Computer simulation of a translational roadblock model - Draft

Gabin Rousseau

2024-05-21

Table of contents

[1 Abstract 1](#_Toc166024682)

[2 Introduction 2](#_Toc166024683)

[2.1 Translation initiation and mRNA scanning 2](#_Toc166024684)

[2.2 Ssd1 and translational control 3](#_Toc166024685)

[2.3 The importance of multiple Ssd1 binding sites 3](#_Toc166024686)

[2.4 TASEP as a translation modelling tool 4](#_Toc166024687)

[2.5 Aims and objectives 5](#_Toc166024688)

[3 Methods 5](#_Toc166024689)

[3.1 Writing the model: rini 5](#_Toc166024690)

[3.2 Density phase profile 7](#_Toc166024691)

[3.3 Corini: a second model for the cooperativity hypothesis 7](#_Toc166024692)

[3.4 Generating comparable data 8](#_Toc166024693)

[3.5 Exploring the parameter space 9](#_Toc166024694)

[4 Results 10](#_Toc166024695)

[4.1 Model functionality 10](#_Toc166024696)

[4.2 Comparing model fits to experimental data 10](#_Toc166024697)

[4.3 Exploring the parameter space and improving fits 12](#_Toc166024698)

[5 Discussion 14](#_Toc166024699)

[6 References 15](#_Toc166024700)

## 1 Abstract

Eukaryotic translation begings by inititation: the ribosome loads onto the 5’ end of the mRNA and scans its 5’UTR to find a start codon. While the scanning ribosome is sensitive to contextual cues as well as secondary structures in the mRNA sequence, RNA-binding proteins are also able to control translation. Ssd1 is a fungal protein that binds a specific motif in 5’UTRs. It is understood to inhibit protein production in cell wall biogenesis and was proposed to do so by directly blocking scanning ribosomes. We wrote a stochastic model on Python, rini, that belongs to the model family of TASEPs. After adjusting the basic framework of the TASEP with targeted, dynamic defects or roadblocks to simulate the proposed role of Ssd1, rini proved to function as expected of a TASEP. Between roadblock independence and positive cooperativity, only cooperativity could explain the experimental data available. Projecting the explored parameter space with randomised block binding probabilities further illustrated the difference between the hypotheses. Results remain preliminary despite highlighting Ssd1 cooperativity: parameters lack significance with real equivalents, and more should be tested. We recommend repeating the flow cytometry experiment with more Ssd1 sites and in different configurations to improve the model and predictions.

(2OO words)

## 2 Introduction

### 2.1 Translation initiation and mRNA scanning

Protein synthesis results from translation, where ribosomes read mRNAs to assemble amino-acids into specific peptide chains. In eukaryotes, a ribosomal small subunit (SSU) first interacts with 11 eukaryotic initiation factors (eIFs) to form a pre-initiation complex (PIC), and load onto the mRNA from the 5’ m7G cap (Dever and Pavitt, 2016). The PIC then scans the 5’ UTR of the mRNA towards its 3’ end until a start codon (AUG) is recognised, which triggers a series of reactions to release the eIFs and enable a large subunit (LSU) to form the 80S ribosome complex with the SSU, now primed to begin elongation by reading the translated region that follows (Kozak, 1986).

The scanning ribosome moves in the 5’-3’ direction with a footprint of 20-30nt, passing the single mRNA strand through the SSU’s mRNA channel until a start codon is detected at the recognition site where the methionyl initiator tRNA (Met-tRNAiMet) is held in place within the PIC by eIF1 (Hinnebusch, 2011). Secondary structures in the mRNA such as double-stranded hairpins would prevent the ribosome from advancing and can be melted by a set of helicases, mainly eIF4A, that benefit from the overall structure of the PIC and hydrolyse ATP to function (García-García *et al.*, 2015; Pisareva and Pisarev, 2016; Gupta *et al.*, 2018). The exact nature of the scanning motion appears to be the object of various models, but the consensus seems to forward Brownian motion as the key thermodynamic principle behind scanning. ATP hydrolisis energy expended by eIF4A likely supports translocation further, which remains directed towards the 3’ end thanks to the PIC’s structural characteristics and possibly dynamic interactions between eIF4 subunits dowstream of the PIC (Spirin, 2009; Marintchev *et al.*, 2009). In addition, the scanning complex is sensitive to context cues in the 5’ UTR that will affect recognition of AUG codons, and therefore which open reading frame (ORF) will be read. For example, preproinsulin mRNAs in vertebrates had significantly impaired initiation with mutations in the consensus optimal sequence 5′-GCCGCC(A/G)CCAUGG-3′ about the cognate start codon (Kozak, 1987). Therefore, how long scanning lasts, whether it is carried out succesfully and where it ends inherently rely on the 5’UTR sequence itself as well as its secondary structures. However, there exists a variety of other mechanisms which control translation and initiation.

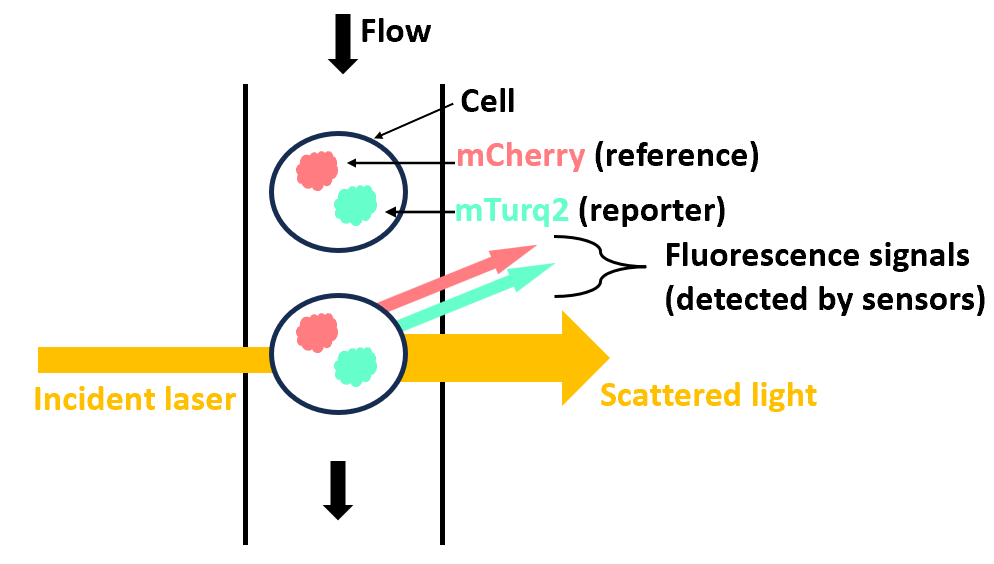
### 2.2 Ssd1 and translational control

RNA-binding proteins (RBPs) may control translation initiation in specific pathways. Ferritin is a ubiquitous protein able to carry iron Fe3+ ions within ferritin complexes. Its subunits’ translation can be suppressed by iron regulatory protein (IRPs) to control iron homeostasis. Ferritin mRNAs indeed contain an iron responsive element (IRE) which forms a stem-loop secondary structure on which IRPs can individually bind, sterically blocking loading of the PIC on the mRNA by interacting with the eIF4F complex located on the 5’ cap (Muckenthaler *et al.,* 1998).

Ssd1 is another RBP targeting bipartite binding motifs (18-24nt footprint) concentrated in the 5’UTRs of mRNAs involved in regulating fungal cell wall synthesis (Uesono *et al.*, 1997; Hogan *et al.*, 2008; Bayne *et al.*, 2022). Ssd1 is an RNAse II homolog for which access to a would be inner catalytic site is blocked due to loop structures; instead, Ssd1 binds its motifs with high affinity *via* specific positively charged regions on the outer face of both of its -barrel cold shock domains (CSDs; Bayne *et al.*, 2022). The binding motif thus contains two different consensus binding sequences separated by a short gap (4-10nt) that are both key to high affinity binding of Ssd1, and unpublished data from Edward Wallace’s lab suggested multiple Ssd1 motifs in single 5’UTRs appear in quick succession. What is Ssd1’s mode of action as an RBP? The kinase Cbk1 was identified as an Ssd1 inhibitor in a reporter expression assay and rescued normal cell phenotypes, which suggested Ssd1 inhibits translation (Jansen *et al.*, 2009). Due to its properties as a highly specific RBP that strongly binds its target motif, Ssd1 has been proposed to function as a roadblock, sterically preventing PICs from progressing through the 5’UTR (Bayne *et al.,* 2022).

### 2.3 The importance of multiple Ssd1 binding sites

In 2021, Edward Wallace analysed results from a flow cytometry experiment in *S. cerevisiae* that measured the production of a fluorescent reporter, mTurq2, which had a variable number of Ssd1 binding motifs in its mRNA’s 5’UTR, from 0 to 2 (Figure 1). The reporter’s signal was normalised with the signal of a reference protein (mCherry) to use the flow cytometry measurements as proxies for protein production efficiency. The data from the experiment showed a small decrease in reporter expression in the presence of 1 binding motifs, whereas two motifs were associated with a much stronger effect: the loss of nearly a third of the standard protein production without Ssd1 motifs. With these results in mind, what experiments could help makes sense of this pattern to unveil Ssd1’s mode of action? This raised the idea of designing a computational model to simulate translational roadblocking, with respect to the proposed mode of action, and to help generate expectations for follow-up research.



**Figure 1. Schematic of Edward Wallace’s flow cytometry experiment** - The signal of both proteins were captured in separate channels. Organism: Saccharomyces Cerevisiae

### 2.4 TASEP as a translation modelling tool

For this model, we need to take into consideration the essential elements involved: if Ssd1 controls translation initiation by blocking ribosomes while they scan an mRNA’s 5’UTR, then our system contains three essential elements. We need an object (ribosome) to follow a one-directional trajectory (mRNA) with a start (5’cap) and an end (AUG codon). Finally, a secondary object (Ssd1) must be able to stand in the way of the first dynamically. If ribosomes reach the end of the 5’UTR, then we could assume succesful protein production and track these exits. Fortunately, these conditions are well adapted to a type of stochastic, one-directional and open-ended diffuse transport model referred to as TASEP (Totally Asymmetric Simple Exclusion Process; Popkov and Schütz, 1999; Bonnin *et al.,* 2017).

TASEPs already have a history with translation modelling, each finding their own individual variations. For example, one TASEP was designed to study the effects of ribosome drop-off mid-way through translation, adding the possibility of having ribosomes (referred to as ‘particles’ in TASEPs) suddenly leaving the mRNA (the ‘lattice’) before reaching the end; another interested in initiation simulated the interaction between scanning PICs and elongating ribosomes that formed due to an upstream ORF in the 5’UTR, involving dynamic and random changes in the ribosomes’ behaviour (Bonnin *et al.,* 2017; Andreev *et al.,* 2018). However, these examples only model specific ribosome behaviours and dynamics that are not directly featured in the basic TASEP concept, and lack a secondary object that could correspond to another RNA-binding element. In 2019, Warclaw *et al.,* suggested an extension to TASEP involving “site-wise dynamic disorder” (2019). Disorder in the lattice-particle system would be created by having obstacles (‘defects’) dynamically and randomly attach to sites on the lattice, blocking the way until they detach. Different sets of rules were suggested for obstacle behaviour, which included constrained defect dynamics. Constrained defect dynamics not only allow defects to block particles, but also particles to prevent defects from binding the site they stand on, which would better reflect the behaviour of physical objects like ribosomes and Ssd1.

### 2.5 Aims and objectives

This project aimed to simulate translational control by roadblock RBPs with the development of an in-house TASEP model with defects, specifically intended to help generate predictions regarding the effect of Ssd1 landscapes on translation initiation. We wrote our model, named ‘rini’ (roadblock in initiation), in python 3, and gathered the necessary code for the model as well as functions that use the model’s output in a Jupyter notebook. We tested the model’s viability as a TASEP by plotting its density phase diagram. A complementary version, corini, was also written to account for the hypothesis that Ssd1 units undergo positive cooperativity based on the flow cytometry results, as opposed to total roadblock independence. These results served as our biological reference to fit model output to. After testing a range of roadblock (defect) binding probabilities, we plotted the best fits for both rini and corini against the flow cytometry data as well as generating a parameter space to identify which parameter sets matched the available quantitative data the most.

## 3 Methods

### 3.1 Writing the model: rini

We wrote rini (short for “roadblock in initiation”) in Python 3 to model the effect of dynamic roadblocks on translation initiation efficiency. The model and all functions that rely on its output were run in a Jupyter notebook (see Appendix for the final version used in this report).

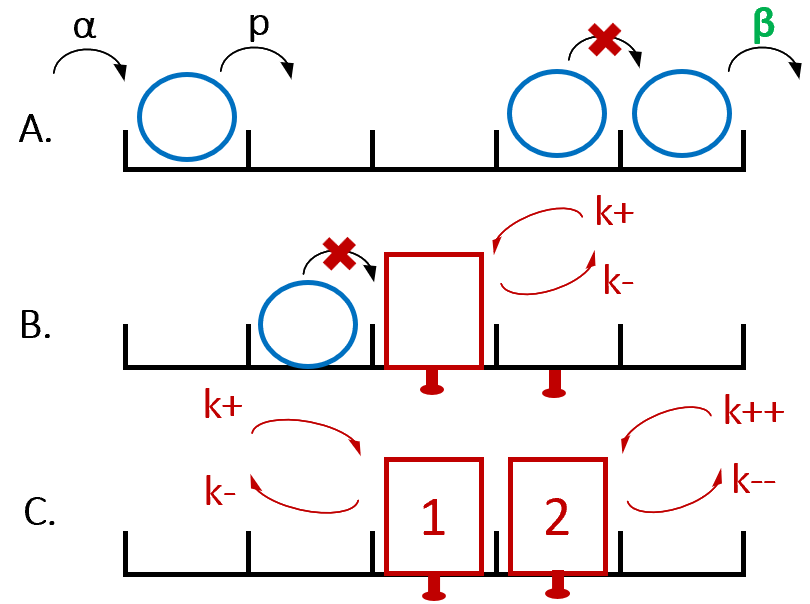
Rini is a TASEP model in which we consider a one-dimensional lattice of length L, the length corresponding to the total number of sites that can be occupied. Particles may enter the lattice, *i.e.* load onto the first site, with a probability of loading . Any particle in the lattice with an unoccupied site ahead of it may hop forward onto that site with a probability of progression (set to the constant value of 1 in this project). Finally, a particle that reaches the end of the lattice can exit the lattice with a probability of unloading (Figure 1A). Entrance and exit probabilities may range from 0 to 1. In translation initiation, particles would correspond to scanning ribosomes, and consequently the lattice would be the mRNA. When a particle exits the lattice, we assume the PIC to have detected a cognate start codon, flagging the start of elongation, which is assumed to result in succesful protein production.

For the model to move forward, the user needs to set a time limit t. rini will create time steps until is reached, starting from runtime . Each time step revisits the same series of key events:

* Checking all possible actions on the lattice and recording the corresponding probabilities as a separate set of propensities , *e.g.* the probability of an action such as particle loading to occur at the current time step.
* Calculating , the sum of the propensities.
* Selecting a random number from an exponential distribution that scales on R to calculate how long the current time step lasts.
* Selecting a random number from a uniform distribution within the bounds (0, 1) that scales on R to determine which of the actions will be executed. This is done by tracing which action appearing in the R distribution corresponds to r2. For example, if in that distribution and the first two propensities and each weigh 0.1, then : action 2 will be performed.
* Executing the chosen action and making appropriate changes to the system as a whole (*i.e.* the lattice, propensities).

This enables the model to function stochastically in semi-continuous time, with respect to the Gillespie algorithm formalism (Gillespie, 1977).

In addition to particles traveling from start to end, rini features targeted, dynamic roadblocks that represent Ssd1. Specific sites (: lattice index of the site) can be defined as block binding targets. When such a site is unoccupied, a block may bind with a probability . A bound block may unbind at a probability . When bound by a block, a site is considered “occupied” the same as when a particle is present, blocking passage (Figure 1B).



**Figure 2. rini concept schematics: from simple TASEP to dynamic defect targetability** - A: Key features of a TASEP model, particles (blue) stochastically enter a lattice containing sites (black), and hop from one site to the next until reaching the exit. Movements are associated with set probabilities. Particles can only move in one direction and can only proceed to the next site if it is unoccupied. B: In rini, sites can be flagged as targets for roadblock binding (also known as “defects”, red). Every flagged site has a probability to get bound or unbound by a roadblock. In the first version, these values are fixed and independent from the roadblock landscape. C: the cooperativity version corini features an additional rule that adapts block dynamics according to the roadblock landscape. The probabilities of block binding and unbinding are modified in the presence of a neighbouring block in favour of the blocks staying on the lattice.

### 3.2 Density phase profile

To observe the viability of rini as a TASEP model, we investigated its phase density profile with the rini\_AvB function (Alpha *vs.* Beta, see Appendix). AvB creates 10 independent rini runs with random combinations of and (values vary between 0.1 and 1); it returns the phase diagram, plotting against , alongside the average occupation density against the expected bulk density for the corresponding density phase (Table 1; Connin *et al.,* 2017). This randomisation may occasionally lead some points to fall into a less predictable phase: Shock Phase, where = and both probabilities are below 0.5; in that case, densities are unpredictable at specific points in time, interchangably high or low along sections of the lattice. In TASEP models, low density (LD) means the lattice will on average be occupied by few particles as limits particle entry. High density (HD) represents the opposite scenario, where significantly limits particle exits, yielding high occupation densities. Otherwise, a TASEP model is said to undergo maximum current (MC) when both and are high enough for the particle flow, and thus occupation density, to be capped by the progression probability rather than entrance or exit.

We are considering the occupation density as , with as the number of occupied sites.

We ran AvB with and with no block sites and selected an output featuring at least 3 points per density phase.

Table 1. Main density phases expected in TASEP models where p = 1.

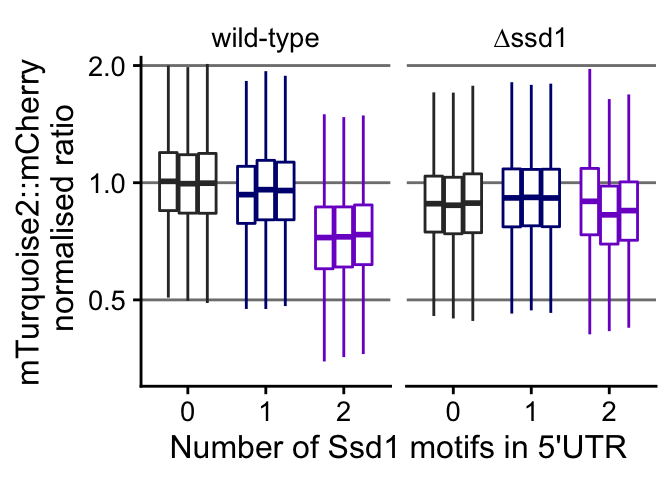
| Density Phase | Conditions | Expected density |
| --- | --- | --- |
| Low Density | , |  |
| Maximum Current | , |  |
| High Density | , |  |

### 3.3 Corini: a second model for the cooperativity hypothesis

After gaining confidence in the model, we observed in the original cell cytometry data that protein production appeared to be more significantly impacted by the presence of 2 Ssd1 binding sites in the reporter protein’s 5’UTR, compared to the effect linked to the presence of 1 site. In other words, experimental data displays a non-linear change in protein production from 0 to 2 Ssd1 binding sites. This difference approximately goes from a drop in protein production efficiency of 4.1% with 1 site, to 25.9% with 2 sites (Figure 3). The nearly 5-fold effect increase would suggest that the presence of adjacent binding sites for Ssd1 multiplies the proteins inhibitory effect on translation. If Ssd1 acts as a direct roadblock for the pre-initiation ribosome, then cooperativity between Ssd1 units could explain this pattern.

However, rini alone does not emulate cooperativity between roadblocks, as and remain strictly constant for any blockable site, independently of the surrounding roadblock landscape. This being likely insufficient to explain the experimental data, rini became the ‘independence’ model. We wrote an alternate’cooperativity’ version of the program, corini, which assumes the presence of a block may improve block binding or decrease block unbinding in its direct neighbourhood, creating simple cooperativity between adjacent roadblock units. If cooperativity becomes active at a block site and at any point in time, then the rates for block binding or unbinding, where applicable, are modified by fixed values in favour of the roadblocks: is added to , and is substracted from (read: “coop-plus” and “coop-minus”; Figure 1C).

These two model versions enabled the testing of two hypotheses, roadblock independence or cooperativity.



**Figure 3. Key flow cytometry data showing the relationship between Ssd1 binding sites and protein production efficiency** - shows data from 3 replicates; mTurquoise2 is the reporter signal associated with up to 2 Ssd1 binding sites in its mRNA’s 5’UTR, it’s signal serves as the proxy for protein production, normalised by the reference mCherry signal and the WT median mean of the signal ratio at 0 binding sites; if we average replicate data, the WT standardised median ratio (protein production efficiency) with 1 and 2 binding sites are respectively 0.959 and 0.741

### 3.4 Generating comparable data

For the model to contribute to further experiments and generate hypotheses on Ssd1 function, we explored fitting the model to the experimental data available. We have data about the link between block sites and protein production from Edward Wallace’s flow cytometry experiment, and chose the WT standardised median signal ratios at 1 and 2 block sites as our experimental values of references (0.959 and 0.741 respectively, Figure 1).

With these values, we can assess how each model version compares to biological data. From a simulation, we can extract the particle exit rate , which we define as the total number of particles that have exited the lattice over a run, divided by . Exiting particles represent ribosomes that succesfully complete initiation and proceed to elongation, which we assume results in protein production. So, we choose *e* to be the initial proxy for the raw protein production. However, we want a comparison with medians of standardised values. For a parameter set , we calculate median values over sampled simulations using this parameter set for an appropriate range of block sites (from 0 to 2). If we use the median at 0 block sites as our model’s standard value, we can divide *e* by this standard to get standardised exit rates , representation *initiation efficiency*, and plot the data as a boxplot to see the corresponding medians alongside their experimental counterpart.

We wrote the function rini\_INIvBLOCK to apply this method on excellent parameter sets we established for each model version (Table 2; Appendix).

Table 2. Key parameter set details used to generate exemplar fits for each model version

| Model | t | n | L |  |  | p | k\_on | k\_off | coop\_p | coop\_m | Targetables sites (first and second added) |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| rini | 1000 | 10 | 5 | 0.75 | 0.75 | 1 | 0.227 | 0.5 | 0.4 | 0.4 | 2 and 3 |
| corini | 1000 | 10 | 5 | 0.75 | 0.75 | 1 | 0.043 | 0.5 | 0.4 | 0.4 | 2 and 3 |

### 3.5 Exploring the parameter space

To understand how well each model is able to explain biological results and what parameters matter the most, we began exploring the model’s input parameter space with the function rini\_paramscore (Appendix).

Exploring a parameter space requires analysing many semi-randomised possibilities and scoring each parameter set based on how close the corresponding data is to the experiment. We tested parameter sets keeping the default values from Table 2 while randomising from a uniform distribution ranging from 0 to 1 (sampling for median calculation: ).

To score each , we calculate three values:

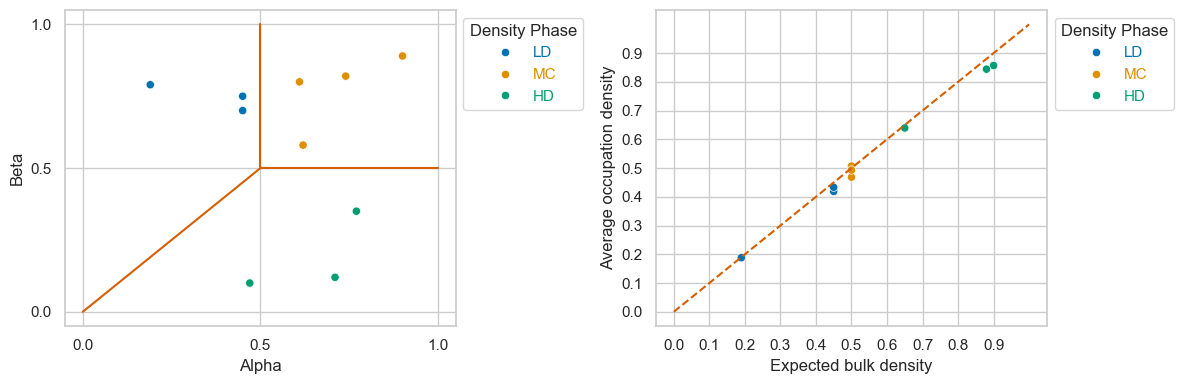
1. (score1) is the difference (or raw error) between the model’s median and the experimental median at 1 site: .
2. is the equivalent at 2 sites.
3. is the quadratic mean of s1 and s2: .

S scales with the difference between s1 and s2 to give an overall score of how close we are to biological data. The closer and are both to 0, the better the fit. Moreover, the sign of and gives us an indication of whether the model’s median output is above or below the expectation. While S is the main score for a model fit, we also plotted and against one another to visualise the parameter space.

## 4 Results

### 4.1 Model functionality

In our assessment of the model’s viability, we rely on phase density profiling, for which we generated 10 independent runs of rini, featured as points in Figure 4. 3 points were expected to be under low density (LD), another 3 in high density (HD), and finally 4 in maximum current (MC). Their expected densities ranged from less than 0.2, to nearly 0.9. The average occupation density calculated over time for each point consistently remains very close to the steady-stade bulk density expected from the associated density phases in a regular TASEP model, and don’t deviate more than about 5% away fom the expected value (Table 1, Figure 4); the point expected to have the highest density, nearly 0.9, yielded the biggest error of 5%, however bigger deviations are not necessarily tied to the highest densities, as a point in LD and another in MC are also noticeably away from the segmented line that represents result-expectation equivalence.



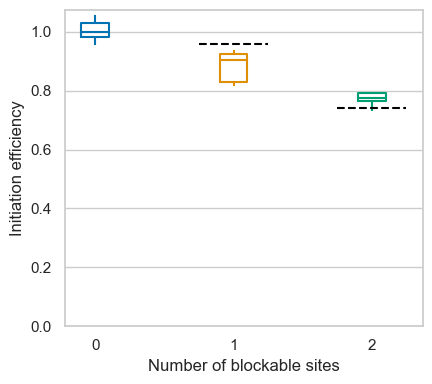
**Figure 4. Density phase profiling of rini** - t=1000, L=100, n=10; Left: density phases expected of each run according to their randomised alpha and beta parameters (LD: Low density, MC: Maximum Current; HD: High Density); Right: average occupation density over each run against the densities expected from density phases (Table 1) \_

### 4.2 Comparing model fits to experimental data

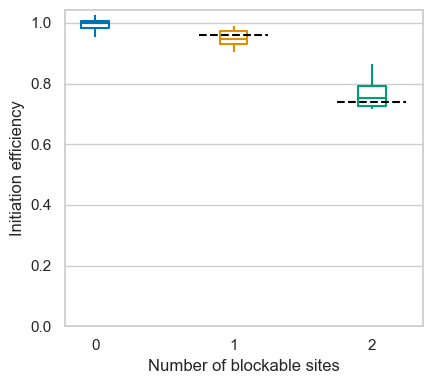
After inspecting whether the model can behave as expected at different density phases, we generated data to be compared with our experiment of reference, and plotted result comparisons for both rini and corini in Figures 5 and 6 respectively. Both Figures display excellent fits from the model versions that resulted from changing the block binding rate .

Adding a second blockable site in rini yielded a seemingly linear decrease in initiation efficiency, doubling the effect of a single blockable site. The data found little to no overlap with the experimental medians and with its linearity failed to capture the trend observed in the flow cytometry experiment; with 1 site, the data undercuts the expectation whereas the model stands above the experimental median with 2 sites.

The corini version of the model was able to produce output that stays consistenly close to biological data (Figure 6); the experimental medians overlap well with the data and remain close to the model medians. Alongside this similarity between the results, we notice that the model follows the same non-linear trend found in the experiment when roadblock cooperativity is taken into consideration.



**Figure 5. Example of an excellent fit of the independence model to experimental data** - Black segmented lines: experimental medians; n=10

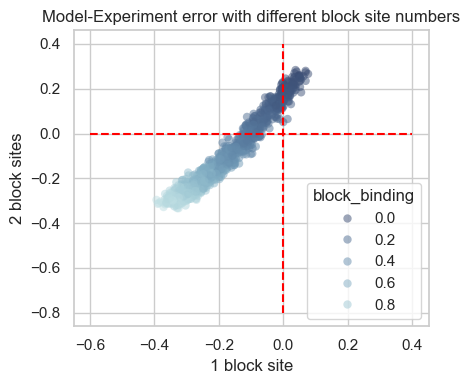


**Figure 6. Example of an excellent fit of the cooperativity model to experimental data** - Black segmented lines: experimental medians; n=10

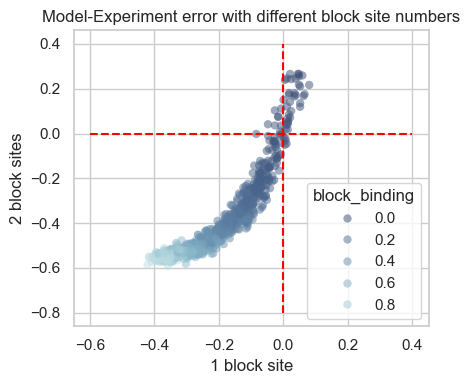
### 4.3 Exploring the parameter space and improving fits

To test a wide-range of parameter sets and observe how well the results fit biological data, we began preparing a function to explore the parameter space and plotted results for 1000 different data sets where we randomly varied the block binding probability in rini (Figure 5) and corini (Figure 6). In both versions, reducing increased initiation efficiency, as seen in the trend going towards positive raw errors. However, the independent model never crossed the ideal (0,0) coordinates, the closest point standing at the coordinates (-0.0394, 0.0177; same as in Figure 3). On the other hand, corini is seemingly able to reach these coordinates, with its best point having the coordinates (-0.0032, 0.0039) and its score being 10-fold lower than the best point from rini (Figure 6; same as in Figure 4).

A key difference between both models is the relationship between and , the results at 1 site and those at 2 sites. For both rini and corini, ranges between -0.4 and about 0.05. The skew towards negatives associated with better block binding shows that for most block binding probabilities in the default parameter set used, block binding would be too strong to fit the biological data. As block binding decreases in the independence model, which makes the model’s median increase closer to the experiment’s at 1 site, the raw error also increases linearly at 2 sites (Figure 5). In the cooperativity model, the 2-blocks data increases following an exponential trend alongside 1-block data, rather than linearly. Indeed, when block binding becomes especially low, cooperativity has an increasingly important effect in the model’s dynamics as it enables strong binding when multiple blocks are present, even though blocks otherwise rarely bind the lattice alone.



**Figure 7. Model-Experiment errors over 1000 randomised parameter sets (independence model)** - Red segmented lines highlight axes with a 0 coordinate; N=1000, n=10



**Figure 8. Model-Experiment errors over 1000 randomised parameter sets (cooperativity model)** - Red segmented lines highlight axes with a 0 coordinate; N=1000, n=10

## 5 Discussion

Rini is a model that aimed to perform simulations that efficiently emulate the essential dynamics expected of the system mRNA-ribosome-Ssd1, and simulate the effect on protein production seen with multiple Ssd1 binding motifs (Figure 3). The density phase profiling step helped build confidence that the model fundamentally behaved as expected of a model of its kind (Popkov and Schütz, 1999), and the values for t and L chosen were enough to mitigate size effects and observe consistent behaviour within a 95% confidence interval. The last simulation results seem to indicate the cooperativity version may be the only way to explain experimental data with the rini TASEP, given that the best fits for both the independence and cooperativity hypotheses respectively contrast a linear effect in adding block sites with close fits to experimental medians (Figures 3 and 4). Defining a rule for roadblock cooperativity succesfully changes block dynamics to explain these initial observations with appropriate parameter sets (Figure 8).

Indeed, we managed to see good simulations upon randomising among a set of default parameters that we selected to keep the model efficient and minimalist (L), while assuming Maximum Current. In these conditions, simulations were most succesful when the base binding affinity of blocks for their sites was very low (less than 0.05), paired to a relatively high unbinding rate and strong affinity improvements in a cooperativity context. This only represents one type of parameter set with which the model can accurately simulate biological data, and is already questionable since Ssd1 is supposed to bind its target with high affinity. Initiation is expected to be rate-limiting in translation, and the configuration of the 5’ UTR sequence (*e.g.,* likelihood to form secondary structures) seems to play an important part (Cambray *et al.*, 2018). Therefore, maximum current appears to be an interesting assumption, where the key limiting rate would be . However, the difference in specific parameter values and the range of randomisation chosen to explore the parameter space lack a clear relationship with the reality we aim to simulate; time is unique to the model and specific rates lie on the arbitrary range of (0,1). Therefore, the observations made in the project essentially highlight the relevance of cooperativity, but how could we harness rini to make predictions on the relevance of each parameter to obtain optimal simulations, and what would that mean biologically? We recommend conducting further parameter space explorations that focus on testing multiple parameters at once, and most importantly to isolate parameter value ranges that would be nonsense biologically, such as excessively low values. Evaluating the significance of time and probability values based on real ribosome scanning rates and effective particle progression rates would be the starting point of these efforts.

Yet, we have now gained interest in the hypothesis that Ssd1 units act as roadblocks capable of positive cooperativity, possibly through protein-protein interaction if proximity is indeed essential. Although we focused on modeling roadblock dynamics for up to 2 binding sites, the model would be able to generate data beyond just 2 sites and help predict results for various scenarios. However, to generate data for 3 block sites or more requires a decision on how cooperativity effects are calculated. In other words, the current rini simply detects cooperativity and modifies block dynamics by set values, and . This might create weaknesses in the model because cooperative RNA-binding elements that interact with one another are likely to see their overall affinity for RNA increase as the number of elements bound increases (Stitzinger *et al.*, 2023).

Knowing this, what would be the next step to elucidate the role of Ssd1 in translational control? While we think that Ssd1 may indeed undergo positive cooperativity when two motifs motifs are juxtaposed in a 5’UTR, it is difficult to predict the trend we might observe with a third site, which bears consequences for the model’s algorithm. Experimental data that would contribute to refining the model as it is now could therefore be a repeat of the Wallace lab’s flow cytometry experiment from 2021; specifically, we expect data for 3 to 4 sites as well as the distances between each site to shed light on the dynamics involved.

(Appendix where I share the github link to the model)

Excluding figures, titles and references, this report contains 4334 words (word limit: 6000)

## 6 References

D. E. Andreev et al., TASEP modelling provides a parsimonious explanation for the ability of a single uORF to derepress translation during the integrated stress response. Elife 7, (2018).

R. A. Bayne et al., Yeast Ssd1 is a non-enzymatic member of the RNase II family with an alternative RNA recognition site. Nucleic Acids Research 50, 2923-2937 (2022).

P. Bonnin, N. Kern, N. T. Young, I. Stansfield, M. C. Romano, Novel mRNA-specific effects of ribosome drop-off on translation rate and polysome profile. Plos Computational Biology 13, (2017).

G. Cambray, J. C. Guimaraes, A. P. Arkin, Evaluation of 244,000 synthetic sequences reveals design principles to optimize translation in Escherichia coli. Nature Biotechnology 36, 1005-+ (2018).

C. García-García, K. L. Frieda, K. Feoktistova, C. S. Fraser, S. M. Block, Factor-dependent processivity in human eIF4A DEAD-box helicase. Science 348, 1486-1488 (2015).

D. T. Gillespie, EXACT STOCHASTIC SIMULATION OF COUPLED CHEMICAL-REACTIONS. Journal of Physical Chemistry 81, 2340-2361 (1977).

N. Gupta, J. R. Lorsch, A. G. Hinnebusch, Yeast Ded1 promotes 48S translation pre-initiation complex assembly in an mRNA-specific and eIF4F-dependent manner. Elife 7, (2018).

A. G. Hinnebusch, Molecular Mechanism of Scanning and Start Codon Selection in Eukaryotes. Microbiology and Molecular Biology Reviews 75, 434-467 (2011).

M. Kozak, POINT MUTATIONS DEFINE A SEQUENCE FLANKING THE AUG INITIATOR CODON THAT MODULATES TRANSLATION BY EUKARYOTIC RIBOSOMES. Cell 44, 283-292 (1986).

M. Kozak, AN ANALYSIS OF 5’-NONCODING SEQUENCES FROM 699 VERTEBRATE MESSENGER-RNAS. Nucleic Acids Research 15, 8125-8148 (1987).

A. Marintchev et al., Topology and Regulation of the Human eIF4A/4G/4H Helicase Complex in Translation Initiation. Cell 136, 447-460 (2009).

M. Muckenthaler, N. K. Gray, M. W. Hentze, IRP-1 binding to ferritin mRNA prevents the recruitment of the small ribosomal subunit by the cap-binding complex elF4F. Molecular Cell 2, 383-388 (1998).

A. S. Spirin, How Does a Scanning Ribosomal Particle Move along the 5′-Untranslated Region of Eukaryotic mRNA? Brownian Ratchet Model. Biochemistry 48, 10688-10692 (2009).

S. H. Stitzinger, S. Sohrabi-Jahromi, J. Söding, Cooperativity boosts affinity and specificity of proteins with multiple RNA-binding domains. Nar Genomics and Bioinformatics 5, (2023).

B. Waclaw, J. Cholewa-Waclaw, P. Greulich, Totally asymmetric exclusion process with site-wise dynamic disorder. Journal of Physics a-Mathematical and Theoretical 52, (2019).