model suggested that increasing T7 RNAP and DNA would further increase yields above the practical limit of 100%. This led to the conclusion that curvature should be introduced to the model for those factors.

3.3 | Iteration 3

The model was then further expanded to reduce the overestimation of maximum relative yield, with complementary experiments based on D-optimal criteria. The curvature effect present in one or both of T7 RNAP and DNA was deemed unlikely to be of quadratic form; in earlier studies, increasing DNA and T7 RNAP concentrations consistently showed a positive effect on reaction rate and process output (Akama et al., 2012; Pregeljc et al., 2023). Curvature in those terms was instead modeled by log-transforming factors before generating new experiments, which allowed the model to level off, rather than introducing optima within the investigated region for T7 RNAP and DNA. Because process yield started to decrease when using Mg:NTP ratios ≥1.2, the candidate set was adjusted to Mg:NTP ratios 0.8–1.2. Design of Iteration 3 (Supporting Information S1: Table S4) supported all ten interactions and square terms (Table 3).

The difference between R^2 and Q^2 decreased significantly compared to previous iterations ($R^2 = 0.98$, $Q^2 = 0.93$), indicating that predictive performance of the model improved (Supporting Information S1: Figure S6). A contour plot of the relative yield response (Figure 3b) demonstrated that the model was now reliable in >95% of the knowledge space. The highest relative yield over-estimation was reduced to 5% (compared to the 20% from the previous iteration), or slightly below 0.5 g/L, within the magnitude of pure error (±0.56 g/L). As such, this model was deemed sufficiently accurate to proceed with cost-efficiency optimization. The final model consisted of the same nine model terms as in Iteration 2: square terms for both Mg:NTP ratio and NTP, along with all three interaction terms involving NTP (Supporting Information S1: Equation S1).

3.4 | Optimization

Agarose gel electrophoresis was used to confirm that mRNA at target molecular size was produced in all 23 reactions (Supporting Information S1: Figure S7). A model was then fitted for dsRNA content using dot blot data collected during Iterations 1–3 (Supporting Information S1: Figure S8, Table S7) following the same statistical approach. Time-resolved dot blot measurements revealed that no

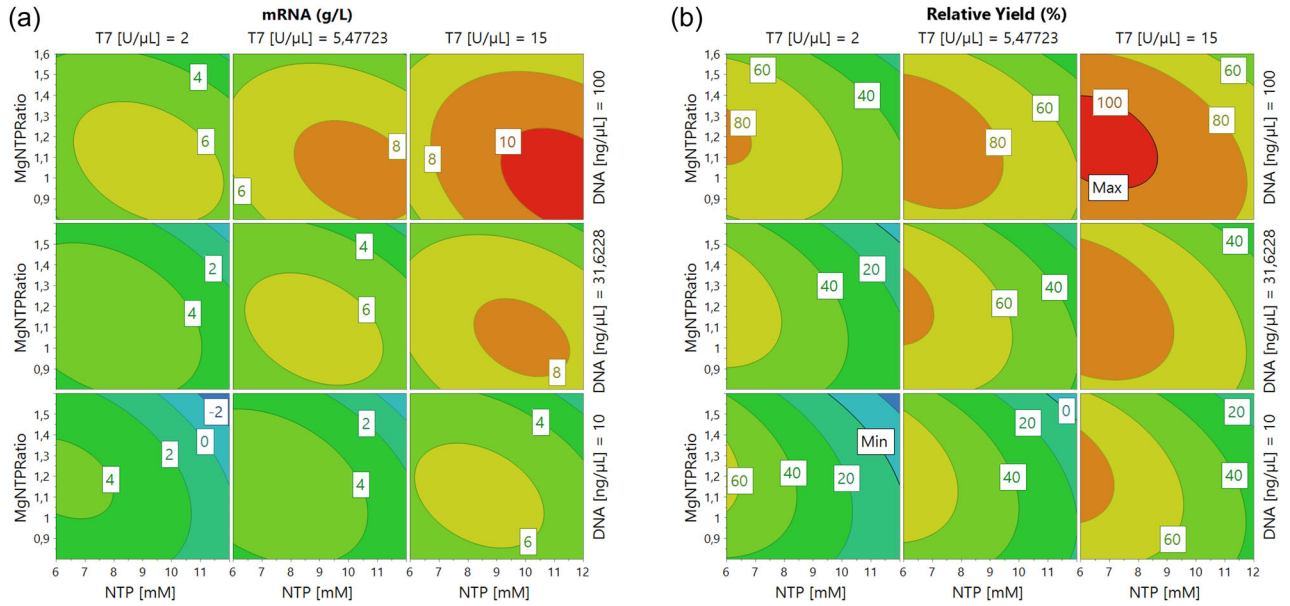

FIGURE 3 Modeling results of Iteration 3. (a) mRNA yield. (b) Relative mRNA yield. The area where the model predicted yields was smaller than after iteration 2, and the maximum predicted yield was 105%.

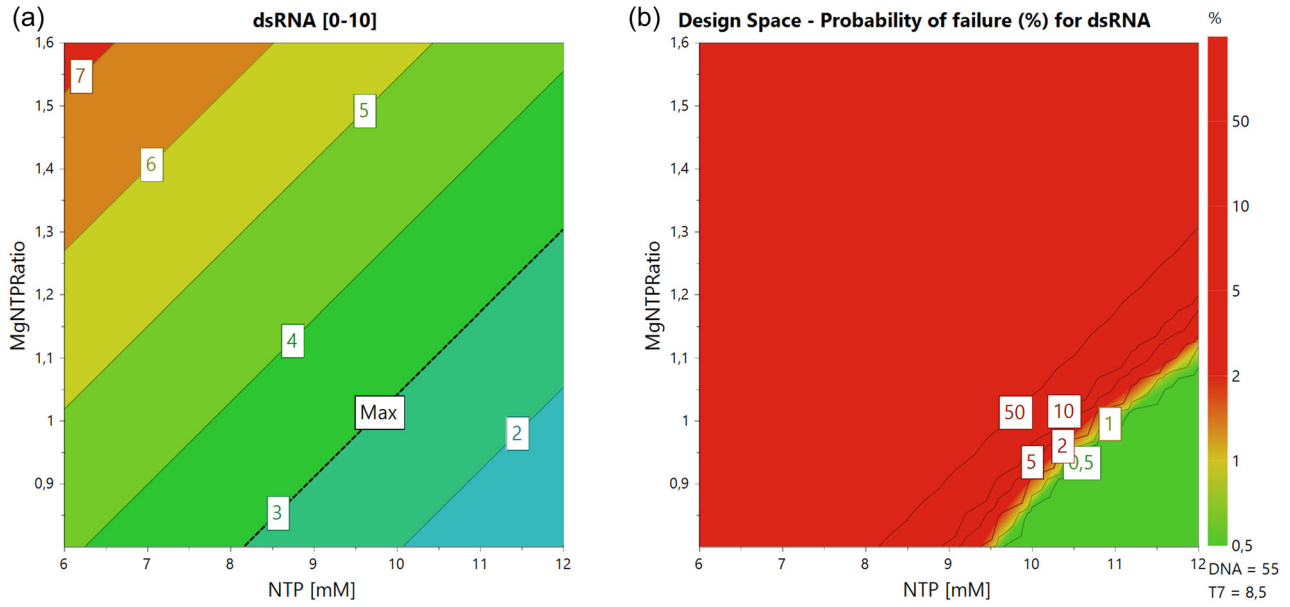

FIGURE 4 The final model for dsRNA content (a) and the design space based on the probability of exceeding the required dsRNA amount (b). Red corresponds to the high probability of failure; green represents the design space; and <1% probability of failure. The only factors that had an impact on dsRNA content were NTP concentration and Mg:NTP ratio.

dsRNA was produced after reaction reached plateau (Supporting Information S1: Figure S9). In the final dsRNA model (Figure 4a, Supporting Information S1: Equation S2) only NTP concentration and Mg:NTP ratio were statistically significant terms (Supporting Information S1: Figure S10). A design space was estimated using Monte Carlo simulations, adding 5% stochastic noise to each factor over 10,000 iterations. A probability of failure <1% was set as the limit for the design space to guarantee dsRNA score <3 (Figure 4b).

Cost efficiency response (Figure 5) was calculated using the model for mRNA yield and price ranges available at the time of writing (Supporting Information S1: Tables S1 and S7). Notably, higher cost efficiency was reached with lower concentrations of T7 RNAP (2–6 U/µL). For Mg:NTP ratio, higher cost efficiency was achieved in middle of the range (0.9–1.3), while for NTP and DNA, the most cost-efficient setting depended heavily on settings of other factors.

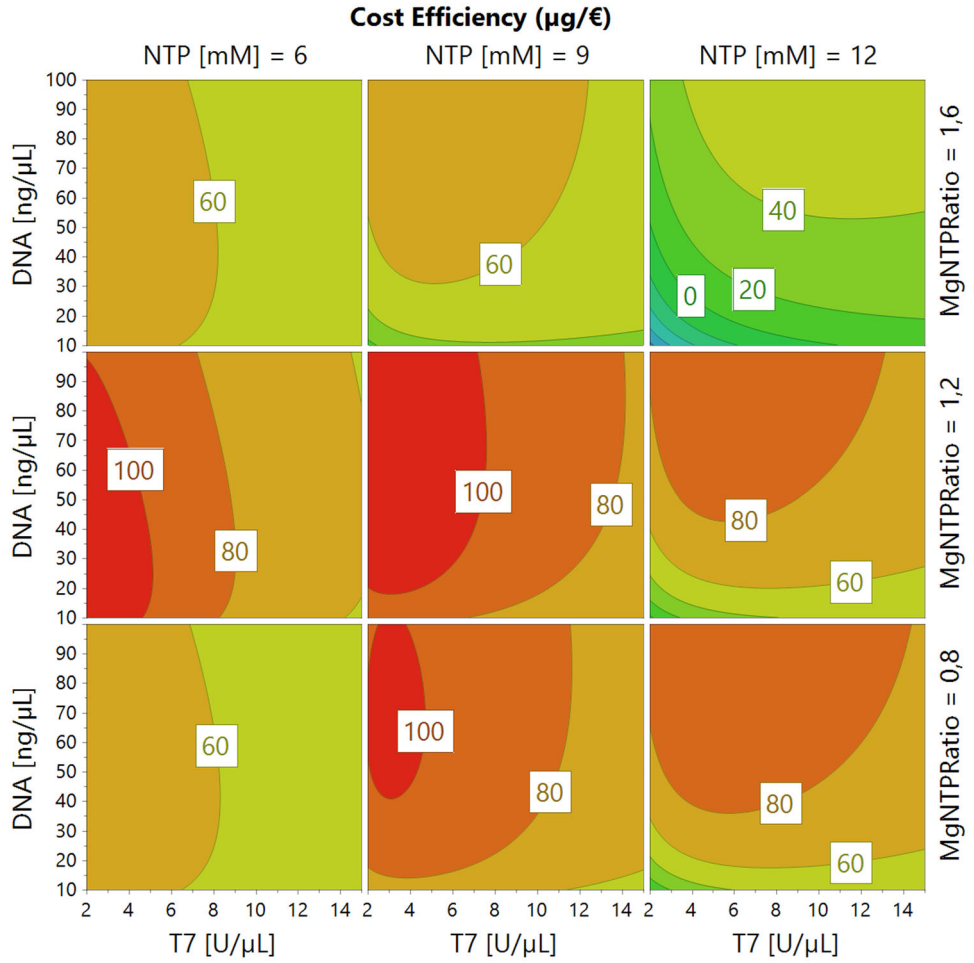

FIGURE 5 The model for cost efficiency (µg/€) indicates that using lower amounts of T7 RNAP is preferred for cost-efficient production of mRNA.

The design space for dsRNA was combined with the cost efficiency model to identify the optimal operating conditions within the design space. An optimal setpoint within the design space was identified using a desirability function applied to the model (Supporting Information S1: Table S8). Comparing responses between the reference and optimal setpoint, the new optimal setpoint had a predicted cost efficiency 14% higher than the reference, at 50% lower dsRNA.

The design space was verified with seven experiments, proximal to design space corners. Each of the four factors was varied around the optimal setpoint; the predicted optimal setpoint was also verified experimentally. Upper and lower limits for the predictions were constructed based on 95% probability of the result landing between the limits. All seven verification experiments were within specification for dsRNA content. Process yield was predicted accurately for five of the seven experiments. For the remaining two, the process output exceeded the predicted upper limit by <4%. Given the higher than predicted result for the optimal reaction conditions, cost efficiency was improved by 44% compared to the reference IVT protocol (Supporting Information S1: Table S9).

3.5 | Model validation with longer mRNA construct

To evaluate whether the model could be applied to different mRNA constructs, we performed a set of verification experiments using a construct of 3969 nt (mFIX; Supporting Information S1: Table S10). It had previously been reported that the transcription reaction is limited primarily by the initiation and termination steps, the impact of elongation rate being significant only at extreme transcript lengths (Arnold et al., 2001). The mFix construct is approximately four times longer than eGFP, therefore production kinetics (in g/L) should be four times faster under identical process settings, provided that DNA template concentration is adjusted for molarity. A time-resolved MLR model for eGFP was built by expanding the model from Iteration 3, adding time as a log-transformed factor and using mRNA concentrations from all time points (0–210 min) measured in the three iterations. Model building was carried out in a similar manner to the end-point model from the DOE iterations. The model fit the eGFP data well ($R^2 = 0.93$, $Q^2 = 0.91$). To predict concentration trajectories for the mFIX construct, the time factor was scaled by the ratio of the two construct lengths, such that the predicted concentration for

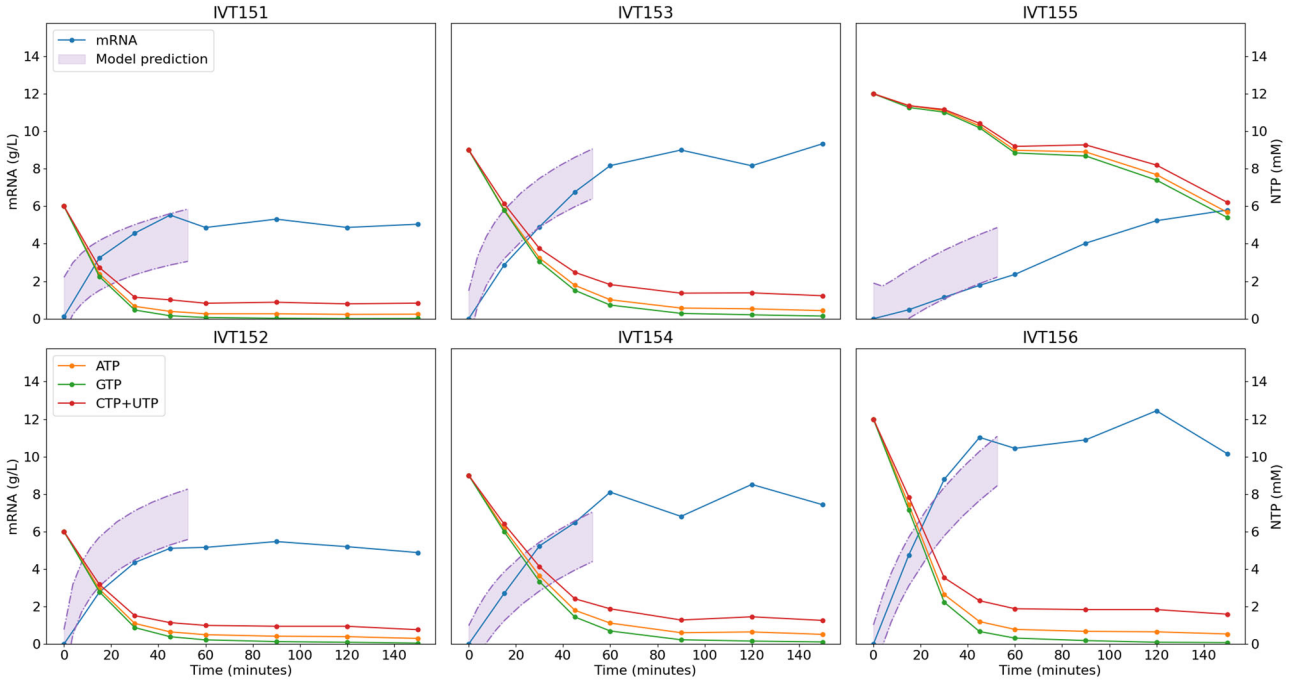

FIGURE 6 Comparison of predicted and measured mRNA concentration trajectories for mFIX mRNA (3969 nt) construct. Concentrations of mFIX mRNA construct (3969 nt) were predicted by adjusting the reaction rate of eGFP mRNA (995 nt) by a factor of 4, corresponding to the ratio between the construct sizes. Reaction trajectories were within predicted limits (95% prediction interval).

mFIX at *x* min was equal to the predicted concentration of eGFP at 4*x* min, thus compensating for the expected fourfold increase in reaction rate for the longer construct. As a consequence, concentrations of mFIX could only be predicted up to 52.5 min, as the time resolved eGFP model was only valid for reactions of up to 210 min.

Six validation experiments were selected from the corners of the knowledge space along with a center point in NTP (Supporting Information S1: Table S11). All six reaction trajectories fell within the 95% prediction intervals of the model (Figure 6). dsRNA levels for all six experiments were low, indicating that dsRNA formation is likely also construct-specific and not dependent solely on IVT conditions (Supporting Information S1: Figure S11).

3.6 | Chloride versus acetate

To verify the impact of Mg^{2+} counterion on reaction kinetics and yield, a separate DOE was performed, which varied Mg:NTP ratio and Mg^{2+} counterion over 12 total experiments (Supporting Information S1: Table S6). At Mg:NTP ratio of 1.25, counterion did not affect reaction kinetics, while at ratio 2.5, Cl^- inhibited the reaction more than OAc^-, with both yielding 4 g/L mRNA when NTP concentration was 5 mM. At ratio 2.5 and 10 mM NTP, Cl^- inhibited the reaction, while OAc^- yielded 3 g/L, albeit at a low rate (Figure 7a). Low Mg:NTP with very high Mg^{2+} and NTP levels favored OAc^-, resulting in almost doubled reaction rate and yield (Figure 7b). A time-resolved model was built on the process yield (Supporting Information S1: Figure S12), confirming that OAc^- leads to a higher reaction rate than Cl^- at higher Mg^{2+} concentrations.

3.7 | Maximizing yield

Finally, we combined our findings in an attempt to surpass the highest previously reported production yield of 12 g/L. Starting NTP concentrations in IVT matched the relative nucleotide abundance of the mRNA construct (25 mM ATP, 22.3 mM GTP, 21.4 mM CTP and 20.4 mM UTP). High-concentration NTPs stocks (200 mM) were used to minimize volume. Based on DOE findings, Mg:NTP ratio was 0.8, T7 RNAP and DNA were at 15 U/μL ('high') and 100 ng/μL ('high'), respectively. 1 U/μL RNase inhibitor and 1 U/mL pyrophosphatase were used. OAc^- was used as the Mg^{2+} counterion, and pH of 10 × IVT buffer (400 mM Tris, 20 mM spermidine, 100 mM DTT, pH 7.9) was adjusted with acetic acid rather than hydrochloric acid so as not to introduce inhibitory Cl^- counterion. DTT concentration in reaction was 10 mM since we observed a positive effect on reaction rate when increasing from 1 to 10 mM, while further increases to 20 mM did not improve the reaction (Supporting Information S1: Figure S13). mRNA concentration in high productivity IVT reaction was initially determined by CIMac PrimaS at different time points, but high viscosity of reaction mixture at late time points together with low sampling volumes resulted in increased quantification error (Supporting Information S1: Figure S14). Therefore, concentration in IVT reaction after 270 min was determined in diluted inactivated sample and by mass recovery from Oligo dT chromatography (Supporting Information S1: Figure S15): after 270 min, the reaction was quenched with EDTA and loaded onto a 1 mL Oligo dT column. Mass of polyadenylated mRNA in Oligo dT elution confirmed that on

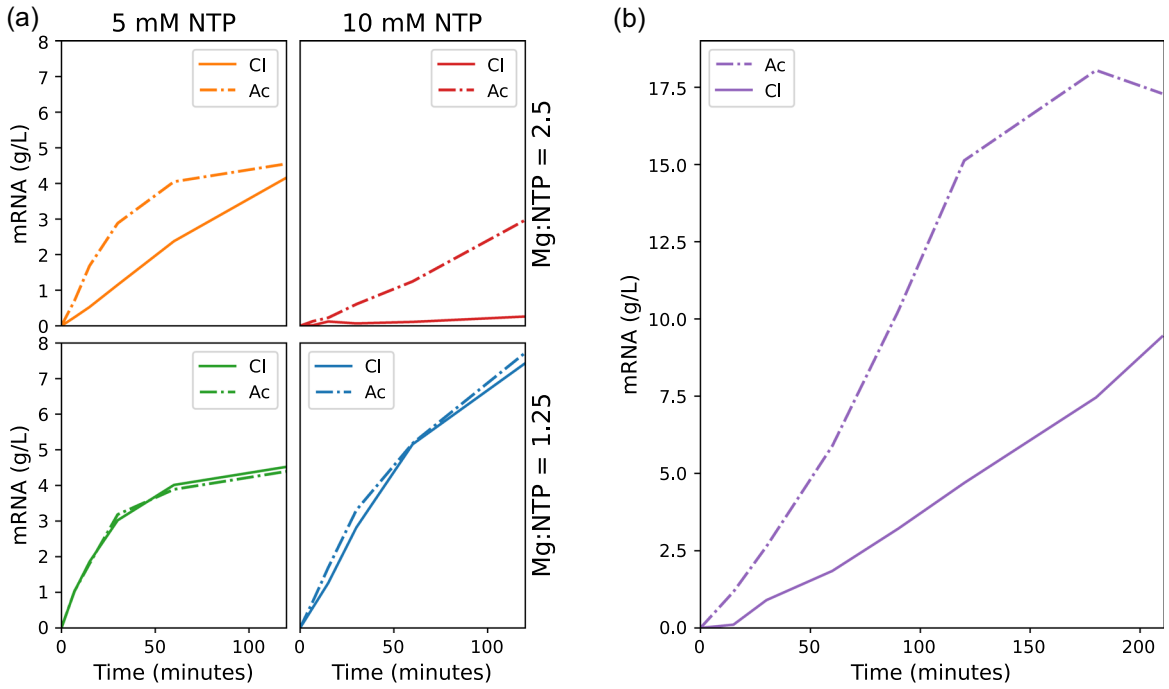

FIGURE 7 Effect of Mg^{2+} counterion on IVT kinetics and yield. (a) NTP concentration and Mg:NTP ratio was varied with chloride or acetate as the counterion. At Mg:NTP ratio of 1.25, there was no difference. At Mg:NTP ratio of 2.5, the reaction rate was substantially lower when using chloride, with the reaction having essentially stopped completely at 10 mM of each NTP. (b) The negative impact of Cl^- was verified in a high-productivity reaction. Mg:NTP ratio 0.8, 76 mM NTP in total (mix adjusted based on mRNA sequence), 100 ng/µL DNA, 15 U/µL T7 RNAP and 20 mM DTT were used. Reaction studying Cl^- used $MgCl_2$ and HCl for pH adjustment; reaction studying acetate used $Mg(OAc)_2$ and acetic acid for pH adjustment.

average, 24.9 ± 1.5 g/L (95% confidence level) was produced in IVT (Supporting Information S1: Table S12), corresponding to a cost efficiency of 173 µg/€, which is 84% higher compared to the reference protocol. Additionally, dsRNA was below detection by dot blot, consistent with the dsRNA model which predicted that high NTP at low Mg:NTP ratio leads to low dsRNA.

4 | DISCUSSION

We set out to optimize the IVT reaction for production of mRNA. Using a DOE approach, 23 experiments in three iterations were performed to derive a model with high R^2 and Q^2 values for both mRNA production and dsRNA estimation with the goal to identify reaction parameters required to exceed 12 g/L, maximize yield and decrease dsRNA and cost. The final concentration and dsRNA content at 210 min were used for optimization, but time-course data was collected for each reaction as well, enabling analysis of reaction kinetics, identification of outliers by comparing mRNA and NTP levels, and deriving a time-resolved model that could be used to predict reaction trajectories transferrable between mRNA constructs.

NTP concentration was selected as a factor since it directly impacts the maximum process yield of the IVT reaction due to the stoichiometric nature of the reaction. As such, Equation (1) should inform the design of IVT reactions by selecting NTP concentrations to target a desired final yield of mRNA. Since the fraction n/n_{LIM} is construct-specific, the maximum amount of mRNA that can be produced with a given NTP concentration varies between constructs. When mRNA construct contains exactly equal amounts of each NTP, $n/n_{LIM} = 4$, thus allowing for a maximum of 1.28 g/L of mRNA for each mM of NTP mixture. There seems to be a lack of regard in the field for the inherently stoichiometric nature of IVT reaction; yields that exceed theoretical maximum have been reported, suggesting that either NTP, mRNA concentration, or both, have been erroneously quantified, potentially leading to misinterpretation of factors governing IVT reaction (Table 1).

We used 'high' setting for NTP at 12 mM to reach process outputs of at least 12 g/L. The final model contains a square term and all interaction terms for NTP, suggesting that its effect on process yield is complex (Supporting Information S1: Figure S6). Thus, a change in [NTP] impacts the effects of all other factors on the process yield. Most importantly, we identified that optimal Mg:NTP ratio decreases with increasing [NTP] (Figure 3), possibly explaining the variation in reported optima in previous reports (Table 1). We conclude that to transfer conclusions between different IVT studies, information on NTP levels is essential for correct interpretation of reported findings, which may only be valid for the [NTP] range studied (Supporting Information S1: Figure S1).

The importance of [Mg^{2+}] was demonstrated in several previous studies (Table 1, Supporting Information S1: Figure S1), but its impact on the responses was not yet fully understood. Unlike prior approaches, we used the Mg:NTP ratio as a design factor rather than Mg^{2+} alone, because meta-analysis of prior literature pointed to a strong link between Mg^{2+} and NTP; total Mg^{2+} concentration 5–10 mM higher than total NTP concentration was reported multiple times, corresponding to molar ratio Mg:NTP between 1 and 2 (Table 1). When selecting the boundaries for DOE, we therefore selected the lower Mg:NTP ratio setting of 0.8 and upper setting of 1.6. Interestingly, modeling Mg:NTP ratio rather than absolute concentrations revealed multiple previously unexplained findings. Notably, the optimal Mg:NTP ratio changed as [NTP] were increased. For low [NTP] (6–8 mM) the optimal Mg:NTP ratio for maximum mRNA yield and reaction rate was approximately 1.2. This agrees with several previous findings (Kanwal et al., 2018; Kartje et al., 2021; Kern & Davis, 1997; Young et al., 1997) which suggest that total [Mg^{2+}] should be slightly above the total [NTP], and appears to be in line with the theory that elongation by T7 RNAP requires both $MgNTP^{2-}$ and Mg^{2+} ions; with a ratio of 1.2, NTP is predominantly present in solution as $MgNTP^{2-}$, and the remaining Mg^{2+} is available to the enzyme. However, at the highest [NTP] evaluated in our DOE (12 mM), the optimal Mg:NTP ratio was 0.9. At this ratio, a very low amount of Mg^{2+} should be available for binding to T7 RNAP, which should lead to a decrease in reaction rate, and suggests the presence of a competing effect that reduces the reaction rate even further at higher [Mg^{2+}]. The competing effect may be the increased ionic strength and higher concentration of Cl^-ions, which several earlier studies reported to have an impact on the rate and/or yield of IVT reaction via anions competing with DNA for binding sites on RNAP; most results indicate that OAc^- should be the preferred counterion due to higher inhibitory effect of Cl^- (Kern & Davis, 1997; Maslak & Martin, 1994; Young et al., 1997). We show that at very high [Mg^{2+}] and total [NTP] (100 mM and 40 mM, respectively, ratio 2.5), both reaction rates are lower compared to 1.25, but OAc^- is less inhibiting than Cl^- (Figure 7). On the contrary, at Mg:NTP ratio 1.25, there was a negligible difference in reaction rate between the two anions when using 10 mM of NTP. We can conclude that the choice of counterion is unlikely to affect rate or yield for IVT reaction designs targeting <12 g/L, while for >12 g/L, OAc^- is the preferred counterion.

Mg^{2+} was also reported to significantly affect dsRNA levels (Mu et al., 2018). It was reported that [Mg^{2+}] < 10 mM results in lower dsRNA content when dsRNA is formed by hybridization of run-off transcript and antisense transcript. Since the study did not quantify mRNA, the effect of decreased [Mg^{2+}] on mRNA production was unclear. We show that Mg:NTP ratio, rather than [Mg^{2+}] alone, influences dsRNA content, which has crucial implications for the productivity of the IVT reaction. Mg^{2+} levels above 70 mM resulted in negligible dsRNA content if [NTP] was also high (i.e., Mg:NTP ratio was low), leading to high productivity and high purity (Supporting Information S1: Figure S14). On the contrary, low [Mg^{2+}] at high Mg:NTP ratio resulted in high dsRNA content. Although the mechanism by which Mg^{2+} affects dsRNA formation is not yet understood, it is possible that high [Mg^{2+}] might promote aberrant T7 RNAP conformation which results in promoter-independent transcription and antisense product formation, as speculated by Mu et al. (2018). To our knowledge, we are the first to show that content of dsRNA can be lowered simply by reducing the Mg:NTP ratio in the IVT reaction.

Concentrations of DNA template and T7 RNAP were selected as DOE factors with highest cost; factor settings were based on (Pregeljc et al., 2023), with the modification of lowering the low setting slightly to investigate how limiting [DNA] affects kinetics, since this could have cost efficiency implications. Since the effect of T7 RNAP and DNA concentration on process yield was found to be similar (Figure 3), the optimal reaction setting depends on the cost difference between the two reagents. At the current pricing information available, the cost-optimal point has a lower level of T7 RNAP; in case of future price changes, the model can be used to adjust [T7]/[DNA] to maintain price performance without impacting productivity or quality—in agreement with (Mu et al., 2018) we show that concentrations of T7 RNAP and DNA do not affect dsRNA content.

By verifying the design space, we showed that within this space, process parameters can be changed while satisfying selected requirements. The optimal setpoint improved cost efficiency by 44% compared to starting reference IVT protocol, while decreasing dsRNA content.

5 | CONCLUSIONS

In this study, a QbD approach was used to determine a cost-optimal recipe for mRNA production in IVT reaction. An MLR model was fitted to the data and used to find optimal settings for the four investigated factors, constrained to a design space limited by the amount of the by-product dsRNA, and the design space and proposed optimal settings were verified experimentally. We defined a stoichiometric relationship between [NTP] and maximum theoretical reaction yield (in g/L), with an approximate yield of 1.1 g/L for each 1 mM of an equimolar NTP mix. We developed a methodological DOE approach to design and optimize a process for mRNA production and identified previously unappreciated interactions in the IVT reaction that have important effects on the reaction kinetics and process output. In particular, we demonstrate for the first time the importance of Mg:NTP ratio for reaction yield. We observed that the optimal Mg:NTP ratio decreased for high NTP concentrations. Furthermore, we show that it is possible to balance the cost drivers DNA and T7 RNAP within the design space and still achieve high process yield, allowing for post-approval cost optimization without triggering change notifications to regulating authorities.

DOE results confirm an optimal recipe for cost efficiency within the investigated ranges for all four factors, which improved cost efficiency by 44% over the reference recipe, while reducing the amount of dsRNA. Mg^{2+} was previously thought to increase dsRNA levels; we show that this is Mg:NTP ratio-dependent, with important

implications for productivity. Low Mg:NTP ratio decreased dsRNA levels, while keeping productivity high.

We studied the effect of Mg^{2+} counterion on IVT yield to unify previous reports; OAc^- was found to be the preferred counterion for process designs that would require $[Mg^{2+}] > 50$ mM, that is, for high-productivity reaction conditions, while selection of counterion was not important for reaction designs which target <12 g/L.

We combined the findings on Mg:NTP ratio, counterion, stoichiometry and DTT in batch IVT to reach 25 g/L mRNA, doubling the currently highest reported IVT yield. The cost efficiency of this reaction design is 84% higher than the reference IVT protocol, pointing to a need for future DOE exercises in ultrahigh producing reaction conditions with potential for further cost-savings. The model can in the future also be expanded both by number of factors and CQAs, especially for constructs susceptible to degradation (e.g., saRNA), where reaction time can be crucial for product quality. Models can also be built for different IVT strategies (e.g., co-transcriptional capping, fed-batch reaction design). We also suggest a strategy to adapt our data-driven regression model for eGFP to predict mRNA production of other constructs by adjusting for molarity and construct size. This could serve as a basis for future development of a digital twin for the IVT reaction.

AUTHOR CONTRIBUTIONS

Jimmy Boman: Methodology, investigation, writing—original draft, writing—review and editing. **Tjaša Marušič**: Methodology, investigation, writing—original draft, writing—review and editing. **Tina Vodopivec Seravalli**: Methodology, investigation. **Janja Skok**: Methodology, investigation, writing—review and editing. **Fredrik Pettersson**: Methodology, investigation. **Kristina Šprinzar Nemec**: Methodology, investigation. **Henrik Widmark**: Methodology, investigation, supervision, writing—original draft, writing—review and editing, project administration. **Rok Sekirnik**: Conceptualization, supervision, project administration, writing—original draft, writing—review and editing.

ACKNOWLEDGMENTS

Authors would like to thank Biomay AG for providing plasmid encoding for eGFP mRNA constructs, Prof. Tomaž Bratkovič (Faculty of Pharmacy, University of Ljubljana) and Lennart Eriksson (Sartorius SSDA) for constructive feedback during manuscript preparation.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Tjaša Marušič https://orcid.org/0009-0004-6085-7839
Rok Sekirnik http://orcid.org/0000-0001-6008-9500

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), 548–561. https://doi.org/10.1002/1097-0290(20010305)72:5<548::AID-BIT1019>3.0.CO;2-2

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), Article 6. https://doi.org/10.1038/s41587-022-01294-2

Chamberlin, M., & Ring, J. (1973). Characterization of T7-specific ribonucleic acid polymerase. *Journal of Biological Chemistry*, *248*(6), 2235–2244. https://doi.org/10.1016/S0021-9258(19)44211-7

Cho, E., Namgung, J., Lee, J. S., Jang, J., Jun, B.-H., & Kim, D.-E. (2023). Mesoporous silica particle as an RNA adsorbent for facile purification of in vitro-transcribed RNA. *International Journal of Molecular Sciences*, *24*(15), 12408. https://doi.org/10.3390/ijms241512408

Daniel, S., Kis, Z., Kontoravdi, C., & Shah, N. (2022). Quality by design for enabling RNA platform production processes. *Trends in Biotechnology*, *40*(10), Article 10. https://doi.org/10.1016/j.tibtech.2022.03.012

Dousis, A., Ravichandran, K., Hobert, E. M., Moore, M. J., & Rabideau, A. E. (2023). An engineered T7 RNA polymerase that produces mRNA free of immunostimulatory byproducts. *Nature Biotechnology*, *41*(4), 560–568. https://doi.org/10.1038/s41587-022-01525-6

EMA/604040/2016EMA/CHMP/CVMP/QWP/354895/2017. (n.d.).

Eriksson, L. (2008). *Design of experiments: Principles and applications* (3rd rev. and enl. ed). Umetrics.

Gholamalipour, Y., Johnson, W. C., & Martin, C. T. (2019). Efficient inhibition of RNA self-primed extension by addition of competing 3′-capture DNA-improved RNA synthesis by T7 RNA polymerase. *Nucleic Acids Research*, *47*(19), e118. https://doi.org/10.1093/nar/gkz700

Gößringer, M., Helmecke, D., Köhler, K., Schön, A., Kirsebom, L. A., Bindereif, A., & Hartmann, R. K. (2014). Enzymatic RNA synthesis using bacteriophage T7 RNA polymerase. In R. K. Hartmann, A. Bindereif, A. Schön, & E. Westhof (Eds.), *Handbook of RNA biochemistry* (1st ed., pp. 1–28). Wiley. https://doi.org/10.1002/9783527647064.ch1

Guth-Metzler, R., Mohamed, A. M., Cowan, E. T., Henning, A., Ito, C., Frenkel-Pinter, M., Wartell, R. M., Glass, J. B., & Williams, L. D. (2023). Goldilocks and RNA: Where Mg^{2+} concentration is just right. *Nucleic Acids Research*, *51*(8), 3529–3539. https://doi.org/10.1093/nar/gkad124

Hengelbrock, A., Schmidt, A., Helgers, H., Vetter, F. L., & Strube, J. (2023). Scalable mRNA machine for regulatory approval of variable scale between 1000 clinical doses to 10 million manufacturing scale doses. *Processes*, *11*(3), Article 3. https://doi.org/10.3390/pr11030745

Kanwal, F., Chen, T., Zhang, Y., Simair, A., Rujie, C., Sadaf Zaidi, N. 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

Karikó, K., Muramatsu, H., Welsh, F. A., Ludwig, J., Kato, H., Akira, S., & Weissman, D. (2008). Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability. *Molecular Therapy*, *16*(11), Article 11. https://doi.org/10.1038/mt.2008.200

Kartje, Z. J., Janis, H. I., Mukhopadhyay, S., & Gagnon, K. T. (2021). Revisiting T7 RNA polymerase transcription in vitro with the broccoli RNA aptamer as a simplified real-time fluorescent reporter. *Journal*

of Biological Chemistry, *296*, 100175. https://doi.org/10.1074/jbc.RA120.014553

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

Kis, Z., Kontoravdi, C., Shattock, R., & Shah, N. (2020). 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

Kis, Z., Tak, K., Ibrahim, D., Daniel, S., Van De Berg, D., Papathanasiou, M. M., Chachuat, B., Kontoravdi, C., & Shah, N. (2022). Quality by design and techno-economic modelling of RNA vaccine production for pandemic response, *Computer aided chemical engineering* (Vol. 49, pp. 2167–2172). Elsevier. https://doi.org/10.1016/B978-0-323-85159-6.50361-4

Martin, C. T., & Coleman, J. E. (1987). Kinetic analysis of T7 RNA polymerase-promoter interactions with small synthetic promoters. *Biochemistry*, *26*(10), 2690–2696. https://doi.org/10.1021/bi00384a006

Maslak, M., & Martin, C. T. (1994). Effects of solution conditions on the steady-state kinetics of initiation of transcription by T7 RNA polymerase. *Biochemistry*, *33*(22), 6918–6924. https://doi.org/10.1021/bi00188a022

Mu, X., Greenwald, E., Ahmad, S., & Hur, S. (2018). An origin of the immunogenicity of in vitro transcribed RNA. *Nucleic Acids Research*, *46*(10), 5239–5249. https://doi.org/10.1093/nar/gky177

Mu, X., & Hur, S. (2021). Immunogenicity of in vitro -transcribed RNA. *Accounts of Chemical Research*, *54*(21), 4012–4023. https://doi.org/10.1021/acs.accounts.1c00521

Nacheva, G. A., & Berzal-Herranz, A. (2003). Preventing nondesired RNA-primed RNA extension catalyzed by T7 RNA polymerase. *European Journal of Biochemistry*, *270*(7), 1458–1465. https://doi.org/10.1046/j.1432-1033.2003.03510.x

Nagaraj, S., Stankiewicz-Drogon, A., Darzynkiewicz, E., & Grzela, R. (2022). RNA sensor response in HeLa cells for transfected mRNAs prepared in vitro by SP6 and HiT7 RNA polymerases: A comparative study. *Frontiers in Bioengineering and Biotechnology*, *10*, 1017934. https://doi.org/10.3389/fbioe.2022.1017934

Nelson, J., Sorensen, E. W., Mintri, S., Rabideau, A. E., Zheng, W., Besin, G., Khatwani, N., Su, S. V., Miracco, E. J., Issa, W. J., Hoge, S., Stanton, M. G., & Joyal, J. L. (2020). Impact of mRNA chemistry and manufacturing process on innate immune activation. *Science Advances*, *6*(26), eaaz6893. https://doi.org/10.1126/sciadv.aaz6893

Piao, X., Yadav, V., Wang, E., Chang, W., Tau, L., Lindenmuth, B. E., & Wang, S. X. (2022). Double-stranded RNA reduction by chaotropic agents during in vitro transcription of messenger RNA. *Molecular Therapy—Nucleic Acids*, *29*, 618–624. https://doi.org/10.1016/j.omtn.2022.08.001

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

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), Article 16. https://doi.org/10.1016/j.vaccine.2021.03.038

Sahin, U., Karikó, K., & Türeci, Ö. (2014). mRNA-based therapeutics—Developing a new class of drugs. *Nature Reviews Drug Discovery*, *13*(10), 759–780. https://doi.org/10.1038/nrd4278

Samnuan, K., Blakney, A. K., McKay, P. F., & Shattock, R. J. (2021). Design-of-experiments in vitro transcription yield optimization of self-amplifying RNA (2021.01.08.425833). bioRxiv. https://doi.org/10.1101/2021.01.08.425833

Skok, J., Megušar, P., Vodopivec, T., Pregeljc, D., Mencin, N., Korenč, M., Krušič, A., Celjar, A. M., Pavlin, N., Krušič, J., Mueller, M., McHugh, K., Štrancar, A., & Sekirnik, R. (2022). Gram-scale mRNA production using a 250-mL single-use bioreactor. *Chemie Ingenieur Technik*, *94*(12), Article 12. https://doi.org/10.1002/cite.202200133

Thomen, P., Lopez, P. J., Bockelmann, U., Guillerez, J., Dreyfus, M., & Heslot, F. (2008). T7 RNA polymerase studied by force measurements varying cofactor concentration. *Biophysical Journal*, *95*(5), 2423–2433. https://doi.org/10.1529/biophysj.107.125096

Wu, M. Z., Asahara, H., Tzertzinis, G., & Roy, B. (2020). Synthesis of low immunogenicity RNA with high-temperature in vitro transcription. *RNA*, *26*(3), 345–360. https://doi.org/10.1261/rna.073858.119

Yin, Y., & Carter, C. W. (1996). Incomplete factorial and response surface methods in experimental design: Yield optimization of tRNATrp from in vitro T7 RNA polymerase transcription. *Nucleic Acids Research*, *24*(7), 1279–1286. https://doi.org/10.1093/nar/24.7.1279

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

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. *Frontiers in Molecular Biosciences*, *10*, 1291045. https://doi.org/10.3389/fmolb.2023.1291045

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Boman, J., Marušič, T., Seravalli, T. V., Skok, J., Pettersson, F., Nemec, K. Š., Widmark, H., & Sekirnik, R. (2024). Quality by design approach to improve quality and decrease cost of in vitro transcription of mRNA using design of experiments. *Biotechnology and Bioengineering*, 1–13. https://doi.org/10.1002/bit.28806