Computer simulation of a translational roadblock model - Parameter Space Presentation

Gabin Rousseau

2024-04-03

Table of contents

## 1 Why

### 1.1 Two models, two hypotheses

* After establishing that our model to simulate the effect of roadblocking on translation initiation functions as expected, we observed that it would likely be insufficient to explain experimental data, and therefore wrote two distinct versions that use the same formalism and produce the same output, but differ in their roadblock dynamics. Thus, we named the original program the ‘independent’ model (rini), which assumes roadblock binding and unbinding occur independently from other roadblock units, thus leaving the corresponding rates constant throughout the simulation regardless of the roadblock landscape. On the other hand, the new ‘cooperative’ model (corini) assumes the presence of a block may improve block binding or decrease block unbinding in its direct neighbourhood, creating cooperativity between adjacent roadblock units; knowing this, if cooperativity is active at any block site and any point in time, *then the rates for block binding or unbinding where applicable are modified by fixed values in favour of the roadblocks*.

### 1.2 To explain early results

* In 2021, Edward collected flow cytometry data showing the production efficiency of a protein reporter with 0 to 2 Ssd1 sites in its mRNA’s 5’UTR. We can emulate this in the model since we can calculate particle exit rates directly to serve as proxies for protein production. For the model to contribute to further experiments and help generate hypotheses on Ssd1 function, we should therefore explore fitting the model to the experimental data available.

## 2 What

* We first extracted model data in a form comparable with experimental data, before writing a function to produce a parameter space where we can sample from a wide range of semi-randomised parameter sets to measure how close the model’s results are to the experiment’s.

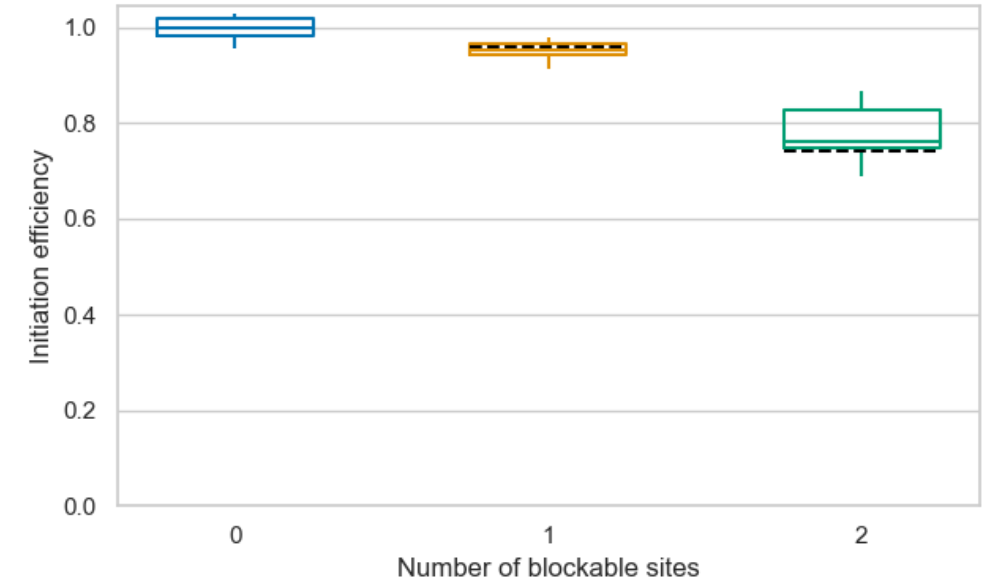
## 3 How

### 3.1 To get proxy data

* We have data about the link between block sites and protein production from Edward’s flow cytometry experiment. We chose to make a proxy comparable to the medians of standardised protein production estimates at 1 and 2 block sites. If the standard is the median protein production at 0 site (standardised value of 1), the experimental median for 1 site is 0.959, and 0.741 for 2 sites. We therefore have two simple values that we can use to assess how each model version is able to match biological data and what that could mean for the underlying biology.
* From a simulation, we can extract the particle exit rate *e*. 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 median standardised values. For a parameter set , we can calculate median *e* values over *n* 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 (Figures 1 and 2).
* Figures 1 and 2 show excellent fits from both rini and corini obtained after randomising the block-binding rate. We were interested in writing corini because the experimental has a non-linear change in protein production from 0 to 2 sites. Indeed, Figure 1 highlights the key issue with the original rini model as although it reaches results that encompass experimental data at 2 sites, it seemingly fails to capture the noticeable drop in protein production between 1 and 2 sites, instead displaying a linear drop that noticeably undercuts experimental data at 1 site. On the other hand, the cooperative model was able to closely match experimental data (Figure 2).
* This observation encouraged us to write a parameter space function that allows us to test a wide range of parameter sets and measure which are associated with the best fits.



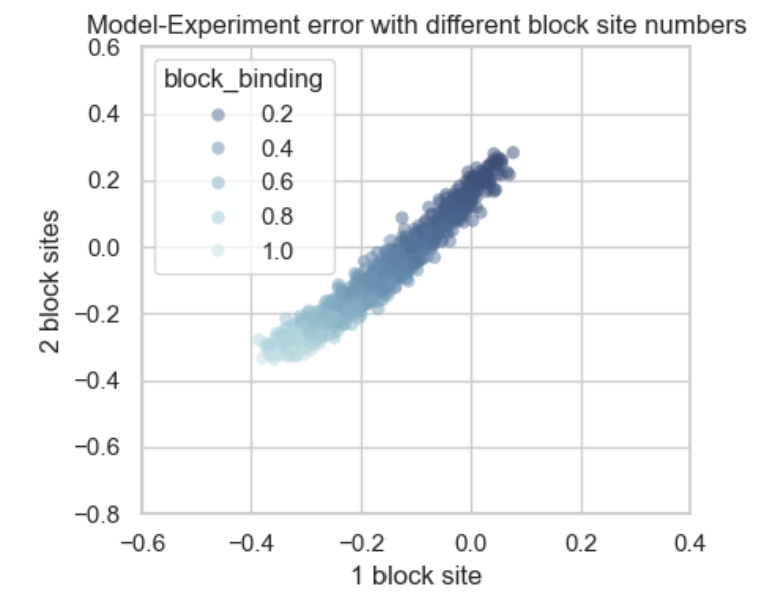
***Figure 1. Example of an excellent fit of the independent model to experimental data*** *- Black segmented lines: experimental medians; n=10*



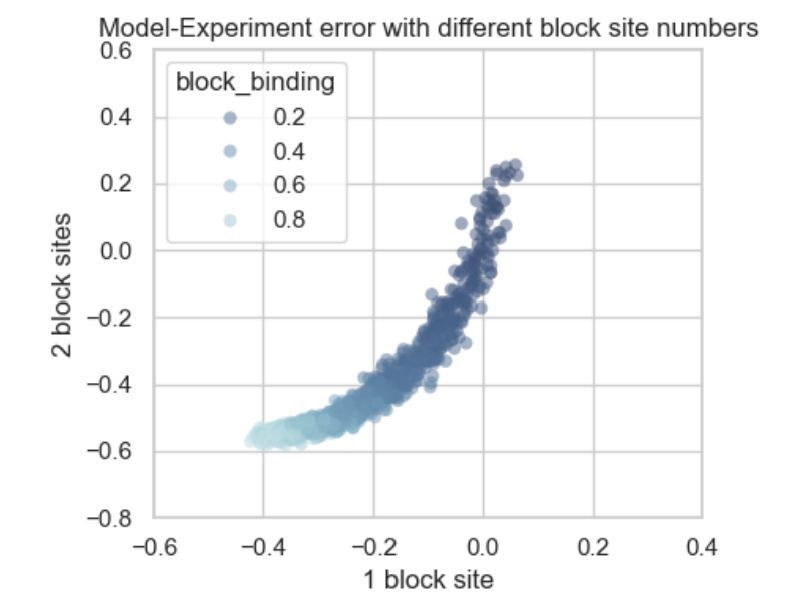
***Figure 2. Example of an excellent fit of the cooperative model to experimental data*** *- Black segmented lines: experimental medians; n=10*

### 3.2 To use a parameter space

* Quickly figuring out which parameter sets create the best fits requires a way to score each parameter set based on how close the associated data is to biological data overall. To do this, we calculate three values:
  1. s1 (score1) is the difference (or raw error) between the model’s median and the experimental median at 1 site: .
  2. s2 is the equivalent at 2 sites.
  3. S 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 s1 and s2 are both to 0, the better the fit. Moreover, the sign of s1 and s2 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, s1 and s2 can be conveniently plotted against one another to visualise the parameter space (Figures 3 and 4).
* Our initial parameter space in Figures 3 and 4 tests N=1000 parameter sets with randomised block-binding rates in rini (Figure 3) and corini (Figure 4). In both versions, reducing the block-binding rate increased initiation efficiency (as seen in the trend going towards positive raw errors). While the independent model never crosses the ideal (0, 0) coordinates, with its best point having the coordinates (-0.0394, 0.0177), corini is seemingly able to do so in this test, suggesting the potential for cooperativity as a key element to explaining the biological data. The best point in the cooperative model has the coordinates (-0.0032, 0.0039), with its S score being 10-fold lower than the best point’s in Figure 3.



***Figure 3. Model-Experiment errors over 1000 randomised parameter sets (independent model)*** *- N=1000, n=10*



***Figure 4. Model-Experiment errors over 1000 randomised parameter sets (cooperative model)*** *- N=1000, n=10*