Modelling cell-free RNA and protein synthesis with minimal systems

Doerr, Anne; de Reus, Elise; van Nies, Pauline; van der Haar, Mischa; Wei, Katy; Kattan, Johannes; Wahl, Aljoscha; Danelon, Christophe

DOI
10.1088/1478-3975/aaf33d

Publication date
2019

Document Version
Accepted author manuscript

Published in
Physical Biology

Citation (APA)
Doerr, A., de Reus, E., van Nies, P., van der Haar, M., Wei, K., Kattan, J., Wahl, A., & Danelon, C. (2019). Modelling cell-free RNA and protein synthesis with minimal systems. *Physical Biology*, *16*(2), Article 025001. https://doi.org/10.1088/1478-3975/aaf33d

Modelling cell-free RNA and protein synthesis with minimal systems

Anne Doerr[1], Elise de Reus[1,2], Pauline van Nies[1], Mischa van der Haar[1], Katy Wei[1], Johannes Kattan[1], Aljoscha Wahl[2], Christophe Danelon[1]

[1]Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, van der Maasweg 9, 2629 HZ, Delft, The Netherlands
[2]Department of Biotechnology, Delft University of Technology, van der Maasweg 9, 2629 HZ, Delft, The Netherlands

*Corresponding author: E-mail: c.j.a.danelon@tudelft.nl
Tel.: +31-152788085, Fax: +31-152781202

ABSTRACT

DNA-guided cell-free protein synthesis using a minimal set of purified components has emerged as a versatile platform in constructive biology. The *E. coli*-based PURE (Protein synthesis Using Recombinant Elements) system offers the basic protein synthesis factory in a prospective minimal cell relying on extant molecules. However, it becomes urgent to improve the system's performance, and to build a mechanistic computational model that can help interpret and predict gene expression dynamics. Herein, we utilized all three commercially available PURE system variants: PURExpress, PURE*frex* and PURE*frex*2.0. We monitored apparent kinetics of mRNA and protein synthesis by fluorescence spectroscopy at different concentrations of DNA template. Analysis of polysome distributions by atomic force microscopy, combined with a stochastic model of translation, revealed inefficient usage of ribosomes, consistent with the idea that translation initiation is a limiting step. This preliminary dataset was used to formulate hypotheses regarding possible mechanisms impeding robust gene expression. Next, we challenged these hypotheses by devising targeted experiments aimed to alleviate the current limitations of PURE*frex*. We identified depletion of key initiation factors by translationally inactive mRNA as a possible inhibitory mechanism. This adverse process could partly be remedied by targeted mRNA degradation, whereas addition of more IFs and of the hrpA RNA helicase had no substantial effects. Moreover, depletion of tRNAs as peptidyl-tRNAs can become limiting in PURE*frex* (but not in PURExpress), which can be alleviated by addition of peptidyl-tRNA-hydrolase (PTH). We attempted to build a new model for PURE system dynamics integrating all experimental observations. Although a satisfying global fit can be obtained in specific conditions (with PTH), a unifying system's level model is still missing.

INTRODUCTION

In vitro gene expression, also commonly referred as coupled *in vitro* transcription-translation (IVTT) or cell-free protein synthesis, has become a rapidly growing research area since its introduction in the 1960's (1, 2) and further developments (3, 4). This technology is now employed in a broad spectrum of applications (5-8), e.g. in *in vitro* diagnostic (5, 6), and has contributed significant advances in our understanding of the fundamental biology of genetic information transfer. IVTT systems are now exploited as a synthetic biology toolkit for the realization of basic genetic circuits (9-14), demonstrating their potential to recapitulate complex biological phenomena directed by the execution of a DNA program *in vitro*. Furthermore, it was proposed that IVTT, in conjunction with liposome compartmentalization, forms the scaffold to build a synthetic minimal cell relying on the extant biology (15-20).

It is however fair to say that the robustness and efficiency of IVTT systems with respect to translation rate, usage of resources and duration of expression in batch mode are very modest compared to *in vivo* reactions (21). Moreover, the expression of multiple – interacting – genes in a predictable manner remains a challenge. To leverage the potential of IVTT in bottom-up synthetic biology, it becomes critical to identify the set of parameters that controls the system's performance and to devise optimization strategies to execute targeted dynamical behaviours, as has been done in the *in vitro* DNA-based programming community (22, 23).

IVTT systems come in two main guises: one derived from cell extracts and the other consisting of purified components. An obvious advantage of crude cell lysates is the ease of manufacture making the approach economically attractive (24, 25). Cellular extracts inherently contain various cytoplasmic compounds that may assist some steps in protein synthesis, like protein folding. However, metabolic resources, such as nucleoside triphosphates (NTPs) and amino acids, are also consumed in non-intended reactions, and remaining nucleases and proteases can cause undesired degradation of nucleic acids and proteins. These effects can limit the yield of synthesized products and shorten the production lifespan. Alternatively, the reconstitution of the transcription and translation apparatus from essential purified components enables effective channelling of nutrients and energy to gene expression and eliminates enzymatic degradation of nucleic acids and proteins (26, 27). The concoction, termed 'Protein synthesis Using Recombinant Elements', or PURE system, consists of individually purified reagents, including the *E. coli* ribosome, translation factors, tRNAs, aminoacyl tRNA synthetases, the phage T7 RNA polymerase and a set of enzymes to regenerate NTPs from nucleoside mono- and diphosphate and creatine phosphate (26). The reaction network is driven out of equilibrium by the enzymatic hydrolysis of pyrophosphate. Both linear (PCR products) and circular DNA can be used as templates. The PURE system is currently available commercially in three variants: The PURExpress from New England Biolabs, PURE*frex* and PURE*frex*2.0 from GeneFrontier Corporation. The composition of PURExpress has been modified from the original version to improve protein synthesis yield, but the supplier does not disclose this information. PURE*frex* offers easier protein purification as the kit proteins are non-his-tagged and it contains low-level contaminations, including of lipopolysaccharides. The

advanced version PURE*frex*2.0 achieves higher protein yield than PURE*frex*. The kit modifications that lead to improved productivity are currently not disclosed by the company.

Because all IVTT systems are stripped of the cellular context, they might intuitively be expected to be more amenable to quantitative description than *in vivo* systems. First, gene expression can be studied independently of physiological factors, such as growth state and cell density. Second, the system's openness and flexibility enable controlled inputs of substrates, metabolites and protein cofactors, which facilitates the screening of initial conditions. Third, unlike the *in cellulo* environment, random fluctuations due to the low-copy number of reactants can be neglected, making deterministic treatment of gene expression kinetics a valid approximation. Therefore, it is tempting to believe that gene expression dynamics and the system's performance would enjoy greater predictability in IVTT platforms than *in vivo*. With the advent of systems biology, along with the development of computational tools and mathematical models (28), one would also expect that IVTT systems – in particular the PURE system due to its defined composition – are amenable to full numerical description. In practice, IVTT systems suffer from batch to batch variability (29) and experiments are designed mostly based on empirical observations (30) more than from established guiding rules.

A comprehensive and predictive model of cell-free gene expression will arguably facilitate the rational design of new targeted experiments and the engineering of IVTT systems with tailored properties. Approaches to model IVTT reactions have varied considerably, from explicit modelling of 'every' elementary reaction (31-34) to phenomenological models that describe salient features of the observed behaviour instead of explaining their microscopic origins (35-38). The former approach is fundamentally more accurate and provides detailed mechanistic insights about the various steps. However, when applied to the complex reaction network of IVTT, such high-dimensional models can become cumbersome, hardly tractable and prohibitively computationally expensive. Moreover, assigning initial values for the numerous reaction rate constants is challenging as most parameters are either not found in the literature or they have been evaluated from different experimental conditions. More coarse-grained kinetic models represent a promising framework, especially when dynamic information about all individual components is not required. In general, model building and validation should integrate a variety of data sets, each reporting on the dynamics of subsystems, in order to provide a quantitatively accurate description of the IVTT kinetics. However, most of the existing models have been constructed on the basis of the apparent time course production of fluorescent proteins under narrow-ranging experimental conditions.

Herein, we adopted a reverse engineering approach to interrogate PURE system dynamics under a wide range of experimental conditions with the ultimate goal to devise a mathematical model. By combining polysome profiling, protein gel analysis of translation products, mass spectrometry analysis of key metabolites and synthesized peptides, and real-time gene expression kinetics at both mRNA and active protein levels, we identified a number of processes that impair sustained production. Attempts to build a whole-system computational model reveal that a new

network structure and/or unaccounted mechanisms should be implemented to accurately describe and predict gene expression in the PURE system.

MATERIALS AND METHODS

DNA constructs

A codon-optimized version (*meYFPco-LL-spinach*) of the *mYFP-LL-spinach* construct described in (39) was used for all experiments except the metabolite mass spectrometry and temperature-dependent kinetic measurements for which the non-optimized version was used. The codon-optimized DNA template was designed and constructed in a pUC57 vector by Eurogentec (Belgium). Sequence optimisation was performed from the start to the stop codon and involved improving codon usage to the *E. coli*-based expression system and avoiding RNA secondary structures (among other criteria, not all of them are disclosed by the company). A linear PCR product (forward primer: GCGAAATTAATACGACTCACTATAGGGAGACC, reverse primer: AAAAAACCCCTCAAGACCCGTTTAGAGG) including a T7 promoter, ribosome binding site, start codon, 6×His-tag, g10-leader, Xpress-tag, meYFP coding sequence, linker, spinach aptamer sequence and a T7 terminator was generated. DNA was purified from PCR reactions using the Wizard PCR clean-up kit (Promega) using milliQ for the final elution step and its concentration was determined by measuring absorbance at 260 nm. The full sequence of the codon-optimized *meYFPco-LL-spinach* linear construct is: 5'-

GCGAAATTAATACGACTCACTATAGGGAGACCACAACGGTTTCCCTCTAGAAATAAT
TTTGTTTAACTTTAAGAAGGAGATATACATATGCGGGGTTCTCATCATCATCATCATC
ATGGTATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGGGATCTGTACGACGAT
GACGATAAGGATCCGATGGTTAGCAAAGGCGAAGAACTGTTTACGGGCGTGGTGCC
GATTCTGGTGGAACTGGACGGCGACGTGAACGGTCACAAATTCAGCGTTTCGGGCG
AAGGTGAAGGCGATGCGACCTATGGTAAACTGACGCTGAAATTTATTTGCACCACC
GGTAAACTGCCGGTGCCGTGGCCGACCCTGGTTACCACGTTTGGTTATGGCCTGCAG
TGTTTCGCGCGCTACCCGGATCATATGAAACAACACGACTTTTTCAAATCTGCCATG
CCGGAAGGTTATGTGCAGGAACGTACGATTTTCTTTAAAGATGACGGCAACTACAA
AACCCGCGCAGAAGTCAAATTTGAAGGTGATACGCTGGTGAACCGTATTGAACTGA
AAGGCATCGATTTCAAAGAAGACGGTAATATCCTGGGCCATAAACTGGAATACAAC
TACAACTCCCACAACGTTTACATCATGGCAGATAAACAGAAAAACGGTATCAAAGT
CAACTTCAAAATCCGCCATAACATCGAAGATGGCTCAGTGCAACTGGCTGACCACT
ACCAGCAAAACACCCCGATCGGTGATGGCCCGGTTCTGCTGCCGGACAATCATTATC
TGAGCTACCAGTCTAAACTGAGTAAAGATCCGAACGAAAAACGTGACCACATGGTC
CTGCTGGAATTTGTGACGGCGGCTGGTATTACGCTGGGCATGGATGAACTGTATAAA
TGAAAGCTTCCCGGGAAAGTATATATGAGTAAAGATATCGACGCAACTGAATGAAA
TGGTGAAGGACGGGTCCAGGTGTGGCTGCTTCGGCAGTGCAGCTTGTTGAGTAGAGT

GTGAGCTCCGTAACTAGTCGCGTCGATATCCCCGGGCTAGCATAACCCCTTGGGGCC
TCTAAACGGGTCTTGAGGGGTTTTTT-3’

The oligonucleotide used for experiments with mRNA turnover has the following sequence: 5’-CATATGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGG-3’

To obtain the *meYFPco-LL-spinach* construct with CGT as the second codon (DNA sequence), site directed mutagenesis PCR with the primers GGAGATATACATATGCGTGGTTCTCATCATCATC and GATGATGATGAGAACCACGCATATGTATATCTCC was performed with the non-mutated codon optimized plasmid as template. PCR products were digested with DpnI, purified and transformed into top10 cells. Mutation was confirmed by sequencing of the purified plasmid and a linear product was generated with the same primers as those used with the standard construct.

PURE system reactions

PURExpress was purchased from New England Biolabs, PURE*frex* and PURE*frex*2.0 were from GeneFrontier Corporation (Japan). Reaction solutions of typically 20 µl were assembled according to the supplier’s protocol.

Fluorescence kinetics of spinach/DFHBI and YFP

To measure gene expression kinetics in different PURE systems at both mRNA and protein levels simultaneously, the reporter construct *meYFPco-LL-spinach* was used. The spinach RNA aptamer (40) was introduced in the 3’-untranslated region as a fluorescent reporter in the presence of 3,5-difluoro-4-hydroxybenzylidene (DFHBI) (39, 41). 20 µl of PURE system reaction mix containing 20 µM DFHBI (40) and specified concentrations of purified DNA or RNA template was prepared on ice and then transferred to a 15-µl cuvette (Hellma). Cuvettes were mounted in a temperature-controlled cell of a Varian Eclipse fluorescence spectrophotometer and fluorescence was measured every 30 s with the following excitation/emission wavelengths: spinach, 460/502 nm; YFP, 515/528 nm. Temperature was set to 37 °C, unless indicated differently.

Phenomenological fitting was carried out by fitting the linear regime of the YFP fluorescence traces and calculating the intersection with the horizontal line denoting the maximum yield. For each curve the maximum translation rate, lifespan and yield are extracted (Fig. 1(c,d)).

Calibration of spinach and YFP concentrations

Calibrations were performed as described in (39). Conversion factors of 4.3 nM/a.u. and 0.35 nM/a.u. were found for mRNA and YFP, respectively.

Quantitative metabolite analysis by mass-spectrometry

180 µl of a standard PURE*frex* reaction supplemented with 7.4 nM *mYFP-LL-spinach* DNA were prepared on ice and incubated at 37 °C. Samples of 20 µl were taken at time points 0, 30, 60, 90, 120, 150 and 180 min, mixed with 50 µl ^{13}C internal standard (42, 43) and 180 µl milliQ,

and snap-frozen in liquid nitrogen. All samples were stored at –20 °C until analysis. Before analysis the samples were filtered with a 10-kDa cut-off spin-filter (Vivacon500). The nucleotide concentration was measured by 5 µl injection into an LC-MS/MS system (Waters, for details of the setup please see (44)) and ID-MS based quantification (45). Quantification was based on standard curves obtained with NTPs or NDPs in 50 mM HEPES/KOH pH7.6, 100 mM potassium-glutamate, 13 mM magnesium acetate.

Amino-acid concentrations were determined using GC-MS analysis. A sample of 50 µl was freeze-dried and derivatised by first incubation with 50 µl of 20 g/l O-methoxyamine hydrochloride for 50 min and then 80 µL of MSTFA-TMCS for another 50 min. The sample was injected (1 µl) in a PTV inlet (Gerstel). The GC-oven was at 70 °C at the start of the run (7890A, Agilent Technologies). The GC column (Zebron ZB-50 column, 30 m×250 µm, 0.25 µm film thickness; Phenomenex, Torrance, CA, USA) was connected to an MS system (5975C Agilent Technologies, Santa Clara, CA, USA). Further details can be found in (45).

In vitro expression and purification of RNA template

mRNA was produced from the linear *meYFPco-LL-spinach* construct with the Promega T7 RiboMAX kit according to the manufacturer's instruction. The mRNA was purified with the Qiagen RNeasy kit using milliQ for the final elution step. Concentration was determined by absorbance measurement at 260 nm.

Polysome analysis by atomic force microscopy

PURE system reactions were incubated at 37 °C and 1 µl samples were taken at different time points. Samples were immediately added to 9 µl of ice-cold AFM buffer (3% sucrose, 20 mM Tris-HCl pH 7.4, 10 mM $MgCl_2$, 100 mM NH_4Cl, 1.1% (v/v) 2-mercaptoethanol and 100 µg/ml chloramphenicol) to stop translation. The next steps of the protocol were as described in (46). Samples were stored at 4 °C for up to 2 days and imaged in air on freshly cleaved mica after settling for 30 s, with a Bruker Multimode AFM equipped with a Nanoscope IIIa controller and a SCANASYST-FLUID+ cantilever (Bruker). Each field of view was 1.46 µm × 1.46 µm and 512 by 512 pixels. Scan speed was set to 1.97 Hz. For each time point about 15 fields of view were imaged to collect sufficient statistics (more than 1000 ribosomes per time point). A custom MatLab (MathWorks) script was used for image analysis. Images were background-corrected by first subtracting a fitted plane, then subtracting a line-by-line median offset, and finally subtracting a fitted 4×2 polynomial using MatLab's built-in *fit* function. Data points with more than a 1.8-fold standard deviation difference from the mean were ignored for calculating the background subtractions. A median filter of size 1×3 was then applied to remove stripes. To find particles a 'spot finding' algorithm as described in (47) was used. The maximum value of each particle in the background-subtracted image was taken as the particle height. Particle height distributions were generated (Fig. 7(b)) and a cut-off was defined to select only 50S and 70S particles for the calculation of polysome distributions. Cut-off values varied between images acquired on different days. Clusters were defined as all 50S and 70S particles with a maximum

distance smaller than 13 pixels (37 nm) to any other particle in the cluster (the effect on the calculated polysome distribution is small when varying this maximum distance between 11 and 15 pixels). To account for random proximity of particles, the distributions of cluster sizes were simulated for each field of view with the same number of particles as identified in the image and random placement of particles (random except disallowing too close proximity which the image analysis algorithm would also not be able to resolve as distinct particles). Average cluster size distributions for 100 simulated random placements of particles were calculated. Polysome distributions were computed by subtracting the simulated random cluster sizes from the observed ones. The concentration of polysomes was calculated based on the known concentration of ribosomes and on the fraction of ribosomes involved in polysomes of different sizes.

Targeted mRNA cleavage

Reactions with mRNA degradation contained 10 µl PURE*frex* solution I, 1 µl PURE*frex* solution II, 1 µl PURE*frex* solution III, 1 µl DFHBI (0.4 mM), 0.75 µl RNaseH (5 U/µl, New England BioLabs), 1 µl antisense oligo (1 µM), 0.5 µl T7 RNA polymerase (1.8 µM), 0.25 µl UTP (100 mM), 0.25 µl CTP (100 mM), 0.25 µl $MgAc_2$ (200 mM) and milliQ up to 20 µl.

Purification of hrpA, IF1, IF2, PTH and T7 RNA polymerase mutant

The enzyme hrpA was purified as described earlier (48). The genes for IF1, IF2 and *E. coli* peptidyl-tRNA-hydrolase (PTH) were cloned into pRSET-B (N-terminal His-tag) and expressed in BL21(DE3) *E. coli* cells. Cultures of 50 ml in LB/amp medium were grown at 37 °C and 250 rpm. At OD_{600}=0.4 cells were induced with 0.3 mM IPTG and grown overnight at 20 °C. Cells were harvested by centrifugation, the pellet was resuspended in lysis buffer (50 mM HEPES-KOH pH7.5, 500 mM NaCl, 25 mM imidazole), lysed by sonication (25% amplitude, 10 s on, 30 s off, repeated 10 times) and centrifuged for 30 min at 4 °C and 16000 g. The resulting lysate was applied to a NiNTA mini spin column (Qiagen) and washed twice with lysis buffer according to manufacturer's instructions. The protein was eluted with 50 mM HEPES KOH, 500 mM NaCl and 500 mM imidazole. The protein was stored at –80 °C in 50 mM HEPES-KOH pH 7.5, 100 mM KCl, 10 mM $MgCl_2$, 30% glycerol and 7 mM 2-mercaptoethanol. The gene of the T7 RNA polymerase mutant (P266L/I810S) was cloned in pBAD33 for overexpression in *E. coli* Top10 cells and the protein was purified as previously described (49). Storage buffer consist of 50 mM HEPES-KOH pH 7.5, 100 mM KCl, 10 mM $MgCl_2$, 30% glycerol and 7 mM 2-mercaptoethanol.

Stochastic model for polysome distributions

A Gillespie algorithm with five different types of reaction was used to simulate the distributions of polysomes (Table 1): 1) mRNA synthesis, 2) ribosome binding to the RBS, 3) ribosome inactivation, 4) YFP maturation, 5) ribosome translocation on an mRNA. The first four reaction types are simple reactions, implemented according to the standard Gillespie algorithm (next reaction method). mRNA molecules are generated with a rate derived from the spinach

fluorescence kinetics (Fig. 1). Ribosomes can bind to the RBS and move forward on the mRNA if there is at least a 15-codon distance to the next ribosome on the same mRNA molecule. The model keeps track of positions of ribosomes on all mRNAs. This is implemented in Mathematica (Wolfram Research, version 10.2.0.0) as a *list of lists*, where each mRNA is described as a list of the positions (in codons) where ribosomes are bound (free mRNAs are denoted as empty lists). When reaction type 5 is triggered a translating ribosome is randomly selected and if there is no ribosome blocking it from moving forward (minimum distance of 15 codons) it moves one codon. If the moving ribosome is situated at the last codon it unbinds from the mRNA and the count of full-length non-fluorescent proteins is incremented by one. The last step of chromophore maturation will generate fluorescent YFP. Ribosomes can be inactivated during either the free or bound state. If inactivation occurs during translation, the ribosome is modelled as falling off the mRNA so that subsequent ribosomes are not blocked. Parameters were fitted manually to match the observed fluorescence curve as well as inactivation data for PURExpress as measured by the delayed DNA addition experiments. The reaction rates and parameter values are reported in Table 1.

Phenomenological modelling of gene expression kinetics

The salient features of the YFP fluorescence kinetics have been parametrized as depicted in Fig. 1(c). We focused on three parameters that describe the experimental data in a compact way: the maximum slope, yield and expression lifespan. The linear regime of the curve was determined by using Durbin-Watson statistics selecting for a test value larger than 1.8 and maximum window size. The maximum yield was either directly determined from the final measurement values (when the curve shows a clear plateau) or by using a fit of the last part of the fluorescence curve to the function $\frac{y\,c(t)}{c(t)+x}$, where $c(t)$ denotes the measured YFP concentration, x and y are fitting parameters and the fitted value for y was taken as the maximum yield.

Deterministic model of gene expression kinetics

An ODE model of the schematic depicted in Fig. 12(a) was implemented in Mathematica 10.2.0.0 with the following equations:

$$
\begin{aligned}
\frac{\partial \mathrm{mRNA1}(t)}{\partial t} &= -\left(k_{on1}\mathrm{IF}(t)\mathrm{mRNA1}(t) - \mathrm{mRNA1IF}(t)\left(k_{ini1}\frac{\mathrm{tRNA}}{K_m+\mathrm{tRNA}} + k_{off1}\right)\right) - k_1\,\mathrm{mRNA1}(t) - k_{deg}\,\mathrm{mRNA1}(t) + \mathrm{ksyn}\\
\frac{\partial \mathrm{mRNA2}(t)}{\partial t} &= -\left(k_{on2}\mathrm{IF}(t)\mathrm{mRNA2}(t) - \mathrm{mRNA2IF}(t)k_{ini2}\left(\frac{\mathrm{tRNA}}{K_m+\mathrm{tRNA}} + k_{off2}\right)\right) + k_1\,\mathrm{mRNA1}(t) - k_2\mathrm{mRNA2}(t) - k_{deg}\,\mathrm{mRNA2}(t)\\
\frac{\partial \mathrm{mRNA3}(t)}{\partial t} &= -\left(k_{on2}\mathrm{IF}(t)\mathrm{mRNA3}(t) - \mathrm{mRNA3IF}(t)k_{ini2}\left(\frac{\mathrm{tRNA}}{K_m+\mathrm{tRNA}} + k_{off2}\right)\right) + k_2\,\mathrm{mRNA2}(t) - k_{deg}\,\mathrm{mRNA3}(t)\\
\frac{\partial \mathrm{mRNA1IF}(t)}{\partial t} &= k_{on1}\mathrm{IF}(t)\mathrm{mRNA1}(t) - \mathrm{mRNA1IF}(t)\left(k_1 + k_{ini1}\frac{\mathrm{tRNA}}{K_m+\mathrm{tRNA}} + k_{off1}\right)\\
\frac{\partial \mathrm{mRNA2IF}(t)}{\partial t} &= k_{on2}\mathrm{IF}(t)\mathrm{mRNA2}(t) + k_1\,\mathrm{mRNA1IF}(t) - k_2\mathrm{mRNA2IF}(t) - \mathrm{mRNA2IF}(t)\left(k_{ini2}\frac{\mathrm{tRNA}}{K_m+\mathrm{tRNA}} + k_{off2}\right)\\
\frac{\partial \mathrm{mRNA3IF}(t)}{\partial t} &= k_{on2}\mathrm{IF}(t)\mathrm{mRNA3}(t) + k_2\mathrm{mRNA2IF}(t) - \mathrm{mRNA3IF}(t)\left(k_{ini2}\frac{\mathrm{tRNA}}{K_m+\mathrm{tRNA}} + k_{off2}\right)\\
\frac{\partial \mathrm{IF}(t)}{\partial t} &= \mathrm{IF}(t)(-(k_{on1}\,\mathrm{mRNA1}(t) + k_{on2}\mathrm{mRNA2}(t) + k_{on2}\mathrm{mRNA3}(t))) + \mathrm{mRNA1IF}(t)\left(k_{ini1}\frac{\mathrm{tRNA}}{K_m+\mathrm{tRNA}} + k_{off1}\right)\\
&\quad + \mathrm{mRNA2IF}(t)\left(\frac{k_{ini2}\mathrm{tRNA}}{K_m+\mathrm{tRNA}} + k_{off2}\right) + \mathrm{mRNA3IF}(t)\left(\frac{k_{ini2}\mathrm{tRNA}}{K_m+\mathrm{tRNA}} + k_{off2}\right)\\
\frac{\partial \mathrm{elo}(t)}{\partial t} &= \left(k_{ini1}\underbrace{(1-k_{PTH}\alpha(t))}_{\substack{k_{PTH}=0\text{ for reactions without PTH2,}\\ \text{otherwise this term accounts}\\ \text{for inactivation of translation machinery}}}\mathrm{mRNA1IF}(t)\frac{\mathrm{tRNA}}{K_m+\mathrm{tRNA}} + k_{ini2}(1-k_{PTH}\alpha(t))\frac{\mathrm{tRNAmRNA2IF}(t)}{K_m+\mathrm{tRNA}}\right) - \mathrm{kelo}\frac{\mathrm{tRNA}}{K_m+\mathrm{tRNA}}\mathrm{elo}(t)\\
\frac{\partial \mathrm{YFPx}(t)}{\partial t} &= \mathrm{kelo}\times\mathrm{elo}(t)\frac{\mathrm{tRNA}}{K_m+\mathrm{tRNA}} - \mathrm{kmat}\times\mathrm{YFPx}(t)\\
\frac{\partial \mathrm{YFP}(t)}{\partial t} &= \mathrm{kmat}\times\mathrm{YFPx}(t)
\end{aligned}
$$

The model was solved numerically using the built-in *NDSolve* function and fitted to the data manually using the *Manipulate* environment. The corresponding parameter values are reported in Table 2 and the script can be found as a supplementary file.

RESULTS

The different commercial PURE systems exhibit quantitatively different gene expression kinetics

We monitored the dynamics of gene expression at the mRNA and protein levels in real-time for various concentrations of DNA template with the three different PURE system versions: PURExpress, PURE*frex* and PURE*frex*2.0. As can be seen in Fig. 1, the profiles of mRNA and protein synthesis are qualitatively similar. The rate of mRNA production is constant for at least 2 h, with a maximum of 0.7 nM s^{-1}, and decreases over time to reach up to 2.5 μM transcript at high DNA concentrations. With PURExpress, but not with either of the PURE*frex* versions, the mRNA fluorescence signal decreases instead of plateauing, suggesting nuclease activity (as already reported in (50)). In all expression systems, the fluorescence time traces of YFP can be decomposed in three main phases in analogy with logistic growth curves: a lag time of a few minutes, a linear regime corresponding to constant translation rate and a plateau. The lag time can be attributed to chromophore maturation (about 20 min (39, 41)) and possibly to charging of tRNAs by their cognate amino acids. Even though all curves can be described by this general scheme, there are significant quantitative variations between the different conditions tested, which we parameterized by phenomenological modelling using three relevant features: the maximum translation rate (v_{max}), the lifespan representing the production period and the yield of synthesized active protein (Fig. 1(c)). Note that all three parameters merely describe *apparent* properties of protein synthesis dynamics as inferred from the fluorescence signal (Fig. S1, S2).

While DNA concentration has little effect on lifespan in PURExpress and in PURE*frex*2.0, lowering the amount of DNA in PURE*frex* is accompanied by longer expression periods, beyond 400 min in some reactions. The higher concentrations of fluorescent protein are about 1.5 μM, 5 μM and 6 μM for PURE*frex*, PURE*frex*2.0 and PURExpress, respectively. The maximum translation rate is reached with PURExpress at high DNA concentration and is about 45 nM min^{-1}. The aggregated data from all three PURE systems show that a wide range of parameter values can be reached (Fig. 1(d)). Moreover, there does not exist a single condition with respect to the PURE system version and DNA concentration that maximizes all three parameters. This result demonstrates that a trade-off in the system's performance has to be found for specific applications. Furthermore, there is no correlation between the protein production lifespan and the yield or maximum translation rate, whereas the yield and maximum translation rate are positively correlated for all three PURE systems.

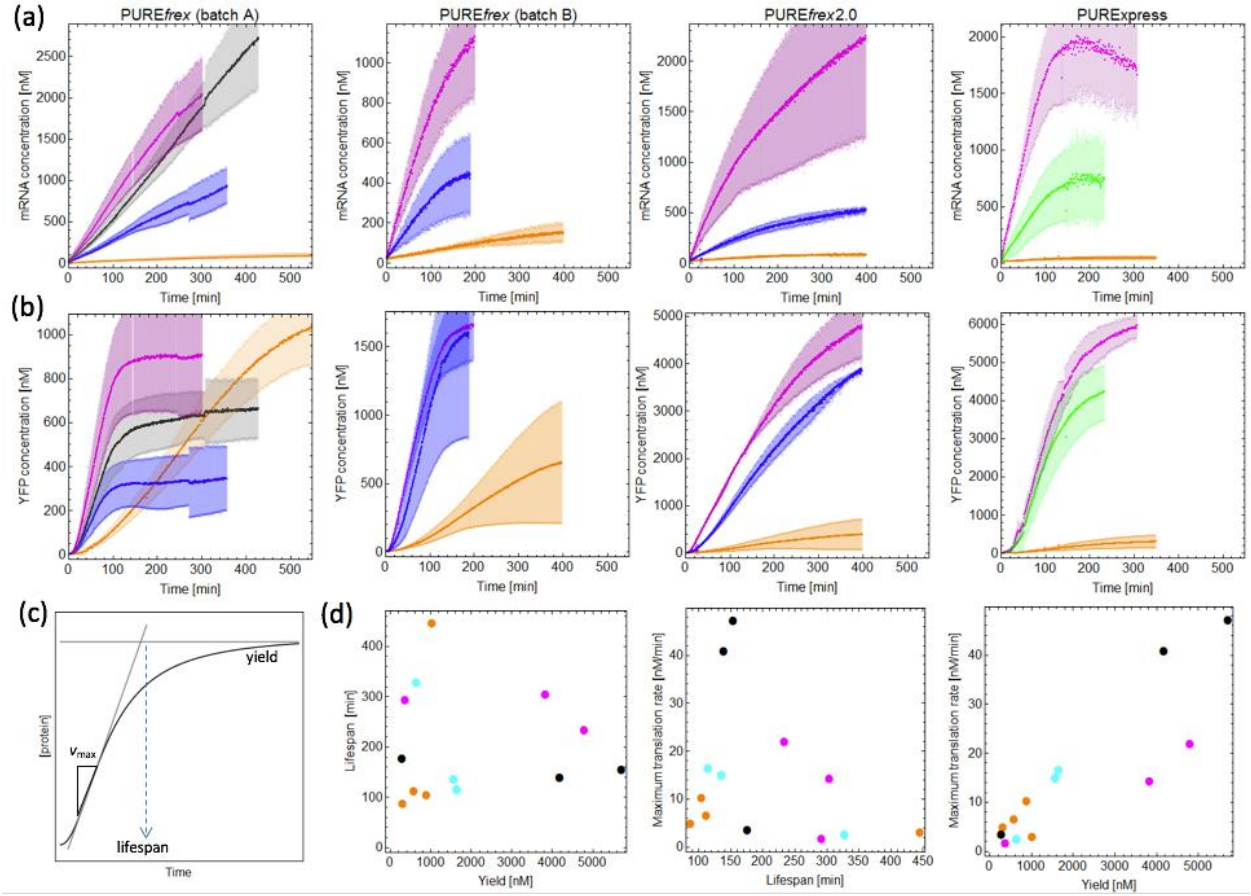

Figure 1: Kinetic profiles of mRNA and protein production with different PURE systems. The YFP fluorescence trajectories are not corrected for maturation time. (a) Time traces of mRNA concentration (calculated from DFHBI/spinach fluorescence) for different DNA concentrations: 7.4 nM (magenta), 3.2 nM (black), 1.5 nM (blue), 0.74 nM (green) and 0.074 nM (orange). Bold lines indicate averages of at least three experiments and shaded areas the average +/– 1 standard deviation (SD). (b) Time traces of YFP concentration for the same reactions and same color-coding as in (a); values are calculated from

measured fluorescence and are not corrected for maturation time. (c) Schematic explanation of the parameters derived from the YFP fluorescence kinetics. (d) Comparison of the protein kinetics parameters for the different PURE systems: PURExpress (black), PURE*frex*2.0 (magenta), PURE*frex* batch A (orange) and PURE*frex* batch B (cyan). Values are derived from the curves shown in (b) as described in (c).

Besides notable technical variability (reactions from same batches of DNA and PURE system), a large batch-to-batch variability was observed with PURE*frex* (Fig. 1 and Fig. S3). In contrast, different batches of PURExpress received over seven years led to more consistent expression kinetics and protein yield. The reported experiments with PURE*frex*2.0 come from a single batch. Experiments from a second batch with 7.4 nM DNA showed a higher yield, approximately 10 μM, while the expression time remained similar (data not shown).

Some of the differences in yield between PURExpress and PURE*frex* can be explained by a higher total protein concentration in PURExpress. Indeed, a coomassie-stained gel of both systems shows that PURExpress contains a significantly higher concentration of ribosomes as well as EF-Tu (Fig. S4(a)). To test if either of the two systems contains additional protein factors that are not listed in the original PURE system composition (51), we performed an LC-MS proteomic analysis. In both systems, we could detect a number of additional proteins, though less in PURE*frex* (the full list can be found in the Table S1 as a separate file). For example, the trigger factor chaperone protein, two hydrogen peroxide removal proteins and ribosome-binding factor A were uniquely identified in PURExpress. They most likely originate from carryover of purification.

Simple kinetic models do not capture the apparent dynamics for PURE*frex* and PURExpress

We next sought to challenge the coarse-grained deterministic models available in the literature with our kinetics data of transcription and translation activity. We noticed that all models have in common that translation rate (v_{TL}) can be described as a function of the concentrations of four species: the active translation machinery, mRNA, aminoacyl-tRNAs (aatRNAs) and NTPs:

$$v_{\mathrm{TL}}(t) = k_{\mathrm{TL}} \times f\big([\text{translation machinery}](t)\big) \times f\big([\mathrm{mRNA}](t)\big) \times f\big([\mathrm{aatRNAs}](t)\big) \times f\big([\mathrm{NTPs}](t)\big), \quad \text{(Eq. 1)}$$

where k_{TL} is the translation rate constant and t is time. Functional dependencies on the latter three concentrations are described with Michaelis-Menten kinetics or approximated as linear relations. The concentration of active translation machinery is modelled either as a decreasing exponential function or with enzymatic decay. Some of the models include additional constants (e.g. to account for conversion between consumed amino acids and amount of full-length protein (38)), or incorporated nucleotides and concentration of mRNA, but the equations could be rewritten by a rescaling of parameters into the one above (Eq. 1). Not all models include all four

functional dependencies, while some account for the concentration of amino acids instead of aatRNAs (37). To test if these models – in their general form – can fit the gene expression kinetics observed in the three different PURE systems, the translation rate (measured by the increase in YFP fluorescence over time) was plotted as a function of mRNA concentration (Fig. 2(a)). The shape of these experimental curves can be compared to simulated data obtained with different parameterizations of Eq. 1 (Fig. 2(b)). Although data from different DNA concentrations can individually be fit with both PURE*frex* (Fig. S5) and PURExpress, no global fitting across the entire range of DNA concentrations can be obtained. PURE*frex* reactions exhibit different maximum translation rates (corresponding to the linear regime in time traces) for different DNA concentrations, but during that period the amount of mRNA also increases. This is seen as plateaus of different heights in Fig. 2(a), a feature that cannot be achieved using Eq. 1 (see simulated model predictions in Fig. 2(b)). The relationships between translation rate and mRNA concentration with PURExpress are also difficult to interpret with existing models. Indeed, the kinetics with intermediate DNA concentration reaches the maximum translation rate at a lower mRNA concentration than reactions with high DNA concentration, whereas in both conditions the maximum translation rate is reached after about 60 min. Therefore, the magnitude of inactivation of the translation machinery should be comparable at that time. Even taking into account higher resource consumption in the reaction with higher DNA concentration (faster consumption of NTPs and amino acids), these profiles cannot be fit globally with Eq. 1. With PURE*frex*2.0, a reasonably good qualitative description is obtained.

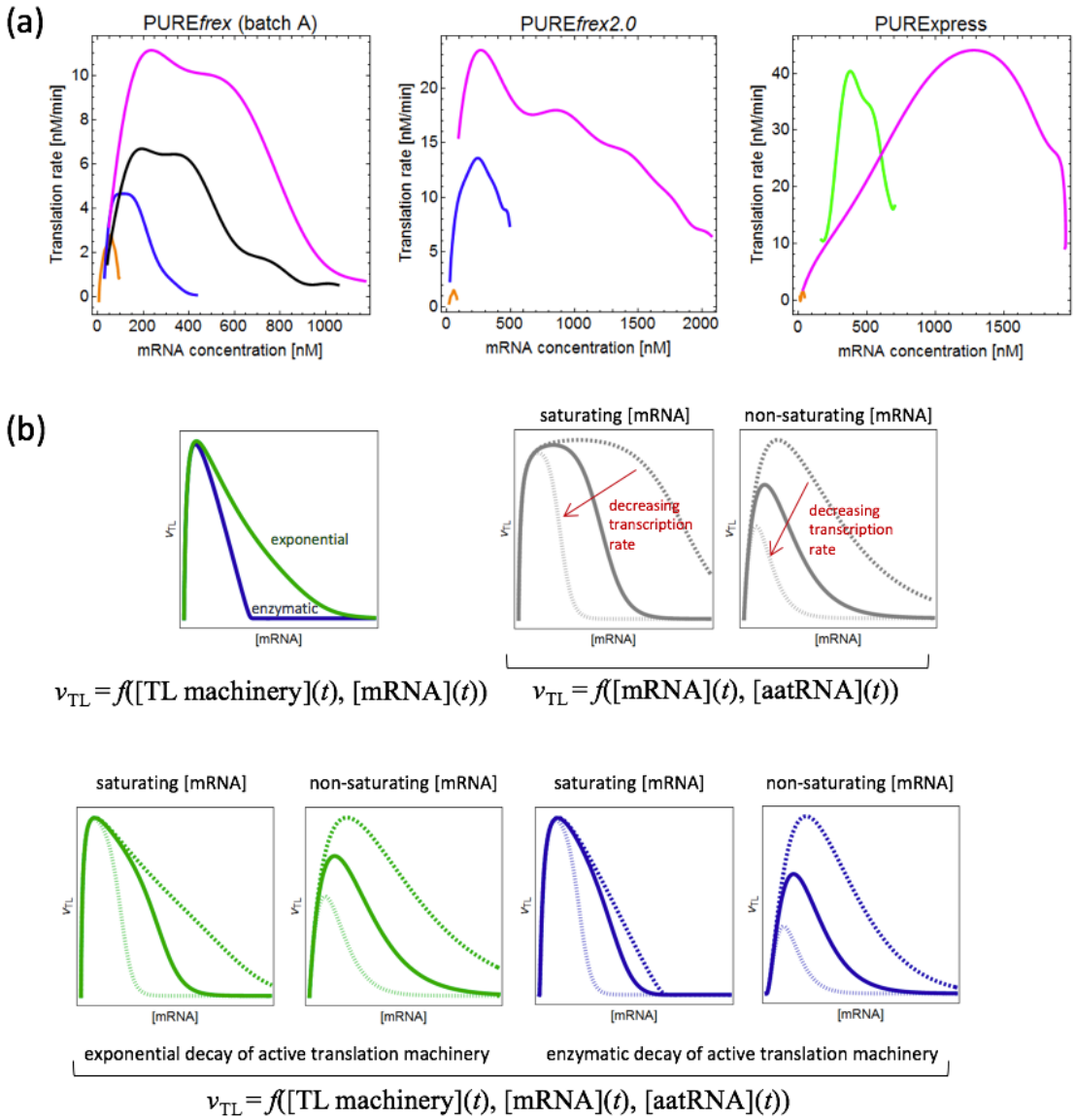

Figure 2: Comparison of generalized model predictions for the expression kinetics with the different PURE systems. (a) Translation rate as a function of mRNA concentration for the three different PURE systems and different DNA concentrations: 7.4 nM (magenta), 3.2 nM (black), 1.5 nM (blue), 0.74 nM (green) and 0.074 nM (orange). (b) Model predictions for translation rate as a function of mRNA concentration for different combinations of control parameters. For all plots mRNA concentration was assumed to increase linearly with time. In the upper left panel a first order ('exponential') or Michaelis-Menten ('enzymatic') type of inactivation of the translation machinery are considered. In the upper middle and right panels two conditions were simulated: [mRNA] lower ('non-saturating') or higher ('saturating') than the concentration above which the translation rate is constant (linear regime). Simulated curves corresponding to three different transcription rates are plotted: dashed, solid and dotted lines, from high to low rate values.

There are two possible explanations why the models do not fully capture the observed IVTT dynamics: (i) The current models lack integration of other 'hidden' (non-apparent) processes, such as DNA inactivation, synthesis of unwanted by-products (e.g. truncated proteins) or depletion of translation factors into inactive complexes. (ii) The model is actually good but the observables do not reliably report on the system's behaviour. Indeed, YFP fluorescence only informs about the concentration of active, mature, protein. Although the plots in Fig. 2 are corrected for the maturation time, a scenario whereby a subset of synthesized proteins may be misfolded or truncated, hence not fluorescent, is not taken into account. If the fraction of non-fluorescent translation products is constant over time and across different reactions, the conclusion that Eq. 1 does not fit the data would still hold. However, we cannot exclude that this fraction may vary.

Most of the experiments described hereafter, as well as the modelling attempts, have been conducted using PURE*frex* (unless specified otherwise). Despite a more pronounced batch-to-batch variability with PURE*frex*, this choice was motivated by the fact that nucleic acids in PURExpress are more prone to degradation due to nuclease contamination. Furthermore, we know the exact composition of PURE*frex* only, which facilitates data interpretation and model construction.

Inactivation of the translation machinery, and depletion of either NTPs or amino acids do not limit translation rate in PURE*frex*

To discriminate the two scenarios mentioned above, we studied the translation rate dependence on the four individual concentration terms of Eq. 1. First, inactivation of the translation machinery was assayed by adding the DNA template after pre-incubating PURE system reactions for 4 h at 37 °C. A slower translation rate after delayed introduction of DNA would indicate inactivation of one of the rate-limiting components of the system over time. With PURE*frex*, no major difference in the YFP kinetics between delayed and non-delayed gene expression was observed (Fig. 3). Therefore, we can rule out a contribution of this term in Eq. 1 on translation rate. Moreover, addition of fresh enzyme or ribosome solution (from the corresponding vials of the kit) at the YFP fluorescence plateau does not restart translation (data not shown). Unlike with PURE*frex*, delayed addition of DNA with PURExpress results in a decrease both in transcription and translation rates (consistent with previous reports (35, 52)), whilst with PURE*frex*2.0 a reduced translation rate was measured (Fig. 3). These results suggest that transcription-translation-independent inactivation of the translation machinery reduces protein synthesis rate in these two systems, but not in PURE*frex*.

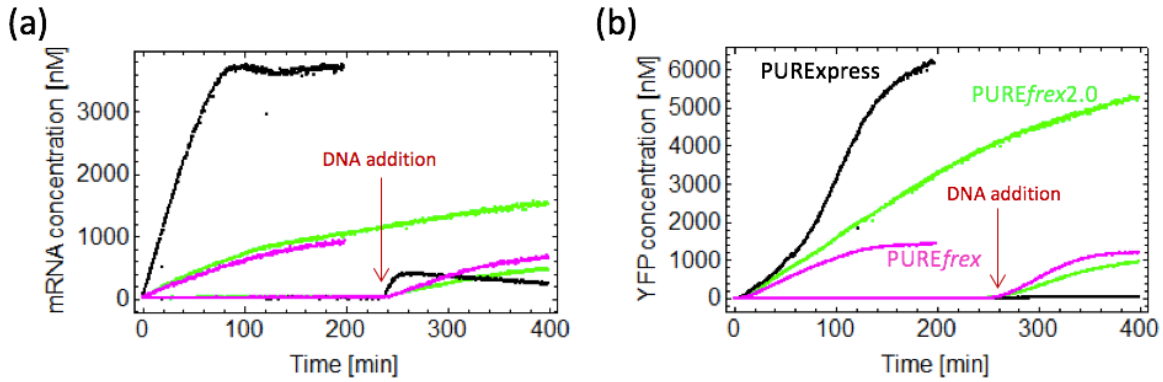

Figure 3: Inactivation of the translation machinery over time differs for the three PURE systems. (a) mRNA concentration as a function of time for reactions that were either started immediately or pre-incubated at 37 °C for 4 h before addition of DNA to a concentration of 7.4 nM. PURExpress (black), PURE*frex*2.0 (green) and PURE*frex* batch B (magenta). (b) Time traces of YFP concentration for the reactions shown in (a). Color-coding is the same as in (a).

Mass spectrometry analysis of the PURE*frex* reaction mix at different time points indicates that neither NTPs nor amino acids are depleted significantly enough to explain the observed translation rate profiles (Fig. S6). GTP and ATP are consumed, but their concentrations after 3 h are still two orders of magnitude higher than typical K_M values of aaRS (ATP) (www.brenda-enzymes.org) and equilibrium dissociation constants of GTP and elongation factors (53, 54). Build-up of ADP and GDP was detectable but seems too low to quantitatively explain the reduction in translation rate over time (about 20-fold excess of ATP, GTP over ADP, GDP after 3 h). NMP concentrations were below the detection limit. This result is consistent with previous measurements on the original PURE system (26). Amino acid depletion in PURE*frex* is negligible (Fig. S6(c)). This is not unexpected as the maximum amount of synthesized fluorescent protein lies only between 1 and 2 µM (depending on the PURE*frex* batch); for the most frequent amino acid glycine (with 28 residues per protein), this means that between 5 and 10% of this amino acid are consumed for the production of full-length mature YFP (all amino acids are present at equimolar amounts at the start of the reaction). Inhibition of translation through accumulation of toxic by-products has been proposed in IVTT reactions with cell lysates (55-57). To test this hypothesis, we performed PURE*frex* reactions in a dialysis chamber with a membrane cut-off of 20 kDa and measured fluorescence over time (Fig. S7). The external dialysis solution acting as feeding reservoir contained all components of PURE*frex* solution I except tRNAs. Kinetics of reactions inside the dialysis chamber were similar as those monitored in batch format. This finding rules out inhibitory effects of side products with low molecular mass (unless they precipitate and remain in the reaction chamber) and corroborates that nutrient depletion is not rate-limiting in PURE*frex* reactions.

Depletion of tRNAs and decrease of the translational ability of mRNA affect translation rate

Because the translation rate depends on the concentration of charged aatRNAs, and not directly on amino acid concentrations, one can envisage two possible bottlenecks beside amino acid depletion: tRNAs themselves are depleted or the aminoacylation step becomes rate limiting. Remarkably, addition of fresh tRNAs at plateau restarts translation for reactions with high or intermediate DNA concentrations (Fig. 4(a)), as well as for reactions with high or intermediate concentrations of purified mRNA used as template to bypass transcription (Fig. 4(b)), but not with low concentration of DNA (Fig. 4(a)). As discussed above, the sole inactivation of the translation machinery (including tRNAs as well as aaRSs) over time as a common bottleneck for all initial conditions can be ruled out. The data show that tRNAs are depleted faster with higher mRNA concentrations (Fig. 4(b,c)), which is accompanied by a lower yield of synthesized protein (Fig. 4(d)). Moreover, tRNA depletion is mediated by the presence of a ribosome binding site in the mRNA sequence (Fig. 4(b)). These results suggest that the translation process itself participates in tRNA depletion.

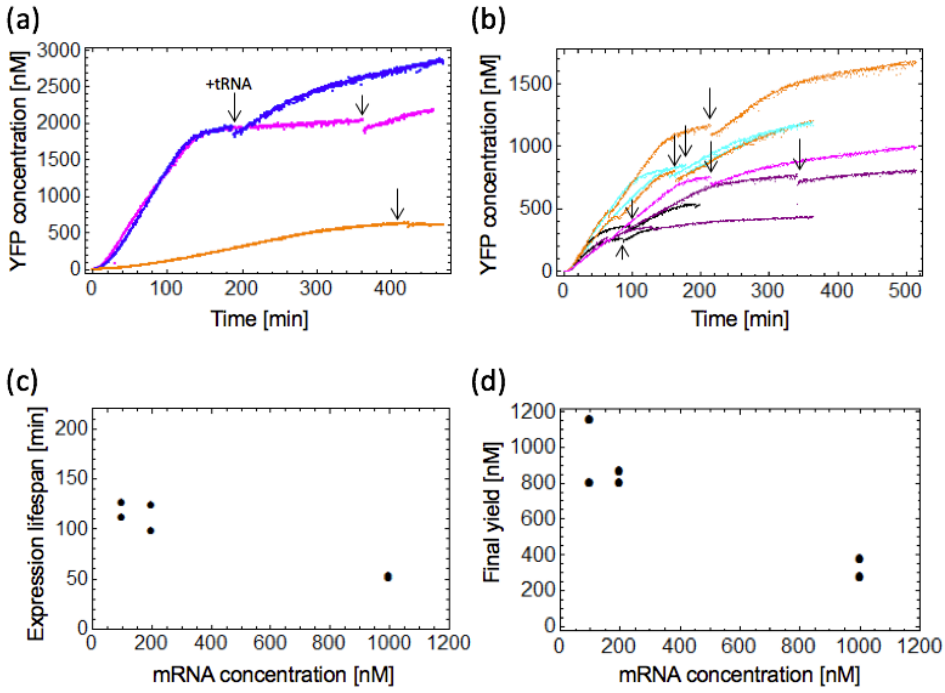

Figure 4: Increasing lifetime or translation rate in PURE*frex* (batch B) reactions with tRNA. (a) Addition of tRNA (1 µl at 0.6 mM, indicated by arrows) to PURE*frex* reactions with different DNA concentrations: 7.4 nM (magenta), 1.5 nM (blue) and 0.074 nM (orange). (b) PURE*frex* reactions starting with purified mRNA at different concentrations: 1 µM (black), 200 nM (cyan), 100 nM (orange), 100 nM mRNA + 1 µM mRNA without RBS (magenta), and 50 nM (purple). At plateau, tRNA (1 µl at 0.6 mM) was added to the reaction (indicated by an arrow). Traces of the same color are duplicates. (c,d) Expression lifespan (c) and final yield (d) as a function of mRNA concentration calculated from the reactions shown in (b) where tRNA is limiting, i.e. excluding the condition with 50 nM mRNA.

To further analyse the dependence of translation rate on mRNA concentration, purified mRNA was used as a template at different concentrations to bypass transcription. The maximum translation rate was reached at ~100 nM mRNA indicating saturation of the translation machinery (Fig. 4). Interestingly, the maximum translation rate is a factor of two lower than that for reactions starting with DNA template (Fig. 1). This result suggests that some form of transcription-translation coupling enhances translation rate. Alternatively, the folding state of mRNA, hence its propensity to efficiently serve as a template for translation, may differ between the two types of reactions. To test this hypothesis, experiments were performed, in which transcription was stopped after 30 min by DNase addition (Fig. 5). At high and medium concentration of DNA, premature arrest of transcription is accompanied by an increased lifespan of translation, up to >9 h with 7.4 nM DNA (Fig. 5(c)). However, translation rate is lower compared to that in the linear regime with unsuppressed transcription. At high DNA concentration, a higher end-point yield of YFP was reached, ~3 µM compared to ~2 µM (Fig. 5(c)). Given that 400 nM mRNA are produced, along with >2 µM YFP, the corresponding protein-to-mRNA ratio of about 5 is significantly higher compared to values <2 at 7.4 nM DNA without DNase addition (Fig. 1). For all DNA concentrations tested the translation rate decreases relatively quickly after DNase addition by about 2-fold (some of the delay is due to YFP maturation). This reduced rate is comparable to the translation rate observed in reactions triggered with purified mRNA template. This is consistent with the idea of a positive effect of transcription-translation coupling on translation rate and could in part explain the different dependence of translation rate on mRNA concentration for the different DNA concentrations: reactions with higher DNA concentrations have a higher transcription rate and therefore a higher concentration of nascent mRNA. If this nascent mRNA acts as a better template for translation, reactions with higher transcription rate would also have a higher translation rate. Moreover, translation rate after suppressed transcription is not constant (as also observed in the reactions with purified mRNA) but continuously decreases, which suggests further mRNA inactivation or other inhibitory effects. Like reactions with low DNA concentrations, reactions with DNase cannot be restarted by spiking more tRNAs at the plateau phase (data not shown), suggesting a similar cause of translation cessation.

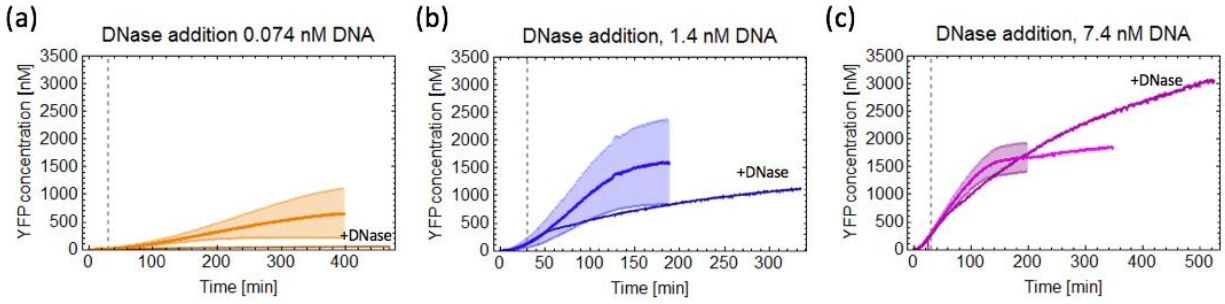

Figure 5: YFP expression kinetics after DNase addition in PURE*frex*. Average values +/– 1 SD (from three experiments) for control reactions without DNase addition are shown for comparison. 1 µl of RQ1

RNase-free DNase (1 U/µl) was added at t = 30 min (vertical dashed line), which was sufficient to stop new transcription immediately (corresponding spinach fluorescence kinetics are not shown). PURE*frex* batch B was used.

As an alternative method to test the effect of transcription-translation coupling on translation rate, reactions with delayed ribosome addition were performed (58) under conditions where the amount of tRNA is not limiting (Fig. 6). Transcription could proceed for 60 or 100 min before translation was triggered by adding ribosomes. Interestingly, translation rate was lower than in control reactions, where IVTT occurs from the beginning, even though translation started with a higher concentration of mRNA (Fig. 6). The extent of the decrease in translation rate varies between different batches of PURE*frex*. To rule out an inhibitory effect of the transcription process itself on translation, a control reaction was performed, where a DNA construct encoding an mRNA without functional RBS was expressed for 100 min without ribosomes; then ribosomes along with the regular DNA template were added (Fig. 6(a,c)). Although the translation rate is lower as compared to a control reaction containing ribosomes from the start, this decrease is much less than in a reaction where the RBS-containing mRNA is expressed before ribosome addition (Fig. 6(a)).

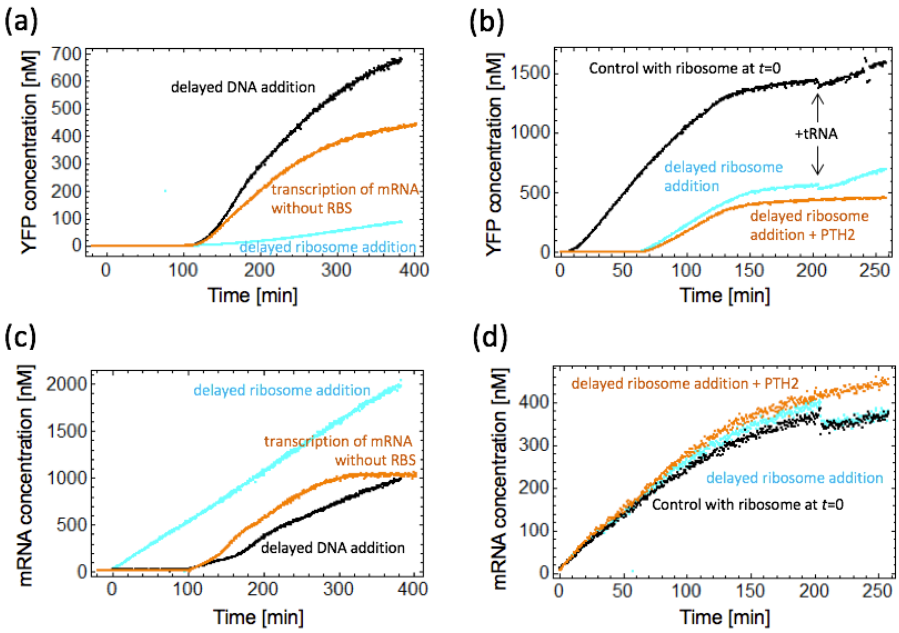

Figure 6: Comparison of mRNA and YFP expression kinetics after delayed ribosome addition in PURE*frex*. (a,c) PURE*frex* batch A. Ribosomes (cyan, orange) or DNA (black) were added after the reaction had been incubated for 100 min at 37 °C. All reactions contained 3.9 nM DNA and 100 nM PTH2 to circumvent tRNA limitation (see Fig. 8 for details). The reaction depicted in orange contained DNA encoding an mRNA without functional RBS from the beginning. After 100 min, at the same time as ribosome addition, the *meYFPco-LL-spinach* DNA construct was added. (b,d) PURE*frex* batch B.

Ribosome addition after 60 min (cyan, orange) and a control reaction (black) with ribosomes present from the start. A tRNA mix was added at plateau at the indicated times.

Collectively these observations are consistent with a scenario in which the mRNA generated in the absence of ribosomes (more generally, the full-length synthesized mRNA released from the RNA polymerase in coupled IVTT) is less prone to be successfully translated, while competing with fresh transcripts for translational resources.

Polysome analysis indicates inefficient ribosome usage during PURE*frex* and PURExpress reactions

Translation rate can be expressed as a function of the number of actively translating ribosomes and of elongation rate. The number of translating ribosomes is itself a function of initiation rate. To examine the efficiency of ribosome usage during PURE system reactions, we analysed the distribution of polysome sizes by atomic force microscopy (59). Polysome profiles were monitored at different time points during PURE*frex* and PURExpress reactions. Ribosomes translating the same mRNA molecule are seen as clusters of particles of about 14-nm height. We were able to resolve 70S, 50S, and 30S particles (Fig. 7(b)), but could not distinguish single translating 70S ribosomes from free ones. Most likely, under the conditions used for attaching the molecules to the mica surface the majority of ribosome-bound mRNA molecules assumes a highly folded conformation and they cannot be resolved from the bound ribosome particle(s) (although we could observe free mRNA molecules in extended conformations with a height of approximately 0.5 nm). Large polysome sizes were measured during PURExpress reactions, whereas only disomes could be detected in PURE*frex* reactions. Combining with the apparent translation rate obtained from the mRNA and YFP fluorescence kinetics (at time points between 15 and 30 min), the polysome distributions can be used to derive upper and lower bounds for the average translation time (comprising the initiation, elongation and termination steps): $v_{\mathrm{TL}}(t)$ = [translating ribosomes]/translation time. The two bounds are obtained by either taking the concentration of ribosomes engaged in polysomes as the total concentration of translating ribosomes, or by assuming that in addition all mRNA molecules that are not observed as polysomes are occupied by a single translating ribosome. Because not all translating ribosomes may generate full-length proteins (see below), the translation time might be underestimated. For PURExpress the calculated bounds on translation time are 17 and 27 min. From the delay time observed in the YFP fluorescence kinetics it is clear that the translation time must be closer to the lower bound. It is a surprisingly long time for a 272-amino acid protein, equivalent to 0.27 amino acids per second per ribosome, which is comparable to the rate measured by Takahashi et al. (60) at 25 °C instead of 37 °C and about 5 times lower (1.5 amino acids per second per ribosome) than reported for a cell lysate (21).

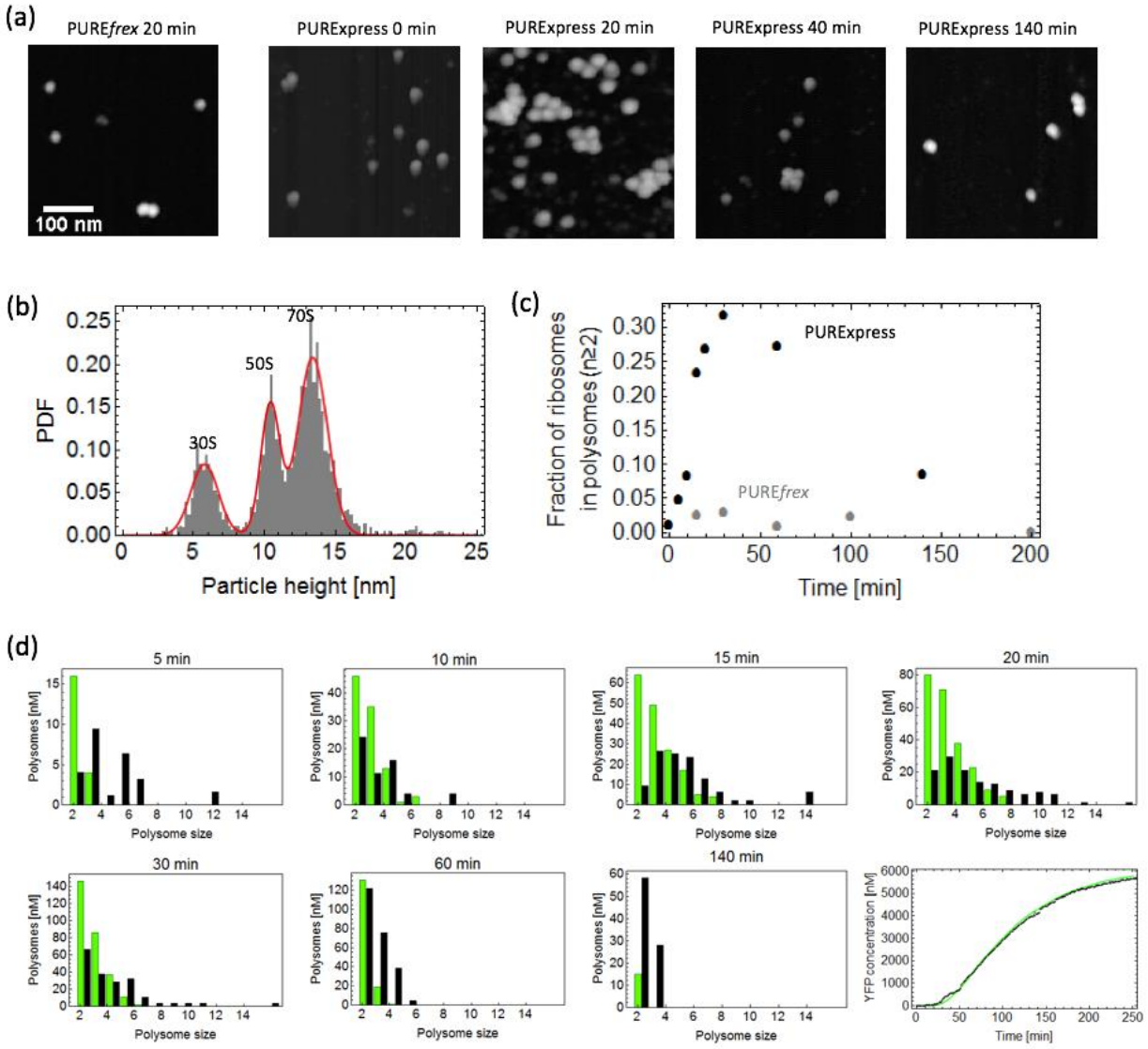

Figure 7: Polysome analysis of PURE*frex* and PURExpress reactions by AFM. (a) Representative AFM images at different time points, with the exception of the PURExpress 20 min condition for which an area with multiple polysomes was selected. All images have the same size. Reactions were performed with 7.4 nM *meYFPco-LL-spinach* DNA. (b) Example for a particle height distribution derived from the AFM images; in red the fit with three Gaussian distributions. (c) Fraction of ribosomes in polysomes with PURExpress (black) and PURE*frex* (grey) at different time points. (d) Polysome distributions at different time points for a PURExpress reaction (black) and simulation results (green). Polysome size refers to the number of ribosomes per polysome. The last panel displays the experimental (black) and simulated (green) time traces of YFP production.

To test if a simple mechanism could explain the observed polysome size distribution in PURExpress a stochastic model of translation was implemented (Table 1). As can be seen in Fig. 7(d), the simulated fraction of ribosomes in polysomes matches the observed AFM data

reasonably well. However, the model does not reproduce the very large polysome sizes of up to 12 ribosomes observed in the experiments. In PURE*frex* reactions, the translation time was estimated to be between 1 and 7 min. From this analysis, one can deduce that at the maximum translation rate only ~10% of ribosomes are actively translating in PURE*frex* (compared to 80% *in vivo* and 70% in an *E. coli* extract (21). This result suggests a severe bottleneck in translation initiation. Indeed, a bottleneck during early translation elongation would also lead to the formation of few polysomes because the firstly engaged ribosome would block the RBS. However, if initiation was otherwise efficient, higher translation rates would be achieved by addition of more mRNA, which is not supported by the experimental data.

Reaction type	**Reaction rate**	**Parameter values**
mRNA synthesis	$v_{tr} = k_{syn}(t)$	$k_{syn}(t)$ is obtained as derivative of the fitted mRNA concentration over time
Ribosome binding to RBS	$v_{ini} = k_{ini} \times [RBS] \times [ribosome_{free}]$	$k_{ini} = 0.0002\ min^{-1}\ nM^{-1}$
Ribosome inactivation	$v_{deg} = k_{deg} \times [ribosome_{total}]$	$k_{deg} = 0.014\ min^{-1}$
YFP maturation	$v_{mat} = k_{mat} \times [YFP_{non\text{-}fluorescent}]$	$k_{mat} = 0.0462\ min^{-1}$
Ribosome translocation	$v_{elo} = k_{elo} \times [ribosome_{translating}]$	$k_{elo} = 20\ min^{-1}$ (per codon)

Table 1: Reaction scheme and parameters for the stochastic model of translation.

PURE system reactions generate heterogeneous translation products

As the tested models cannot globally reproduce the observed fluorescence kinetics (Fig. 2) it seems likely that either the fluorescence measurements do not reliably report on mRNA and protein production or that additional processes need to be included in the model. To directly visualize all translation products, including non-fluorescent ones, fluorescently labelled lysine residues were co-translationally incorporated during IVTT and the newly synthesized polypeptides/proteins (with the exception of very short peptides) were imaged by gel electrophoresis (Fig. S4(b)). Besides the full-length protein, all three PURE systems generate side products within 40 min. The lower bands most likely correspond to an incomplete translation product. A similar result was observed when expressing constructs coding for different proteins, the number of short fragments being template-dependent (not shown). It appears that specific shunt proteins are formed, as opposed to random truncation events that would lead to a smear. No obvious correlation with possible difficult-to-translate sequences, e.g. proline-rich motifs (PVPWP and PIGDGPVLLP in the YFP sequence) (61), or with the presence of internal ribosome binding sites (62), could be found. Moreover, C-terminal truncation due to stalled ribosomes on degraded mRNA seems unlikely since the level of nuclease activity in both

PURE*frex* versions is very low. Interestingly, one (in PURExpress) or two (in PURE*frex* and PURE*frex*2.0) higher bands indicate formation of products longer than YFP. This could happen by stop codon read-through (63) or frameshifting over the stop codon (64). Analysing the DNA template sequence, read-through until the next stop codon would lead to an additional mass of 1.8 kDa, while frameshifting would append 0.8 kDa. The corresponding longer translation products cannot however explain the observed mass difference with the upper band. The mature and immature full-length YFP could conceivably run differently on gel. Furthermore, mass spectrometry was employed to detect targeted translation products and a frameshift peptide was measured under limiting amino acid concentrations (Supplementary Methods and Fig. S8). In conclusion, all three PURE systems generate unintended translation products, which must be accounted for to extract actual translation rates. As the fraction of side products seems to remain constant throughout the expression lifespan, the conclusion that the tested models do not fit the expression kinetics remains valid.

tRNAs depleted through peptidyl-tRNA drop-off can be recycled by adding a peptidyl-tRNA-hydrolase

One mechanism that could explain shortage of the tRNA pool observed in PURE*frex* reactions starting with intermediate or high DNA (Fig. 4(a)) or mRNA (Fig. 4(b)) concentrations is peptidyl-tRNA drop-off and thereby sequestration of tRNAs in the form of peptidyl-tRNAs (Fig. 8(a)). Drop-off occurs when a ribosome terminates translation prematurely without the action of a release factor. Peptidyl-tRNA drop-off has been observed *in vivo* (65-71) and *in vitro* (72-74) and, in *E. coli* mostly but not exclusively, occurs during the first few codons of a transcript. *In vivo* peptidyl-tRNAs are recycled by the essential enzyme peptidyl-tRNA-hydrolase (PTH). This enzyme has been used before in PURE system reactions as part of a screening (75), but its effect was not characterized in detail. Here, the human enzyme PTH2 was used in most experiments. For high and intermediate DNA concentrations (Fig. 8(c,d)), as well as for high concentrations of purified mRNA (Fig. 8(f)), reactions with 100 nM PTH2 have a longer lifespan and a slightly higher yield than control reactions without PTH2 (Fig. 8(e)), and cannot be restarted with fresh tRNA (data not shown), demonstrating that tRNA limitation can partly be relieved. In contrast, for reactions with low DNA concentrations, addition of PTH2 lowers the yield and shortens the linear regime of YFP fluorescence increase (Fig. 8(e)).

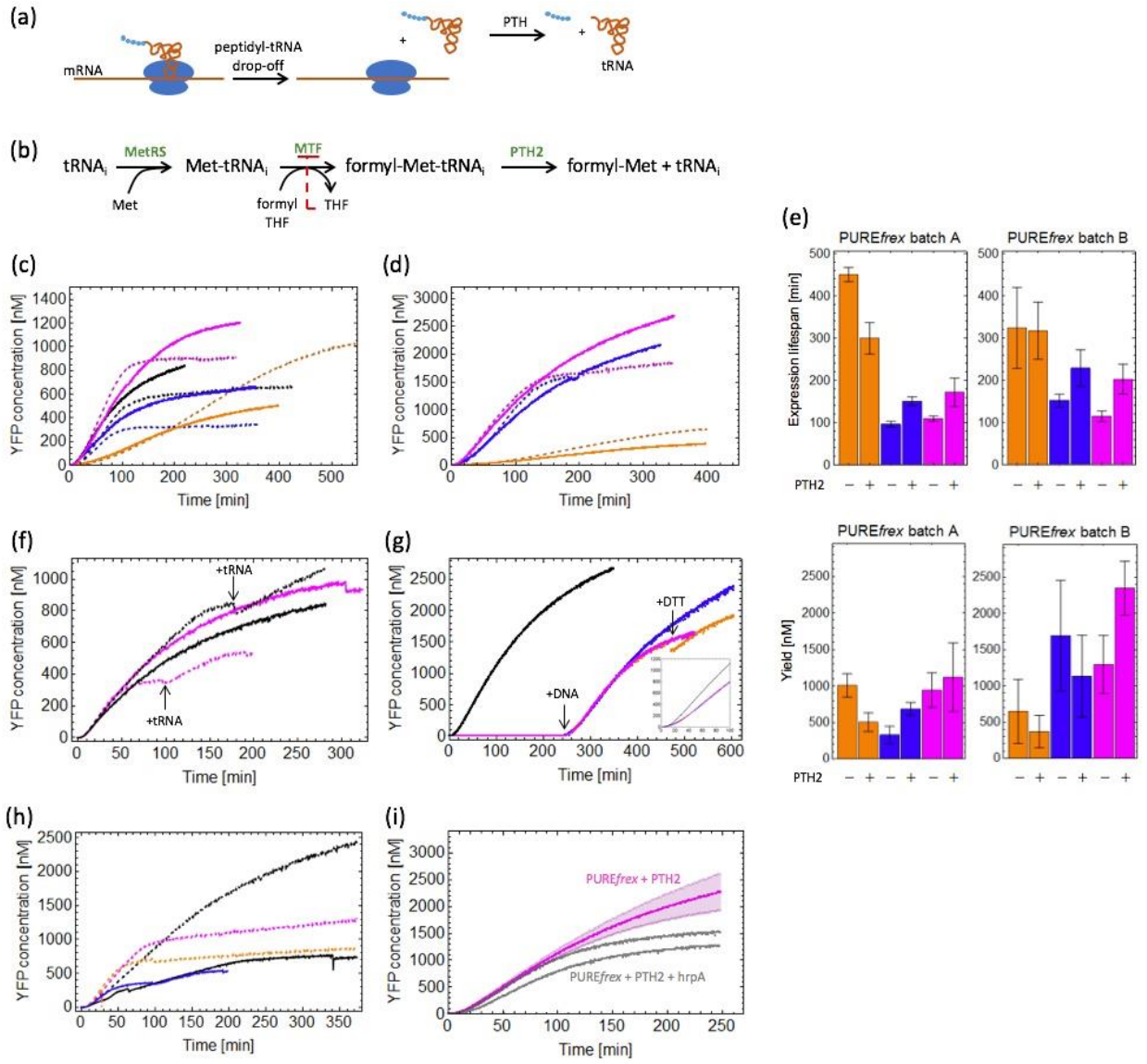

Figure 8: Effects of PTH2 on PURE*frex* reactions. (a) Schematic of peptidyl-tRNA drop-off and recycling of free tRNA by PTH. (b) Schematic of initiator tRNA charging and formylation with product inhibition by tetrahydrofolate (THF). (c) PURE*frex* batch A reactions at different DNA concentrations with (solid) or without (dashed) PTH2. Data are averages from at least three experiments. DNA concentrations were 7.4 nM (magenta), 3.2 nM (black), 1.5 nM (blue) and 0.074 nM (orange). (d) Same as in (c) but with PURE*frex* batch B. (e) Expression lifespan and yield of the kinetics depicted in (c) and (d). Average values +/– 1 SD are represented. Colour coding is similar as in (c) and (d). (f) PURE*frex* batch B reactions with purified mRNA template with (solid) or without (dashed) PTH2. Concentrations of mRNA: 200 nM (black) and 1 µM (magenta). (g) PURE*frex* batch B reactions supplemented with PTH2 with delayed addition of 7.4 nM DNA: control reaction (average of three reactions, black trace), DNA addition after 240 min (magenta), DNA and DTT (5 mM) addition after 240 min (blue), DNA addition after 240 min and DTT addition indicated by arrow (orange). The inset shows the time-shifted traces for easier visualization of the slower initial translation rate with delayed addition of both DNA and PTH2. (h) HrpA increases translation rate for reactions triggered with purified mRNA template. PURE*frex* reactions

with (dashed) or without (solid) 100 nM hrpA for different mRNA concentrations: 50 nM (black), 300 nM (magenta), 600 nM (orange) and 1 µM (blue). (i) Comparison of PURE*frex* reaction kinetics with 7.4 nM DNA and supplemented with 100 nM PTH2 with (gray) and without (magenta) 100 nM hrpA. The average values +/– 1 SD from three experiments are reported. The two grey traces are duplicates.

We first hypothesized that reactivity of the PTH2 with formyl-Met-tRNA could be the cause for the system's lower performance. While the *E. coli* peptidyl-tRNA-hydrolase does not recognize formyl-Met-tRNA as a substrate, it is known that the yeast enzyme (which is functionally close to the human enzyme we have used) hydrolyses formyl-Met-tRNA. Every hydrolysis event requires recharging of the initiator tRNA with methionine and subsequent formylation of Met-tRNA, which generates tetrahydrofolate as a second reaction product (Fig. 8(b)). It is known that the methionyl-tRNA-formyl-transferase (MTF) is subject to product inhibition by tetrahydrofolate (76). To bypass this mechanism, we purified the peptidyl-tRNA-hydrolase from *E. coli* and found the same effects as with PTH2 (Fig. S9). Hydrolysis of formyl-Met-tRNA as an inhibitory side reaction of PTH2 can therefore be ruled out. A second possible mechanism mediated by PTH2 is suggested by the observation that the linear regime of YFP fluorescence increase is shorter in reactions with PTH2 than without. This result could be explained by an inactivation of the translation machinery over time, which would only become rate-limiting in the presence of PTH2. This hypothesis was tested again by delayed DNA addition (Fig. 8(g)). Reactions were pre-incubated for 4 h either with or without PTH2. In the reaction without PTH2 the enzyme was added together with DNA. Both delayed reactions show a lower translation rate than the control which contains DNA and PTH2 from the beginning (Fig. 8(g), see inset for comparison of the initial translation rates). These results indicate that one or more components of PURE*frex* are inactivated over time, but that they are rate-limiting only in the reactions with PTH2 (see Discussion section).

Attempts to restore transcription-translation coupling failed to enhance efficiency of translation initiation

Having identified that translation initiation is inefficient in PURE*frex*, we next attempted to improve this process. We first hypothesized that translation initiation is hampered by mRNA secondary structures (58, 77-80) that would be more susceptible to form *in vitro* due to poor coupling between transcription and translation. In *E. coli*, the first ribosome translating an mRNA often closely follows the transcribing RNA polymerase (see Castro-Roa and Zenkin (81) for an *in vitro* study with the endogenous *E. coli* RNA polymerase). This prevents the polymerase from stalling but also ensures that ribosome binding to the RBS is not impeded by mRNA secondary structure. In the PURE system, this coupling is lost because the T7 RNA polymerase is faster than the native *E. coli* RNA polymerase and also much faster than a translating ribosome. Following strategies that improved transcription-translation coupling in an *E. coli* cell extract (82), we performed PURE*frex* reactions at lower temperature and we used a slow T7 RNA polymerase mutant P266L/I810S. Reactions at lower temperature exhibit longer

translation lifespan (>10 h at 21 °C) but the yield as well as the protein-to-mRNA ratio is about a factor two lower at 21 °C than at 37 °C (Fig. 9(a-e)). For experiments with the slow T7 RNA polymerase, the enzyme concentration was set to 30 nM for both wild-type and mutant, but the DNA concentration was carefully adjusted such that the initial transcription rate was the same for both reactions (Fig. 9(f,g)). No significant differences in YFP production rate and expression time were found between the wild-type and mutated RNA polymerases. Together, these results suggest that the lack of transcription-translation coupling is not the cause for the relatively slow translation rate observed in PURE*frex*. Alternatively, the tested conditions may not suffice to restore the appropriate degree of coupling.

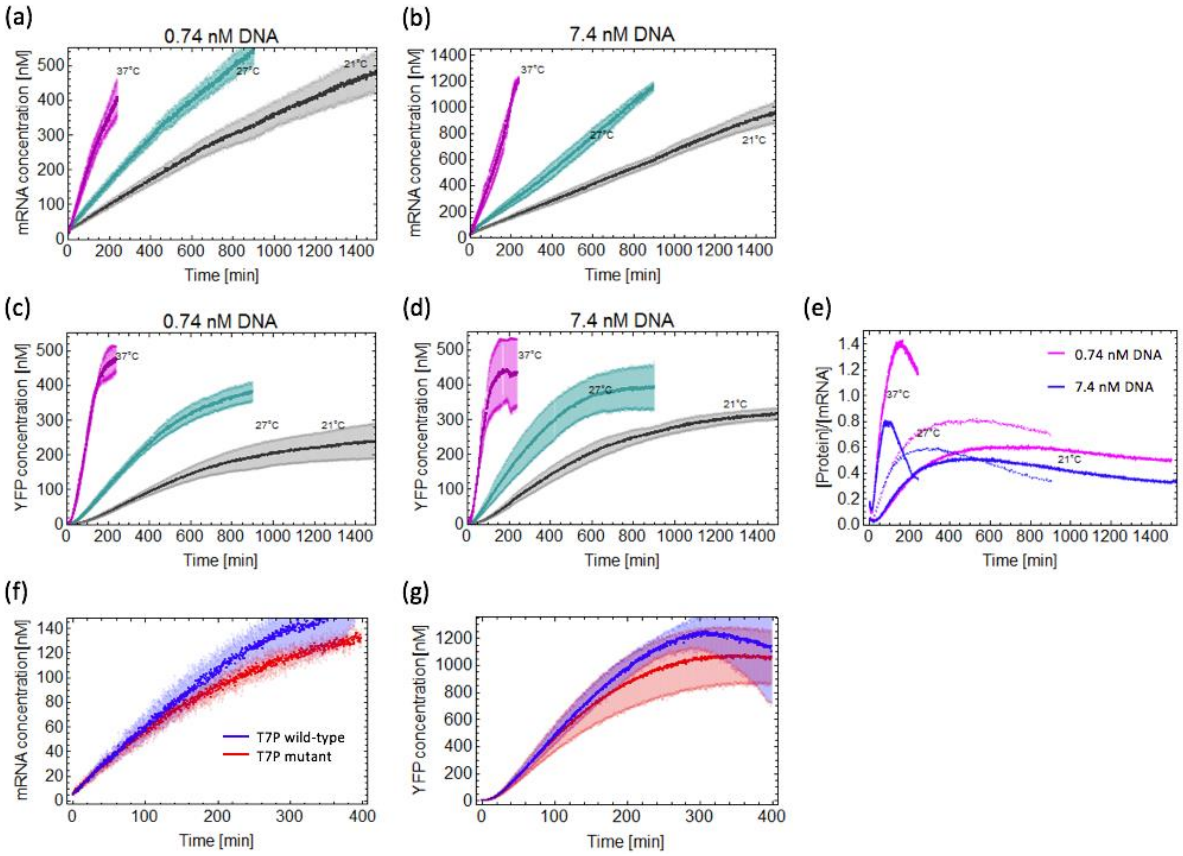

Figure 9: (a-e) Expression kinetics at different temperatures with PURE*frex*. (a,c) DNA concentration was 0.74 nM. Temperature: 37 °C (magenta), 27 °C (blue) and 21 °C (grey). (b,d) DNA concentration was 7.4 nM. Temperature: 37 °C (magenta), 27 °C (blue) and 21 °C (grey). (e) YFP-to-mRNA ratio for the different temperatures and DNA concentrations. (f, g) Comparison of mRNA (f) and protein (g) synthesis curves between the wild-type (blue curves) and slow T7RNA polymerase mutant (red curves) in PURE*frex* batch B. RNA polymerase concentration in all reactions was 30 nM, DNA concentration was 0.37 nM for reactions with wild-type polymerase and 7.4 nM with the mutant RNA polymerase. Bold lines indicate average values of three reactions and shaded regions represent the mean +/– 1 SD.

Another option to improve translation initiation if mRNA secondary structures impede formation of the initiation complex is to use an RNA helicase to unfold inhibitory structures. The *E. coli* helicase hrpA has been found to improve protein synthesis yield in cell-free translation reactions (83). In PURE*frex* reactions with purified mRNA template, hrpA increases the translation rate to the same level as with 7.4 nM DNA (Fig. 8(h)), but the yield and lifespan are lower than in the reaction with DNA. Addition of hrpA has no effect on the translation rate when using a DNA template (Fig. 8(i)). In reactions with PTH2 the yield is slightly lower as the translation rate starts to decrease at an earlier time point (Fig. 8(i)). To rule out a possible effect of ATP depletion (hrpA is an ATPase) more ATP was supplemented to the reaction once the translation rate started to decrease, but no increase in translation rate was observed (data not shown). These results corroborate the idea that mRNA exists in different functional states and that neither a transcription slow-down nor enzymatic clearance of mRNA secondary structures increases translation rate.

Targeted mRNA cleavage, but not increased concentrations of IF1 and IF2, improves the protein yield and lifetime in PURE*frex*

High mRNA concentrations lead to a faster depletion of tRNAs (Fig. 4), but to achieve high translation rates in coupled IVTT reactions high DNA concentrations, i.e. fast transcription rates, are required. To achieve high transcription rate together with long translation lifespan we reasoned that mRNA degradation needs to be implemented. This was realized by adding RNaseH to degrade a target mRNA hybridized to a DNA oligonucleotide that was complementary to the ribosome binding site (Fig. 10(a), Fig. S10). Combined with an increase in the concentration of T7 RNA polymerase to reach higher transcription rate, targeted mRNA degradation improved protein yield and lifespan, but not translation rate (Fig. 10(b)). Consistent with the earlier finding that higher mRNA concentrations lead to faster tRNA depletion, the lifespan and yield slightly decreased in control reactions with no active transcript degradation but with higher transcription rate (Fig. 10(c)).

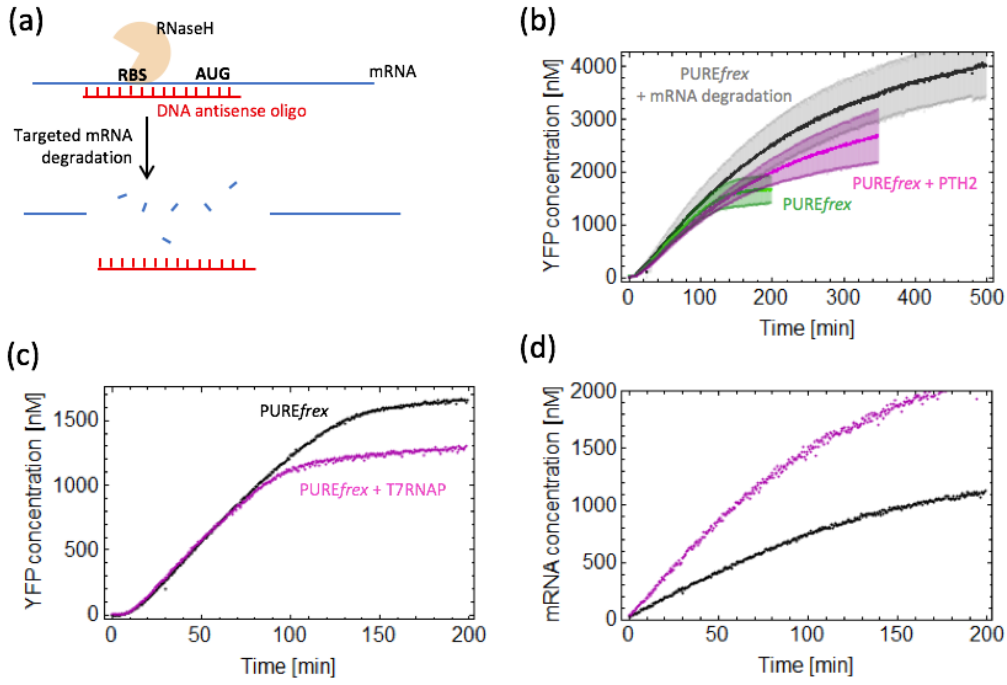

Figure 10: mRNA turnover improves expression time and yield in PURE*frex*. (a) Schematic of mRNA degradation with RNaseH and an antisense oligo. (b) Kinetics of YFP production with targeted mRNA degradation (grey). For comparison, reaction kinetics in standard PURE*frex* (green) and PURE*frex* supplemented with PTH2 (magenta) are appended. The average curves of at least three experiments (PURE*frex* batch B) are displayed, with shaded areas denoting +/– 1 SD. (c,d) Increasing transcription rate (by adding more T7 RNA polymerase, magenta trace) does not improve the yield of synthesized YFP (c) or the expression time. The corresponding mRNA production kinetics are shown in (d).

The longer lifespan in reactions with mRNA turnover is another indication that high mRNA concentrations are inhibitory. Together with the observation that translation initiation is inefficient in PURE system, this result suggests that abundant mRNA could deplete factors involved in translation initiation. An imbalance in the concentrations of mRNA and various components required for initiation, such as formyl-Met-tRNA, initiation factors and ribosomal subunits, could lead to the formation of inactive complexes, which would deplete the system from those factors. Such inactive initiation complexes have already been observed *in vitro*, but usually under conditions where one component is omitted (84). Given the composition of PURE*frex* (disclosure agreement with GeneFrontier) (see also in (51)), initiation factors 1 and 2 seem the most susceptible candidates for depletion. We tested this hypothesis in two different PURE*frex* batches by adding purified IF1 or IF2 in the presence of PTH2 at high DNA concentration. While IF1 addition resulted in an approximately two-fold increase in translation rate and higher yield in one batch (Fig. 11(a)), no improvement was measured with the second batch (Fig. 11(b)). Addition of IF2, either alone or in combination with IF1, led to a reduction in translation rate for both batches. The PURE*frex* composition is kept constant between batches in

terms of the concentration of individual components (personal communication with GeneFrontier), but the different effects of IF1 addition suggests that the fraction of active proteins may differ from one batch to another (although production of the standard protein DHFR should be identical, personal communication with GeneFrontier).

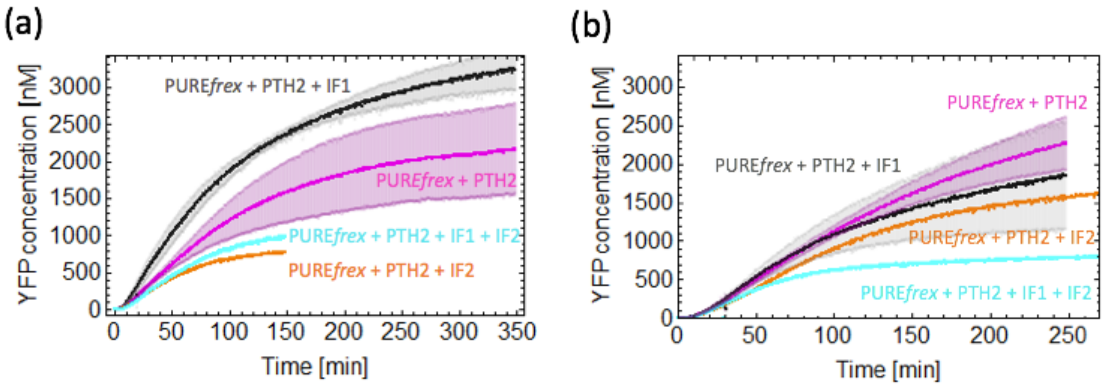

Figure 11: Addition of initiation factors 1 and 2 to PURE*frex* has different effects on batches A and B. (a) PURE*frex* batch A. (b) PURE*frex* batch B. PURE*frex* reactions with 100 nM PTH2 were supplemented with additional IF1 (grey, 1 µM), IF2 (orange, 200 nM) or both (cyan). Control reactions without additional initiation factors are also shown (magenta). Average traces +/– 1 SD from three independent experiments (magenta, grey), or single traces (cyan, orange) are reported.

Although we were able to identify some bottlenecks and improve performance in some specific conditions, overall the data suggest the existence of other (multiple) control mechanisms as none of the attempts led to major improvements under a wide range of initial conditions.

Attempts to build a new integrated mechanistic model

Several experimental observations with PURE*frex* indicate that mRNA can exist in more than one functional state: (i) IVTT kinetics with different DNA concentrations exhibit a different slope of YFP fluorescence intensity versus time in the linear regime, while at that time mRNA concentration is still increasing. (ii) DNase addition leads to altered YFP synthesis kinetics with a fast decrease of the initial translation rate, followed by a further decrease on a slower timescale. (iii) Maximum translation rate in reactions starting with purified mRNA is lower than for reactions with high DNA concentrations, but the maximum rate can be achieved when the RNA helicase hrpA is added to reactions with purified mRNA. To include these findings, a new model with three mRNA states was constructed (Fig. 12(a)). Transcription generates mRNA in its most active state, from which it can be irreversibly converted to a translationally less active state, and finally to a fully inactive state. The reasoning behind this stepwise inactivation of mRNA lies on the idea that nascent mRNA with cleared RBS is less prone to fold into inhibitory secondary structures compared to longer – full-length – mRNA and has a higher probability of ribosome binding. This kinetics scheme is also consistent with the observation that the highest translation

rates are achieved in expression reactions starting with DNA and not with purified mRNA. Although a two-state model (mRNA active or partly inactive) was sufficient to explain the YFP fluorescence kinetics with DNase addition as well as for different DNA concentrations, a third - translationally inactive state - was also included to further reproduce the decreased translation rate measured in the experiments with delayed ribosome addition. Note that unintended mRNA degradation was not considered in the model as no nuclease activity could be detected in PURE*frex* (Fig. 1(a), (41)).

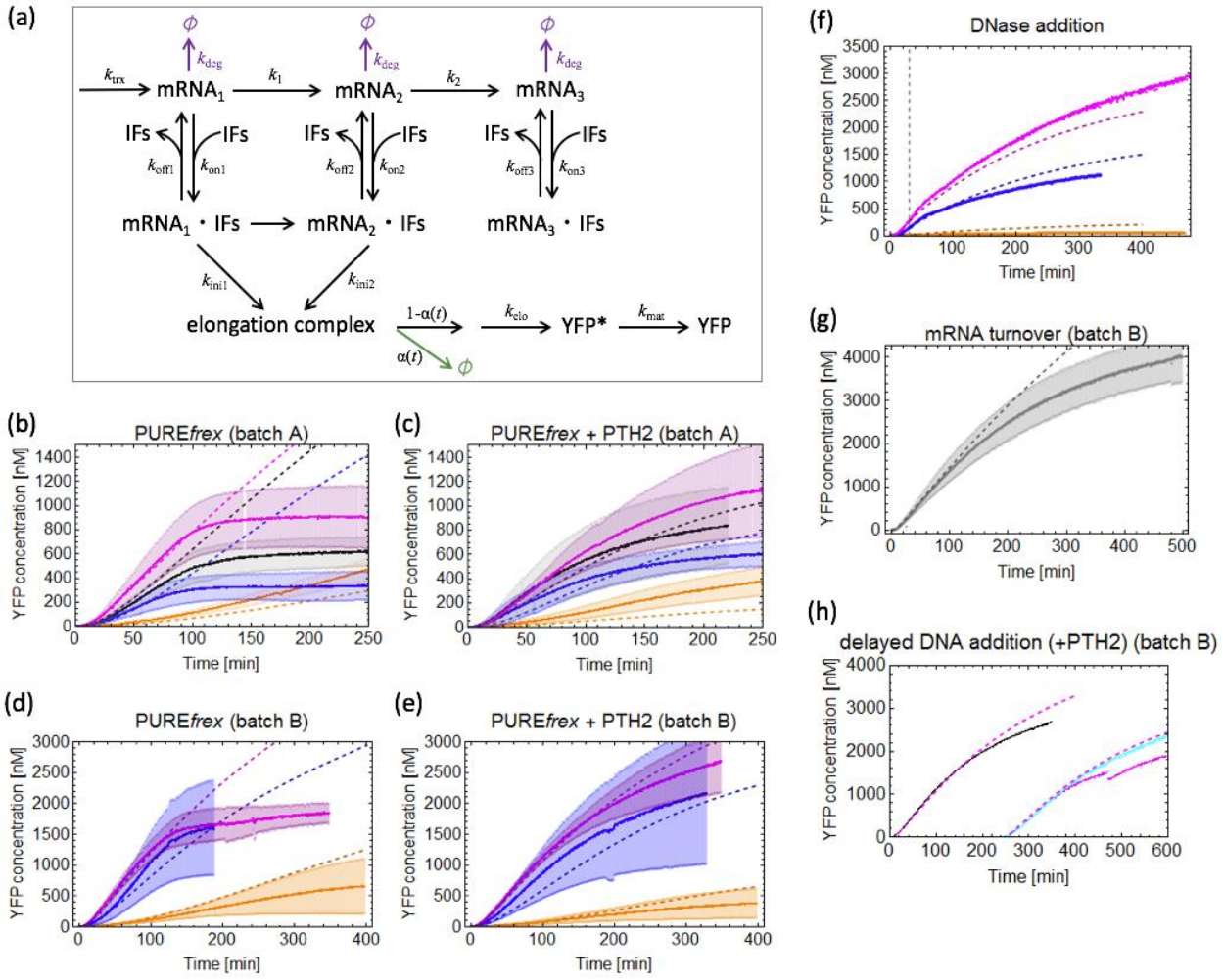

Figure 12: An improved model accounting for various mRNA states captures some, but not all, of the experimental data observed with PURE*frex*. (a) Schematic overview of the proposed ODE kinetic model. The arrow depicted in green represents abortive translation reactions that may occur with a time-dependent probability $\alpha(t)$ and lead to a decrease in production of full-length YFP. Disassembly of the elongation complex can be caused by peptidyl-tRNA drop-off or PTH2-catalyzed hydrolysis of peptidyl-tRNA that is still lodged in the ribosome. As mRNA transitions from states 1 to 2 and 3, its ability to bind IFs and engage into an elongation complex decreases. Abbreviations: IFs, initiation factors, YFP*, non-matured YFP, k_{trx}, transcription rate constant, k_{deg}, rate constant at which mRNA is cleaved by the RNaseH in turnover experiments, $k_{ini1/2}$, translation initiation rate constants, k_{elo}, translation elongation

rate constant, k_{mat}, chromophore maturation rate constant. The parameter values are shown in Table 2. (b,d) Experimental data (solid lines) and model fitting (dashed lines) for reactions with different DNA concentrations. Color-coding is as in Fig. 1. (c,e) Experimental data (solid lines) and model fitting (dashed lines) for reactions supplemented with 100 nM PTH2 with different DNA concentrations. (b-e) Color-coding is as in Fig. 1. Three independent experiments were performed for each condition and the average values +/– 1 SD are reported. (f,g,h) Experimental data (solid lines) and model fitting (dashed lines) for reactions with DNase addition (experimental curves are from Fig. 5, same color coding), mRNA turnover and delayed DNA addition (experimental curves are from Fig. 8(g)), respectively.

Further, translation initiation was modelled by binding of a generic "initiation factor" to mRNA, whose binding constants differ depending on the mRNA's state. The initiation factor in this model represents the rate-limiting component for the initiation step, which could be one of the canonical initiation factors, as well as formyl-Met-tRNA, or either of the 30S or 50S ribosomal subunits. Because inactive mRNAs can also bind to this rate-limiting initiation factor, large amounts of mRNA will inevitably deplete it from the system. The complex formed between the mRNA in a fully or partly active state and the generic initiation factor can then develop into a translation elongation complex. The model describes translation elongation and termination as a single step, since we conclude from the AFM experiments that neither of those steps is rate limiting.

Maturation of YFP to its fluorescent form was modelled by a first order reaction. Although the maturation rate was measured experimentally to be 0.06 min^{-1} (39), a better fit could be obtained with a faster maturation rate of 0.2 min^{-1}, regardless of the nature of the template (purified DNA or mRNA).

We tested a number of different models accounting for tRNA depletion as a process contributing to premature cessation of protein synthesis. For all models we assumed Michaelis-Menten kinetics for the dependence of the translation reaction on tRNAs. Depletion of tRNAs was either modelled as a simple time-dependent process with first order rate or as resulting from peptidyl-tRNA drop-off after an elongation complex has formed. In the latter case, the frequency of drop-off events was either modelled as a fixed probability, or the probability was allowed to either change with time or depend on the 'state' of the mRNA. However, no satisfactory qualitative global agreement (fits within +/– 1 SD) with the complete experimental data could be obtained. Hence, the model depicted in Fig. 12(a) is only meant to describe IVTT dynamics up to and including the steady regime of YFP production. For those time intervals a reasonable fit can be obtained that concurrently describes the apparent translation kinetics at different DNA concentrations, as well as the experiments with DNase addition (Fig. 12(b,d,f)). The corresponding parameter values are reported in Table 2. Transcription rates for all reactions were obtained by fitting the monitored spinach fluorescence signal. To fit the data obtained with the two different batches of PURE*frex*, the initial concentration of the rate-limiting initiation factor was allowed to vary to account for possible differences in the active protein fractions, while all other model parameters were set to the same values for all fits. Such a batch-to-batch

compositional variability could also motivate the adjustment of some rate constant values (e.g. the translation elongation rate constant k_{elo}) as some of them implicitly depend on concentrations of components (e.g. elongation factors).

The experiments with mRNA turnover can be fit by including mRNA degradation in the model. Degradation was modelled as a first order reaction with the same rate constant for all mRNA states. As shown in Fig. 12(g), although the model does predict the initial slope correctly, translation rate decreases faster over time than predicted by the model, a recurrent problem for reactions without PTH2.

For reactions with PTH2 we observed a time-dependent decrease in translation rate (Fig. 8(c,d)). This could be due to a side reaction of PTH2 that catalyses peptidyl-tRNA hydrolysis of substrates on the ribosome (see Discussion). In the model this side-reactivity was included as a certain (time-dependent) probability for the elongation complex to fall apart without generating protein. The probability for the generation of full-length protein from the elongation complex was described with a decreasing exponential function with the rate chosen to fit the delayed DNA addition experiment (Fig. 12(h)). With this reaction added, the apparent translation kinetics with PTH2 were also fit by the model with the same global fit parameters (Fig. 12(c,e)) and a good agreement was obtained (Table 2).

Parameter name	**Value**
k_1	1 min^{-1}
k_2	0.004 min^{-1}
k_{on1}	30 min^{-1} nM^{-1}
k_{on2}, k_{on3}	0.1 min^{-1} nM^{-1}
k_{off1}	400 min^{-1}
k_{off2}, k_{off3}	400 min^{-1}
k_{ini1}	1.2 min^{-1}
k_{ini2}	0.4 min^{-1}
IF	PURE*frex* batch A: 30 nM, batch B: 45 nM
K_m	60 nM
k_{deg}	Only in reactions with mRNA turnover: 0.5 min^{-1}
tRNA	PURE*frex* batch A: 2400 nM, batch B: 600 nM
k_{mat}	0.2 min^{-1}
$\alpha(t)$	PURE*frex* batch A: $(1 - e^{-0.005t/min})$, batch B: $(1 - e^{-0.001t/min})$
k_{elo}	0.5 min^{-1}

Table 2: Parameter values of the gene expression model shown in Fig. 12(a).

DISCUSSION

Harnessing IVTT systems for controlled synthesis of multiple proteins is essential for proper execution of biological operations in a ribosome-based synthetic minimal cell. It implies to better apprehend the molecular mechanisms governing transcription and translation, and to integrate them in a biochemically reasonable, yet conceptually simple, kinetic model. In particular, the reasons for translation cessation need to be identified to rationally design new experiments for prolonged protein production. In addition, biological and technical variability is inherent to cell-free gene expression experiments, as reported here with the PURE system (Fig. 1) and elsewhere with cell lysates (29). This imposes the explicit treatment of uncertainty when assigning confidence levels of model fitting and robustness of predictions.

Practical obstacles contributing to technical variability include the handling of small volumes (pipetting, evaporation, large surface/interface effects), mixing of multiple reagents and the exact nature of the measurement cuvette (cleanness, composition, geometry). The principal factors contributing to biological variability are the DNA or mRNA template (purity, integrity, possible mutations introduced by PCR) and PURE system batch-to-batch variability. Despite possible variability in the activity levels of the tRNA mix, ribosome and individual proteins, consistency of PURE*frex* productivity across different batches is systematically validated by the supplier by quantifying the amount of synthesized dihydrofolate reductase protein with radioactivity (personal communication). Therefore, the extent of variability of the experimental output might be DNA construct-dependent. It could also depend on the exact expression conditions (test tube vs. cuvette) and on the type of readout (protein gel, absorbance, radioactivity, fluorescence spectroscopy).

We here discuss a number of mechanisms that could lead to premature cessation or slowing down of translation: exhaustion of the pool of tRNAs through peptidyl-tRNA drop-off, depletion of translation initiation factors by inactivated mRNA, inactivation of one of the elongation factors. Given that tRNA depletion is a factor contributing to translation arrest in reactions with >1.4 nM DNA template in PURE*frex*, we first examine possible mechanisms causing tRNA inactivation or depletion. A first order (time) inactivation of tRNAs, e.g. by unfolding, can be ruled out by the delayed DNA addition experiments (Fig. 3). Alternatively, tRNA might be subject to oxidative damage (e.g. via the hydrogen peroxide released during YFP chromophore maturation) affecting the 2-thiouridine (85, 86). The fact that tRNA depletion can be remedied by addition of peptidyl-tRNA-hydrolase, however, indicates occurrence of peptidyl-tRNA drop-off events. Moreover, the shortened linear regime observed in some of the reactions with peptidyl-tRNA-hydrolase suggests occurrence of unintended side reactions. It has been reported that even though the main substrate is free peptidyl-tRNA, the hydrolase also has an activity on ribosome-bound peptidyl-tRNA, in particular on stalled translation complexes or when the elongation rate is slow (87, 88). Both the deviation from the linear regime and the decreased translation rate in reactions with delayed DNA addition could be explained if there is a time dependent decrease in the translation elongation rate, e.g. by inactivation of one of the elongation

factors, which is not rate limiting under standard conditions but leads to increased catalyzed ribosome drop-off over time in reactions with peptidyl-tRNA-hydrolase.

Noteworthily, the *yfp* gene sequence used in this study was codon-optimized to match the abundance of the *E. coli* MRE600 tRNAs present in the PURE systems. A mismatch of codon usage with respect to the expression apparatus is known to deplete low-abundance tRNAs, which may slow down translation elongation and cause the ribosomes to stall (89). Although expression of the codon-optimized DNA template clearly remains susceptible to adverse processes involving tRNAs, the apparent translation rate and maximum yield of fluorescent YFP are respectively about two times and 1.5 time higher than with the non-optimized construct (Fig. S11), showing that coding sequence optimization is a valid strategy to ameliorate productivity in the PURE system. We noticed that codon usage at the +2 position could be further improved by replacing the CGG codon (corresponding to the low abundant $tRNA_{Arg3}$ (Table S2)) by the synonymous CGU codon. However, introducing the corresponding mutation in the DNA did not improve the gene expression properties (Fig. S12), indicating that a possible peptidyl-tRNA drop-off at that position (69) is not limiting the reaction. In general, DNA sequence optimization to avoid rare codons, while alternating codon usage for the most abundant amino acids to lessen solicitation of the same aatRNAs should prove to be a valuable design strategy.

Unfortunately, simple models of tRNA depletion failed to describe the observed translation kinetics. Because tRNA depletion is faster with higher mRNA concentrations (Fig. 4(c)), even at mRNA levels that exceed the concentration at which the translation rate saturates, the probability that an (early) elongation complex either continues or drops off, cannot be a fixed value. Drop-off events are expected to be more frequent when the translation elongation rate is slow due to kinetic competition between drop-off and normal elongation. The observed relationship between YFP production rate and expression lifetime for different mRNA concentrations (Fig. 4(c)) might be explained if one assumes that translation elongation is slower in reactions with higher mRNA concentration, e.g. because elongation factors distribute over a higher total number of translating ribosomes, thereby lowering the concentration of free elongation factors. This is however unlikely for two reasons. First, only a low fraction of ribosomes is actively translating (Fig. 7(c)), implying that the associated change in free concentration of elongation factors is minor. Second, the YFP production rate can be increased by adding the RNA helicase hrpA which most likely affects the rate of initiation but not of elongation.

Inactivation of mRNA by intramolecular secondary structures is a well-acknowledged cause for inefficient translation in cell-free systems (58, 79, 90). Intramolecular base pairing occluding the Shine-Dalgarno (SD) sequence can greatly reduce the probability of translation initiation and affect protein levels *in vivo* (91). Whereas tight coupling between the rate of transcription and translation initiation limits formation of inhibitory RNA structures and ensures efficient usage of mRNA, the level of cooperation between these two processes is low in the T7 RNA polymerase-based PURE system. Computer predictions of equilibrium mRNA secondary structures reveal that local folding may sequester the SD sequence (Fig. S13). Note that equilibrium folding states might not be the only relevant structures as transient, higher energy folding states could also

affect translation initiation (91). Avoidance of upstream and downstream anti-SD sequences in the vicinity of the SD by synonymous nucleotide substitution might increase translation efficiency. As a novel mRNA-directed inhibitory mechanism, we find here that accumulation of RBS-containing transcripts can cause depletion of some translation factors, reducing the yield and rate of protein synthesis. Notably, coding-sequence optimization of DNA to minimize formation of mRNA secondary structures and to improve codon usage, as performed here, does not abolish these limitations.

The strategies deployed here to increase the PURE system productivity could be combined with previous attempts. Owing to its modularity, the PURE system composition could be tuned by adjusting the concentration of various components, such as tRNAs, ribosomes, some initiation and elongation factors and magnesium acetate (74, 92, 93), or by replenishing low-molecular weight substrates (74, 93, 94). Other factors that are not present in the original composition have also been supplemented, e.g. the bovine serum albumin, crowding agents, EF-P (93, 95, 96), and proteins that recycle stalled ribosomes (93). Regarding the addition of ribosome rescuing proteins, the absence of detectable amounts of degraded mRNA by nuclease contaminants in PURE*frex* suggests that implementation of rescuing systems, such as tmRNA and ArfA, will only have moderate influence. As reported earlier, as well as in the present study, longer-lived protein expression and higher reaction efficiency could be accomplished under certain experimental conditions and with specific reporter templates. However, formulating generic guidelines to design DNA constructs and to customize a PURE system for improved performance clearly remains a challenge.

It is evident that the premature termination of protein production is not the cause of a single limiting factor and is dependent on the exact initial conditions, e.g. DNA template sequence, PURE system composition, temperature and even the volume and type of reaction vessel. At the system's level, the overall stoichiometric imbalance of some translation factors will lead to an increasingly higher fraction of misassembled complexes, which eventually becomes lethal. Scattering of initiation factors that engage into nonproductive complexes, such as the formation of an fMet-tRNA-empty 70S complex in the absence of IF3 is an example of suboptimal usage of the translation machinery.

CONCLUSION

PURE*frex* is a commercially available *E. coli*-based reconstituted IVTT system with a simpler protein hardware than cell lysates but also than PURExpress owing to improved purification steps (Table S1). Despite its minimal nature and the fact that it operates in absence of other constraints of energy and resources, it is fair to acknowledge that PURE*frex* is recalcitrant to modeling. Moreover, simple kinetic models originally developed for other IVTT systems fail to adequately reproduce its basic dynamic behavior (Fig. 12). In PURExpress, (transcription and translation independent) ribosome inactivation was previously identified as the cause for the saturation of protein synthesis (35). Our experimental results are however consistent with a

combined influence of multiple limiting factors, including the conversion of synthesized transcripts into inactive templates, the consecutive sequestration of translation factors and depletion of tRNAs in the form of peptidyl-tRNAs. Translation initiation was identified as a bottleneck, and the poor usage of both mRNA and ribosomes was revealed. We devised a preliminary model that encompasses the aforementioned processes. Simulated protein kinetics show closer agreement to the measured curves across the entire data sets than with simpler models. However, it is clear that next generation kinetic models with higher predictive power have yet to be constructed.

Higher dimensional models have been described in the literature (31, 32, 34, 97). They more realistically account for the numerous biochemical steps and explicitly include additional species, such as ribosome subunits, aminoacyl-tRNA synthetases and energy regeneration enzymes. However, in order to describe the present experimental observations, these models will have to be upgraded to implement the limiting processes revealed here.

ACKNOWLEDGEMENTS

The authors would like to thank Jeremie Capoulade for technical assistance with fluorescence correlation spectroscopy, Allard Katan for technical assistance with the AFM experiments, and Reza Maleki Seifar, Martijn Pinkse, Mervin Pieterse, Angela ten Pierick and Cor Ras from the Biotechnology department for technical assistance with mass spectrometry measurements. We also thank Kasper van Schie and Alex de Mulder for contributing preliminary models of gene expression, Roeland van Wijk for preparing samples for mass spectrometry analysis of amino acids and nucleotides, Ilja Westerlaken for purifying the mutated T7 RNA polymerase, Just van der Wolf for contributing preliminary experiments with the addition of tRNA to restart translation, and Tom Venema for the codon usage analysis shown in supplementary Table 2. We thank all the Danelon lab members who contributed intellectual inputs and experiments over the past eight years, in particular Duco Blanken, Andrew Scott and Zohreh Nourian. We are grateful to Dr. Tomoko Miyagi and Dr. Takashi Kanamori from GeneFrontier Corporation (Chiba, Japan) for fruitful discussions and for the disclosure agreement with regard to PURE*frex* composition. We also thank Isaac Meek and Kiran Gulati from New England Biolabs for discussions on the PURExpress proteomic analysis. This work was financially supported by the Netherlands Organization for Scientific Research (NWO/OCW) through a VIDI grant (project number 723.012.007) to C.D. and the Frontiers of the Nanoscience Program.

AUTHOR CONTRIBUTIONS

E.R. performed the protein/peptide analysis by mass spectrometry. M.H. contributed preliminary experiments with PTH2 activity assays. P.N. performed the gene expression kinetics shown in Fig. 9(a-d) and in Fig. S11, and contributed initial experiments with DNase and with delayed

addition of DNA or ribosomes. J.K. carried out the experiments with the dialysis chamber. K.W. contributed preliminary experiments with delayed addition of DNA and tRNAs, and contributed a preliminary kinetic model. All other experiments have been performed and analysed by A.D., who also developed the deterministic and stochastic models. A.W. supervised the mass spectrometry experiments. C.D. designed and supervised the research. A.D. and C.D. wrote the manuscript. All the authors commented on the manuscript.

ADDITIONAL INFORMATIONS

Supplementary information is available in the online version of the paper.
Supplementary File 1: Supplementary text, figures and tables
Supplementary File 2: Supplementary Table 1
Supplementary File 3: ODE script file

REFERENCES

1. Bretscher MS. Direct Translation of a Circular Messenger DNA. Nature. 1968;220:1088.
2. Salas J, Bollum FJ. Biosynthetic Polydeoxynucleotides as Direct Templates for Polypeptide Synthesis. Journal of Biological Chemistry. 1968;243(5):1012-5.
3. Nevin DE, Pratt JM. A coupled in vitro transcription-translation system for the exclusive synthesis of polypeptides expressed from the T7 promoter. FEBS Letters. 1991;291(2):259-63.
4. Spirin A, Baranov V, Ryabova L, Ovodov S, Alakhov Y. A continuous cell-free translation system capable of producing polypeptides in high yield. Science. 1988;242(4882):1162-4.
5. Pardee K, Green AA, Ferrante T, Cameron DE, DaleyKeyser A, Yin P, et al. Paper-based synthetic gene networks. Cell. 2014;159(4):940-54.
6. Slomovic S, Pardee K, Collins JJ. Synthetic biology devices for in vitro and in vivo diagnostics. Proceedings of the National Academy of Sciences. 2015;112(47):14429-35.
7. Jewett MC, Noireaux V. Tailor-made genetic codes. Nature Chemistry. 2016;8:291.
8. Marshall R, Maxwell CS, Collins SP, Jacobsen T, Luo ML, Begemann MB, et al. Rapid and Scalable Characterization of CRISPR Technologies Using an *E. coli* Cell-Free Transcription-Translation System. Molecular Cell. 2017;69(1):146-57.e3.
9. Niederholtmeyer H, Stepanova V, Maerkl SJ. Implementation of cell-free biological networks at steady state. Proc Natl Acad Sci U S A. 2013;110(40):15985-90.
10. Niederholtmeyer H, Sun ZZ, Hori Y, Yeung E, Verpoorte A, Murray RM, et al. Rapid cell-free forward engineering of novel genetic ring oscillators. Elife. 2015;4:e09771.
11. Shin J, Noireaux V. An E. coli cell-free expression toolbox: application to synthetic gene circuits and artificial cells. ACS Synth Biol. 2012;1(1):29-41.
12. Shin J, Noireaux V. Study of messenger RNA inactivation and protein degradation in an Escherichia coli cell-free expression system. J Biol Eng. 2010;4:9.
13. Garamella J, Marshall R, Rustad M, Noireaux V. The All E. coli TX-TL Toolbox 2.0: A Platform for Cell-Free Synthetic Biology. ACS Synth Biol. 2016;5(4):344-55.
14. Noireaux V, Bar-Ziv R, Libchaber A. Principles of cell-free genetic circuit assembly. Proceedings of the National Academy of Sciences. 2003;100(22):12672-7.
15. The Minimal Cell: The Biophysics of Cell Compartment and the Origin of Cell Functionality. Pier Luigi Luisi, Pasquale Stano (Eds.) Springer, 1st Edition., 2011.
16. Nomura SM, Tsumoto K, Hamada T, Akiyoshi K, Nakatani Y, Yoshikawa K. Gene expression within cell-sized lipid vesicles. ChemBioChem. 2003;4(11):1172-1175.
17. Noireaux V, Libchaber A. A vesicle bioreactor as a step toward an artificial cell assembly. Proc Natl Acad Sci USA. 2004;101(51):17669-17674.
18. Nourian Z, Roelofsen W, Danelon C. Triggered gene expression in fed-vesicle microreactors with a multifunctional membrane. Angew Chem Int Ed. 2012;51(13):3114-3118.
19. Scott A, Noga MJ, de Graaf P, Westerlaken I, Yildirim E, Danelon C. Cell-Free Phospholipid Biosynthesis by Gene-Encoded Enzymes Reconstituted in Liposomes. PLoS One. 2016;11(10):e0163058.
20. van Nies P, Westerlaken I, Blanken D, Mencia M, Salas M, Danelon C. Self-replication of DNA by its encoded proteins in liposome-based synthetic cells. Nat. Commun. 2018;9(1):1583-1594.
21. Underwood KA, Swartz JR, Puglisi JD. Quantitative polysome analysis identifies limitations in bacterial cell-free protein synthesis. Biotechnol Bioeng. 2005;91(4):425-35.
22. van Roekel HW, Meijer LH, Masroor S, Felix Garza ZC, Estevez-Torres A, Rondelez Y, et al. Automated design of programmable enzyme-driven DNA circuits. ACS Synth Biol. 2015;4(6):735-45.
23. Baccouche A, Montagne K, Padirac A, Fujii T, Rondelez Y. Dynamic DNA-toolbox reaction circuits: A walkthrough. Methods. 2014;67(2):234-49.

24. Krinsky N, Kaduri M, Shainsky-Roitman J, Goldfeder M, Ivanir E, Benhar I, et al. A Simple and Rapid Method for Preparing a Cell-Free Bacterial Lysate for Protein Synthesis. PLOS ONE. 2016;11(10):e0165137.
25. Kwon Y-C, Jewett MC. High-throughput preparation methods of crude extract for robust cell-free protein synthesis. Scientific Reports. 2015;5:8663.
26. Shimizu Y, Inoue A, Tomari Y, Suzuki T, Yokogawa T, Nishikawa K, et al. Cell-free translation reconstituted with purified components. Nature Biotechnology. 2001;19:751.
27. Kung HF, Redfield B, Treadwell BV, Eskin B, Spears C, Weissbach H. DNA-directed in vitro synthesis of beta-galactosidase. Studies with purified factors. Journal of Biological Chemistry. 1977;252(19):6889-94.
28. Karr JR, Sanghvi JC, Macklin DN, Gutschow MV, Jacobs JM, Bolival B, Jr., et al. A whole-cell computational model predicts phenotype from genotype. Cell. 2012;150(2):389-401.
29. Takahashi MK, Chappell J, Hayes CA, Sun ZZ, Kim J, Singhal V, et al. Rapidly Characterizing the Fast Dynamics of RNA Genetic Circuitry with Cell-Free Transcription–Translation (TX-TL) Systems. ACS Synthetic Biology. 2015;4(5):503-15.
30. Matsuura T, Kazuta Y, Aita T, Adachi J, Yomo T. Quantifying epistatic interactions among the components constituting the protein translation system. Mol Syst Biol. 2009;5:297.
31. Matsuura T, Tanimura N, Hosoda K, Yomo T, Shimizu Y. Reaction dynamics analysis of a reconstituted Escherichia coli protein translation system by computational modeling. Proc Natl Acad Sci U S A. 2017;114(8):E1336-E44.
32. Niess A, Failmezger J, Kuschel M, Siemann-Herzberg M, Takors R. Experimentally Validated Model Enables Debottlenecking of in Vitro Protein Synthesis and Identifies a Control Shift under in Vivo Conditions. ACS Synth Biol. 2017;6(10):1913-21.
33. Calviello L, Stano P, Mavelli F, Luisi PL, Marangoni R. Quasi-cellular systems: stochastic simulation analysis at nanoscale range. BMC Bioinformatics. 2013;14 Suppl 7:S7.
34. Frazier JM, Chushak Y, Foy B. Stochastic simulation and analysis of biomolecular reaction networks. BMC Syst Biol. 2009;3:64.
35. Stogbauer T, Windhager L, Zimmer R, Radler JO. Experiment and mathematical modeling of gene expression dynamics in a cell-free system. Integr Biol (Camb). 2012;4(5):494-501.
36. Karzbrun E, Shin J, Bar-Ziv RH, Noireaux V. Coarse-grained dynamics of protein synthesis in a cell-free system. Phys Rev Lett. 2011;106(4):048104.
37. Sokolova E, Spruijt E, Hansen MMK, Dubuc E, Groen J, Chokkalingam V, et al. Enhanced transcription rates in membrane-free protocells formed by coacervation of cell lysate. Proceedings of the National Academy of Sciences. 2013;110(29):11692-7.
38. Mavelli F, Marangoni R, Stano P. A Simple Protein Synthesis Model for the PURE System Operation. Bull Math Biol. 2015;77(6):1185-212.
39. van Nies P, Canton AS, Nourian Z, Danelon C. Chapter Ten - Monitoring mRNA and Protein Levels in Bulk and in Model Vesicle-Based Artificial Cells. In: Burke-Aguero DH, editor. Methods in Enzymology. 550: Academic Press; 2015. p. 187-214.
40. Paige JS, Wu KY, Jaffrey SR. RNA mimics of green fluorescent protein. Science 2012;333(6042):642-6.
41. van Nies P, Nourian Z, Kok M, van Wijk R, Moeskops J, Westerlaken I, et al. Unbiased tracking of the progression of mRNA and protein synthesis in bulk and in liposome-confined reactions. Chembiochem. 2013;14(15):1963-6.
42. Wu L, Mashego MR, van Dam JC, Proell AM, Vinke JL, Ras C, et al. Quantitative analysis of the microbial metabolome by isotope dilution mass spectrometry using uniformly 13C-labeled cell extracts as internal standards. Analytical Biochemistry. 2005;336(2):164-71.

43. Mashego MR, Wu L, Dam JCV, Ras C, Vinke JL, Winden WAV, et al. MIRACLE: mass isotopomer ratio analysis of U-13C-labeled extracts. A new method for accurate quantification of changes in concentrations of intracellular metabolites. Biotechnology and Bioengineering. 2004;85(6):620-8.
44. Seifar RM, Ras C, van Dam JC, van Gulik WM, Heijnen JJ, van Winden WA. Simultaneous quantification of free nucleotides in complex biological samples using ion pair reversed phase liquid chromatography isotope dilution tandem mass spectrometry. Analytical Biochemistry. 2009;388(2):213-9.
45. Wahl SA, Seifar RM, ten Pierick A, Ras C, van Dam JC, Heijnen JJ, et al. Quantitative Metabolomics Using ID-MS. In: Krömer JO, Nielsen LK, Blank LM, editors. Metabolic Flux Analysis: Methods and Protocols. New York, NY: Springer New York; 2014. p. 91-105.
46. Mikamo-Satoh E, Takagi A, Tanaka H, Matsumoto T, Nishihara T, Kawai T. Profiling of gene-dependent translational progress in cell-free protein synthesis by real-space imaging. Anal Biochem. 2009;394(2):275-80.
47. Itzkovitz S, Lyubimova A, Blat IC, Maynard M, van Es J, Lees J, et al. Single-molecule transcript counting of stem-cell markers in the mouse intestine. Nature Cell Biology. 2011;14:106.
48. Jain C. Overexpression and purification of tagged Escherichia coli proteins using a chromosomal knock-in strategy. Protein Expression and Purification. 2006;46(2):294-8.
49. He B, Rong M, Lyakhov D, Gartenstein H, Diaz G, Castagna R, et al. Rapid Mutagenesis and Purification of Phage RNA Polymerases. Protein Expression and Purification. 1997;9(1):142-51.
50. Niederholtmeyer H, Xu L, Maerkl SJ. Real-time mRNA measurement during an in vitro transcription and translation reaction using binary probes. ACS Synth Biol. 2013;2(8):411-7.
51. Shimizu Y, Kanamori T, Ueda T. Protein synthesis by pure translation systems. Methods. 2005;36(3):299-304.
52. Chizzolini F, Forlin M, Cecchi D, Mansy SS. Gene position more strongly influences cell-free protein expression from operons than T7 transcriptional promoter strength. ACS Synth Biol. 2014;3(6):363-71.
53. Burnett BJ, Altman RB, Ferrao R, Alejo JL, Kaur N, Kanji J, et al. Elongation Factor Ts Directly Facilitates the Formation and Disassembly of the Escherichia coli Elongation Factor Tu·GTP·Aminoacyl-tRNA Ternary Complex. The Journal of Biological Chemistry. 2013;288(19):13917-28.
54. Wilden B, Savelsbergh A, Rodnina MV, Wintermeyer W. Role and timing of GTP binding and hydrolysis during EF-G-dependent tRNA translocation on the ribosome. Proceedings of the National Academy of Sciences. 2006;103(37):13670-5.
55. Jewett MC, Swartz JR. Substrate replenishment extends protein synthesis with an in vitro translation system designed to mimic the cytoplasm. Biotechnology and Bioengineering. 2004;87(4):465-71.
56. Kim D-M, Swartz JR. Efficient production of a bioactive, multiple disulfide-bonded protein using modified extracts of Escherichia coli. Biotechnology and Bioengineering. 2004;85(2):122-9.
57. Schoborg JA, Hodgman CE, Anderson MJ, Jewett MC. Substrate replenishment and byproduct removal improve yeast cell-free protein synthesis. Biotechnol J. 2014;9(5):630-40.
58. Hansen MM, Ventosa Rosquelles M, Yelleswarapu M, Maas RJ, van Vugt-Jonker AJ, Heus HA, et al. Protein Synthesis in Coupled and Uncoupled Cell-Free Prokaryotic Gene Expression Systems. ACS Synth Biol. 2016;5(12):1433-40.
59. Mikamo-Satoh E, Takagi A, Tanaka H, Matsumoto T, Nishihara T, Kawai T. Profiling of gene-dependent translational progress in cell-free protein synthesis by real-space imaging. Anal Biochem. 2009;394(2):275-80.
60. Takahashi S, Tsuji K, Ueda T, Okahata Y. Traveling time of a translating ribosome along messenger RNA monitored directly on a quartz crystal microbalance. J Am Chem Soc. 2012;134(15):6793-800.

61. Pavlov MY, Watts RE, Tan Z, Cornish VW, Ehrenberg M, Forster AC. Slow peptide bond formation by proline and other *N*-alkylamino acids in translation. Proceedings of the National Academy of Sciences. 2009;106(1):50-4.
62. Whitaker WR, Lee H, Arkin AP, Dueber JE. Avoidance of Truncated Proteins from Unintended Ribosome Binding Sites within Heterologous Protein Coding Sequences. ACS Synthetic Biology. 2015;4(3):249-57.
63. Poole ES, Brown CM, Tate WP. The identity of the base following the stop codon determines the efficiency of in vivo translational termination in Escherichia coli. EMBO Journal. 1995;14(1):151-8.
64. Wenthzel A-MK, Stancek M, Isaksson LA. Growth phase dependent stop codon readthrough and shift of translation reading frame in Escherichia coli. FEBS Letters. 1998;421:237-42.
65. Cruz-Vera LR, Magos-Castro MA, Zamora-Romo E, Guarneros G. Ribosome stalling and peptidyl-tRNA drop-off during translational delay at AGA codons. Nucleic Acids Res. 2004;32(15):4462-8.
66. Sin C, Chiarugi D, Valleriani A. Quantitative assessment of ribosome drop-off in E. coli. Nucleic Acids Res. 2016;44(6):2528-37.
67. Karimi R, Pavlov MY, Heurgué-Hamard V, Buckingham RH, Ehrenberg M. Initiation factors IF1 and IF2 synergistically remove peptidyl-tRNAs with short polypeptides from the P-site of translating Escherichia coli ribosomesaaEdited by J. Karn. Journal of Molecular Biology. 1998;281(2):241-52.
68. Heurgué-Hamard V, Karimi R, Mora L, MacDougall J, Leboeuf C, Grentzmann G, et al. Ribosome release factor RF4 and termination factor RF3 are involved in dissociation of peptidyl-tRNA from the ribosome. The EMBO Journal. 1998;17(3):808-16.
69. Karimi R, Ehrenberg M. Dissociation Rate of Cognate Peptidyl-tRNA from the A-Site of Hyper-Accurate and Error-Prone Ribosomes. European Journal of Biochemistry. 1994;226(2):355-60.
70. Gonzalez de Valdivia EI, Isaksson LA. Abortive translation caused by peptidyl-tRNA drop-off at NGG codons in the early coding region of mRNA. FEBS J. 2005;272(20):5306-16.
71. Singh NS, Varshney U. A physiological connection between tmRNA and peptidyl-tRNA hydrolase functions in Escherichia coli. Nucleic Acids Research. 2004;32(20):6028-37.
72. Dinçbas V, Heurgué-Hamard V, Buckingham RH, Karimi R, Ehrenberg M. Shutdown in protein synthesis due to the expression of mini-genes in bacteria 11Edited by D. E. Draper. Journal of Molecular Biology. 1999;291(4):745-59.
73. Heurgué-Hamard V, Dinçbas V, Buckingham RH, Ehrenberg M. Origins of minigene-dependent growth inhibition in bacterial cells. The EMBO Journal. 2000;19(11):2701-9.
74. Kang TJ, Suga H. Translation of a histone H3 tail as a model system for studying peptidyl-tRNA drop-off. FEBS Lett. 2011;585(14):2269-74.
75. Kazuta Y, Matsuura T, Ichihashi N, Yomo T. Synthesis of milligram quantities of proteins using a reconstituted in vitro protein synthesis system. J Biosci Bioeng. 2014;118(5):554-7.
76. Dickerman HW, Smith BC. Aminoacyl formylmethionyl transfer ribonucleic acid transformylase. Inhibitor studies and the reversal of the reaction. Biochemistry. 1970;9(5):1247-55.
77. Studer SM, Joseph S. Unfolding of mRNA Secondary Structure by the Bacterial Translation Initiation Complex. Molecular Cell. 2006;22(1):105-15.
78. Del Campo C, Bartholomäus A, Fedyunin I, Ignatova Z. Secondary Structure across the Bacterial Transcriptome Reveals Versatile Roles in mRNA Regulation and Function. PLOS Genetics. 2015;11(10):e1005613.
79. Pop C, Rouskin S, Ingolia NT, Han L, Phizicky EM, Weissman JS, et al. Causal signals between codon bias, mRNA structure, and the efficiency of translation and elongation. Molecular Systems Biology. 2014;10(12):770.
80. Chizzolini F, Forlin M, Yeh Martin N, Berloffa G, Cecchi D, Mansy SS. Cell-Free Translation Is More Variable than Transcription. ACS Synth Biol. 2017;6(4):638-47.

81. Castro-Roa D, Zenkin N. In vitro experimental system for analysis of transcription-translation coupling. Nucleic Acids Res. 2012;40(6):e45.
82. Iskakova MB, Szaflarski W, Dreyfus M, Remme J, Nierhaus KH. Troubleshooting coupled in vitro transcription-translation system derived from Escherichia coli cells: synthesis of high-yield fully active proteins. Nucleic Acids Res. 2006;34(19):e135.
83. Kazuta Y, Adachi J, Matsuura T, Ono N, Mori H, Yomo T. Comprehensive analysis of the effects of Escherichia coli ORFs on protein translation reaction. Mol Cell Proteomics. 2008;7(8):1530-40.
84. Antoun A, Pavlov MY, Lovmar M, Ehrenberg M. How Initiation Factors Maximize the Accuracy of tRNA Selection in Initiation of Bacterial Protein Synthesis. Molecular Cell. 2004;23(2):183-93.
85. Nawrot B, Sochacka E, Duchler M. tRNA structural and functional changes induced by oxidative stress. Cell Mol Life Sci. 2011;68(24):4023-32.
86. Sochacka E, Kraszewska K, Sochacki M, Sobczak M, Janicka M, Nawrot B. The 2-thiouridine unit in the RNA strand is desulfured predominantly to 4-pyrimidinone nucleoside under in vitro oxidative stress conditions. Chem Commun (Camb). 2011;47(17):4914-6.
87. Vivanco-Dominguez S, Bueno-Martinez J, Leon-Avila G, Iwakura N, Kaji A, Kaji H, et al. Protein synthesis factors (RF1, RF2, RF3, RRF, and tmRNA) and peptidyl-tRNA hydrolase rescue stalled ribosomes at sense codons. J Mol Biol. 2012;417(5):425-39.
88. Menninger JR, Mulholland MC, Stirewalt WS. Peptidyl-tRNA hydrolase and protein chain termination. Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis. 1970;217(2):496-511.
89. Ross BJ, Ian S. Halting a cellular production line: responses to ribosomal pausing during translation. Biology of the Cell. 2007;99(9):475-87.
90. Freischmidt A, Liss M, Wagner R, Kalbitzer HR, Horn G. RNA secondary structure and in vitro translation efficiency. Protein Expression and Purification. 2012;82(1):26-31.
91. Bhattacharyya S, Jacobs WM, Adkar BV, Yan J, Zhang W, Shakhnovich EI. Accessibility of the Shine-Dalgarno Sequence Dictates N-Terminal Codon Bias in E. coli. Mol Cell. 2018;70(5):894-905.
92. Li J, Gu L, Aach J, Church GM. Improved cell-free RNA and protein synthesis system. PLoS One. 2014;9(9):e106232.
93. Li J, Zhang C, Huang P, Kuru E, Forster-Benson ETC, Li T, et al. Dissecting limiting factors of the Protein synthesis Using Recombinant Elements (PURE) system. Translation. 2017;5(1):e1327006.
94. Jackson K, Kanamori T, Ueda T, Fan ZH. Protein synthesis yield increased 72 times in the cell-free PURE system. Integr Biol (Camb). 2014;6(8):781-8.
95. Peil L, Starosta AL, Lassak J, Atkinson GC, Virumae K, Spitzer M, et al. Distinct XPPX sequence motifs induce ribosome stalling, which is rescued by the translation elongation factor EF-P. Proc Natl Acad Sci U S A. 2013;110(38):15265-70.
96. Ude S, Lassak J, Starosta AL, Kraxenberger T, Wilson DN, Jung K. Translation Elongation Factor EF-P Alleviates Ribosome Stalling at Polyproline Stretches. Science. 2013;339(6115):82-5.
97. Arnold S, Siemann-Herzberg M, Schmid J, Reuss M. Model-based Inference of Gene Expression Dynamics from Sequence Information. 2005;100:89-179.