Physical Modeling of Translation with a Ballistic Model: Application to the Estimation of the Kinetic Parameters from Ribosome Sequencing Experiments

Carole Chevalier, Jerome Dorignac, Yahaya Ibrahim, Armelle Choquet, Alexandre David, Julie Ripoll, Eric Rivals, Frederic Geniet, Nils-Ole Walliser, John Palmeri, et al.

PHYSICAL MODELING OF TRANSLATION WITH A BALLISTIC MODEL: APPLICATION TO THE ESTIMATION OF THE KINETIC PARAMETERS FROM RIBOSOME SEQUENCING EXPERIMENTS

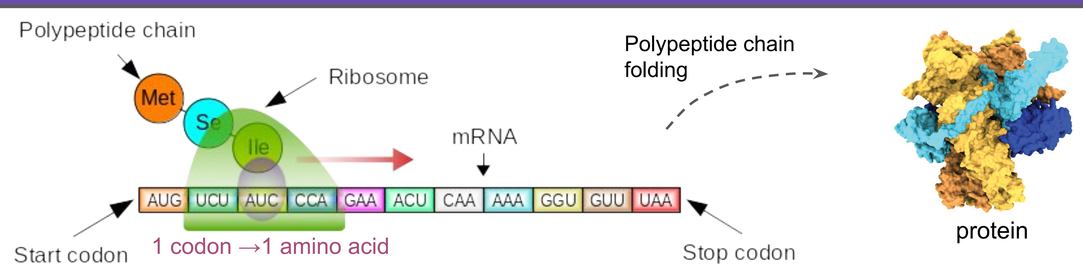
Carole Chevalier[1], Jérôme Dorignac[1], Yahaya Ibrahim[4], Armelle Choquet[3], Alexandre David[3], Julie Ripoll[2], Eric Rivals[2], Frédéric Geniet[1], Nils-Ole Walliser[1], John Palmeri[1], Andrea Parmeggiani[1] and Jean-Charles Walter[1]

[1]Laboratoire Charles Coulomb (L2C) [2]Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier (LIRMM) [3]Institut de Génomique Fonctionnelle (IGF)
[4]Department of Physics, Faculty of Natural and Applied Sciences, Umaru Musa Yar'adua University, Katsina, Nigeria

Key words : mRNA, ribosome, translation, deep sequencing, transport, kinetics, initiation rate, elongation rate, fit, minimisation

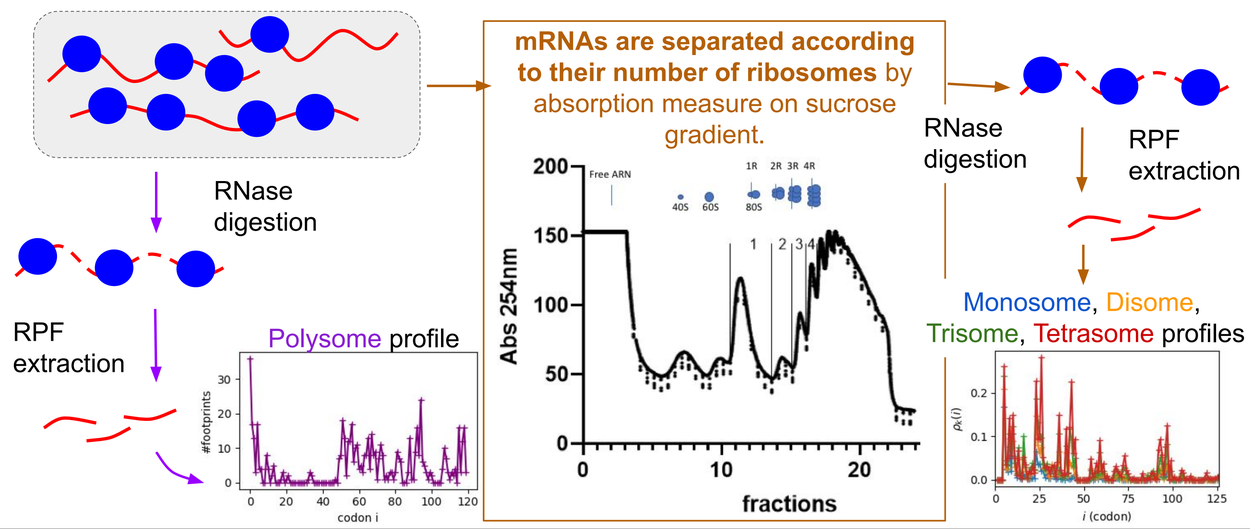
Abstract : Gene expression consists in the synthesis of proteins from the information encoded on DNA. One of the two main steps of gene expression is the translation of messenger RNA (mRNA) into polypeptide sequences of amino acids. Here, by taking into account mRNA degradation, we model the motion of ribosomes along mRNA with a ballistic model where particles advance along a filament without excluded volume interactions. Unidirectional models of transport have previously been used to fit the average density of ribosomes obtained by the experimental ribosome sequencing (Ribo-Seq) technique. In this case an inverse fit gives access to the kinetic rates: the position-dependent speeds and the entry rate of ribosomes into mRNA. The degradation rate is not, however, accounted for and experimental data from different experiments are needed to have enough parameters for the fit. Here, we propose an entirely novel experimental setup and theoretical framework consisting in splitting the mRNAs into categories depending on the number of ribosomes from one to four. We analytically solve the ballistic model for a fixed number of ribosomes per mRNA, study the different regimes of degradation, and propose a criterion for the quality of the inverse fit. The proposed method provides a high sensitivity to the mRNA degradation rate. The additional equations coming from using the monosome (single ribosome) and polysome (arbitrary number) Ribo-Seq profiles enables us to determine all the kinetic rates in terms of the experimentally accessible mRNA degradation rate.

1. Translation as a transport phenomenon

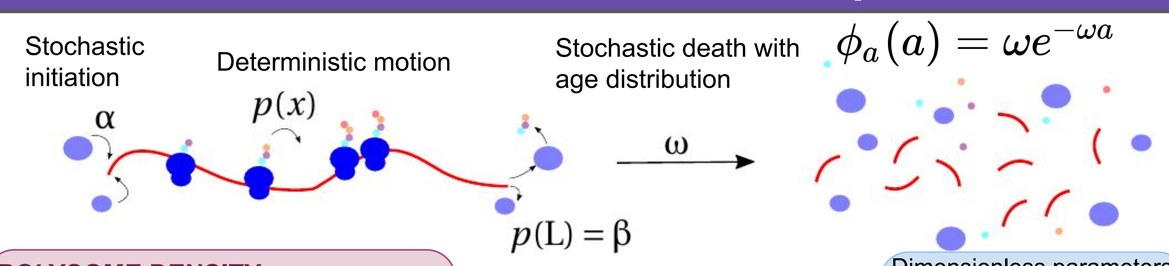


2. Ribosome sequencing on polysome and "k-somes"

Ribosome-profiling strategy is based on the deep sequencing of ribosome-protected mRNA fragments (RPF).



3. Ballistic Model of a mRNAs' soup



POLYSOME DENSITY

$$\rho(x) = \frac{\alpha}{p(x)} e^{-\omega \mathcal{T}(x)}$$

Valleriani A., Ignatova Z., Nagar A. & Lipowsky R. (2010). European Physical Letters, 89 58003.

Time spent from position 0 to x :

$$\mathcal{T}(x) = \int_0^x \frac{dy}{p(y)}$$

Dimensionless parameters

$$\tilde{\alpha} = \alpha \mathcal{T}(L)$$
$$\tilde{p}(x) = p(x)\mathcal{T}(L)$$
$$\tilde{\omega} = \omega \mathcal{T}(L)$$

K-SOMES DENSITY

k-some density for a mRNA with age *a* :

$$\rho_k(x|a) = \begin{cases} \frac{k}{a p(x)} H(x(a) - x), & a \le \mathcal{T}(L) \text{ , Transient state} \\ \frac{k}{\mathcal{T}(L) p(x)}, & a \ge \mathcal{T}(L) \text{ , Stationary state} \end{cases}$$

$\int_0^\infty \cdot \, p(a|k)\, da$ *Integration over the ages*

$$\rho_k(x) = \underbrace{\frac{\tilde{\omega}}{P_k \tilde{p}(x)} \left(\frac{\tilde{\alpha}}{\tilde{\alpha}+\tilde{\omega}}\right)^k \frac{\gamma(k,(\tilde{\alpha}+\tilde{\omega})) - \gamma(k,(\tilde{\alpha}+\tilde{\omega})\frac{\mathcal{T}(x)}{\mathcal{T}(L)})}{(k-1)!}}_{\mathcal{F}_k(x)} + \underbrace{\frac{\tilde{\alpha}^k}{P_k\, \tilde{p}(x)} \frac{e^{-(\tilde{\alpha}+\tilde{\omega})}}{(k-1)!}}_{\mathcal{S}_k(x)}$$

Transient to stationary mRNA ratio : $\mathcal{R}_k(x) = \frac{\mathcal{F}_k(x)}{\mathcal{S}_k(x)}$ characterizes the whole mRNA population (see below)

4. Finite lifetime effect analysis

Dimensionless parameters calculated from rates found in literature [*]

Species	$\omega\mathcal{T}(L)$	$\alpha\mathcal{T}(L)$	α/ω	$\mathcal{R}_1(0)$
E coli	0.04-0.07	-	-	-
S cerevisiae	0.02-0.08	4-13	170	0.3-3700
M musculus	0.002	-	-	-
H sapiens	0.003	9.6	3×10^3	5

POLYSOME

$$\rho(x) = \frac{\alpha}{p(x)} e^{-\omega \mathcal{T}(x)}$$

The exponential decay
- does not depend on α
- is negligible for biological values

K-SOMES

$$\rho_k(x) = \frac{k}{\tilde{p}(x)} \left[1 + \frac{\omega}{\tilde{\alpha}^k e^{-\tilde{\alpha}}} \left(\gamma(k,\tilde{\alpha}) - \gamma(k,\tilde{\alpha}\frac{\mathcal{T}(x)}{\mathcal{T}(L)}) - \gamma(k+1,\tilde{\alpha})\right)\right] + \mathcal{O}(\tilde{\omega}^2)$$

$$\mathcal{R}_k(0) = \frac{\omega(k-1)!}{\tilde{\alpha}^k e^{-\tilde{\alpha}}}$$

Ratio $R_k(0)$ and Taylor expansion of the k-some density lead to the same criterion for the finite mRNA lifetime impact on density profiles.

We identify **3 regimes**:

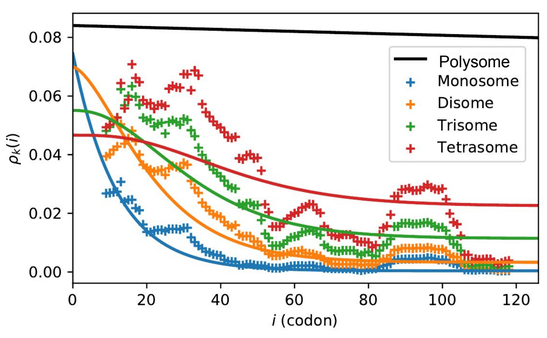
A. No finite lifetime effect on polysome nor k-some densities,
B. Finite lifetime effect on k-some densities but negligible on polysome density. Regarding the existing biological measures, this regime seems to be **the biological regime**,
C. Important finite lifetime effect on all densities (k-somes and polysome).

3 regimes profiles

[*] - Sharova et al. 2009. DNA Research 16, 45-58. - Bernstein et al. 2002. *Proceedings of the National Academy of Sciences* 99 (15): 9697–9702. - Wang et al., Proc. Natl Acad. Sci. USA 99:5860, 2002. - Schwanhäusseret et al. 2011. Nature 473 (7347): 337–42. - Young R et al. 1976 Nov 15;160(2):185-94. doi: 10.1042/bj1600185. - Karpinets et al. BMC biology vol. 4 30. 5 Sep. 2006, doi:10.1186/1741-7007-4-30. - Ingolia, Nicholas T. et al. Cell 147, 4 (11 November 2011): 789-802. - Yan et al. 2016. Cell 165 (4): 976–89. - Trösemeier et al. 2019. *Scientific Reports* 9 (1). - Ciandrini L et al. 2013. PLOS Computational Biology 9 (1): e1002866.

5. Comparison model-experiment

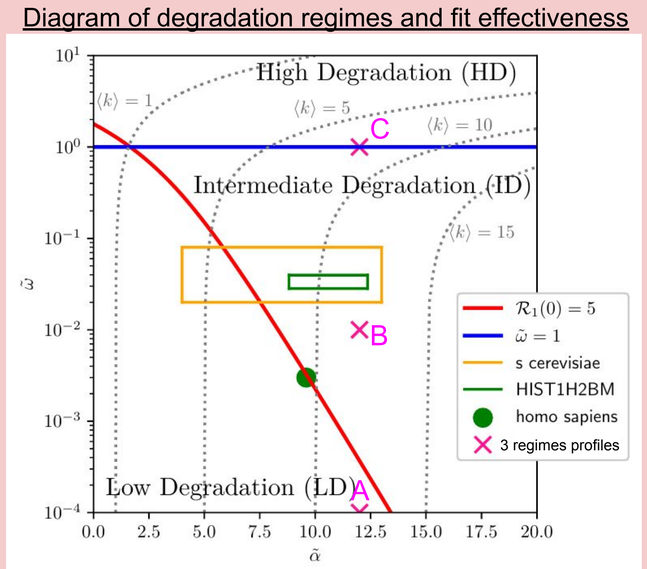
- Ribo-Seq data on k-somes of a histone gene (profiles smoothened over 19 codons)
- Ballistic model densities fit to data for a constant p and the corresponding polysome density (α=0.06 s^{-1}, ω=3.10^{-4} s^{-1}, p=0.7 codons/s and L=127 codons)



→The finite lifetime effect of the model describes Ribo-Seq data on k-somes.

8. Summary and Conclusion

Diagram of degradation regimes and fit effectiveness



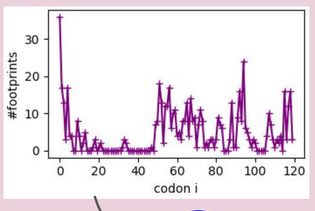
Main results:

- Model of ribosome translation in a mRNA "soup" including mRNA finite lifetime,
- Characterisation of the finite lifetime effect on the polysomes and k-omes using the model,
- Characterisation of the biological regime and evidence for degradation from Ribo-Seq data,
- New method to extract absolute kinetic parameters of translation from Ribo-Seq data.

6. Fitting Method (uses discrete form of equations)

ω measured by another experiment

Polysome profile

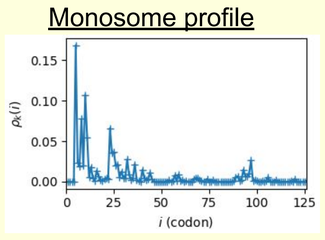


p_i

In the Ribo-Seq experiment on polysomes, the number of mRNA are not known. Therefore the real density of ribosomes from Ribo-Seq profiles can't be deduced. However, and luckily, an arbitrary normalization (arbitrarily to 1) erases the dependence on the parameter α and one can estimate the p_i from :

$$\rho_i = \frac{e^{-\omega \mathcal{T}_{i-1}} - e^{-\omega \mathcal{T}_i}}{1 - e^{-\omega \mathcal{T}_L}}$$

Monosome profile



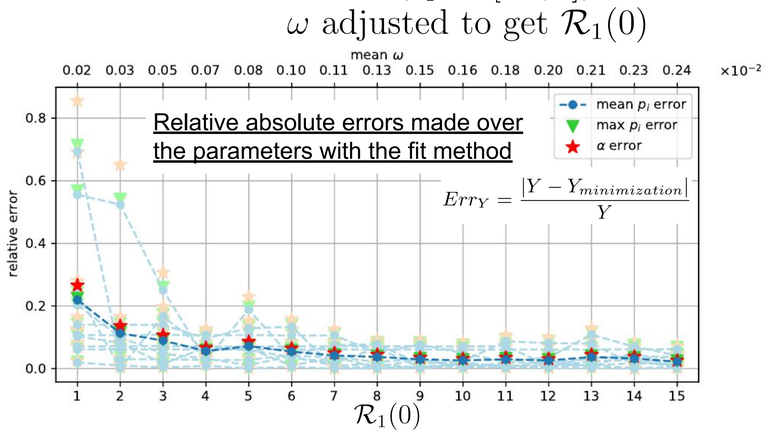
1. Density variations that are not due to p_i are due to α.
2. Finite lifetime effect need to be high enough for the density to depend on α.

α

7. Test of the fitting method in function of $R_1(0)$

Fit test of the ballistic model by the ballistic model: 10 tests profiles with same α, p_{min}, p_{max} and *L*, but p_i values are picked from different random draws.

Parameters of the test: $\alpha = 0.08,\ p_i \in [0.2, 4],\ L = 100$ (α : parameters to find; ω adjusted to get $\mathcal{R}_1(0)$)



Pale colors : errors on the estimated parameters with 10 different test profiles
Saturated colors : average errors of the 10 tests.

Errors lower than 10% for $\mathcal{R}_1(0) \gtrsim 5$