

Stochastic protein production and time-dependent current fluctuations

Peer-reviewed author version

GORISSEN, Mieke & VANDERZANDE, Carlo (2011) Stochastic protein production and time-dependent current fluctuations. In: Garrido, Pedro L.; Marro, J.; de los Santos, F. (Ed.). Non-equilibrium Statistical Physics Today, American Institute of Physics, p. 237-238.

DOI: 10.1063/1.3569513

Handle: <http://hdl.handle.net/1942/13114>

Stochastic protein production and time-dependent current fluctuations

Mieke Gorissen and Carlo Vanderzande

Faculty of Sciences, Hasselt University, 3590 Diepenbeek, Belgium

Abstract. Translation is the cellular process in which ribosomes make proteins from information encoded on messenger RNA. We model this process using driven lattice gases and take into account the finite lifetime of mRNA. The stochastic properties of the translation process can then be determined from the time-dependent current fluctuations of the lattice gas model. We illustrate our ideas with a totally asymmetric exclusion process with extended objects.

Keywords: Stochastic protein production, current fluctuations, driven lattice gases

PACS: 05.70.Ln, 87.10.Mn, 87.18.Tt

The production of proteins from genes occurs in two steps. In the first one, called *transcription*, information encoded on the genes is used to synthesize messenger RNA (mRNA). In the second step of *translation*, a ribosome uses the information stored on the codons of mRNA to make a new protein.

Cellular processes are noisy because of the small number of molecules involved and because they occur in a crowded environment [1]. In this contribution, we study the stochastic aspects of the translation process. In the standard model (Fig.1, left-side), proteins are produced with rate k while the mRNA decays with a rate λ . The number of proteins produced by one mRNA then follows a geometric distribution with average k/λ [2]. This approach does not take into account that at a particular moment several ribosomes are attached to a single mRNA which can lead to delays due to jamming and is an extra source of stochasticity [3]. A simple model that describes these effects [4] is the totally asymmetric exclusion process (TASEP) with extended objects [4, 5]. In the TASEP, mRNA is modelled as a one-dimensional lattice of N sites, each site corresponding with one codon. Ribosomes are large in comparison with a codon and are therefore represented as an extended 'particle' that covers l lattice sites. Following [5], we will take $l = 12$. Each site can be occupied by at most one ribosome. The dynamics of the TASEP is that of a Markov process: a ribosome can move forward one codon with rate 1, translation is initiated with rate α and terminates with rate β . The total number of proteins produced up to time t therefore equals the time integral $J(t)$ of the particle current $j(t)$ through the last site: $J(t) = \int_0^t j(t') dt'$. To take into account the finite lifetime of mRNA, we add an extra empty state \emptyset to the configuration space of the process. This absorbing state can be reached from any other state with rate λ describing the decay of mRNA and the end of the translation process. Our model is summarized on the right side of Fig. 1.

We are interested in the probability $P(Q)$ that Q proteins are produced by one mRNA. This quantity can easily be obtained from simulations of our model. It is however also possible to gain some further insight in this distribution from a link with known

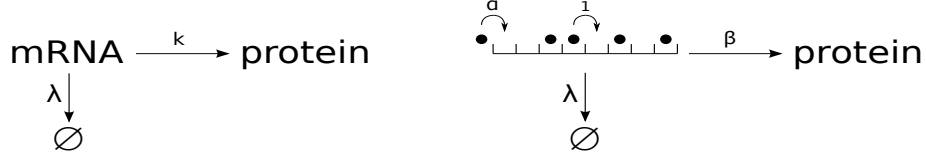


FIGURE 1. Standard model (left) and our model (right) for translation.

properties of current fluctuations.

Using a simple argument starting from the master equation, one can derive a relation between the generating function of Q , $Z(s) = \langle e^{sQ} \rangle$ and that of $J(t)$, $W(s, t) = \langle e^{sJ(t)} \rangle$, where $\langle \cdot \rangle$ is the average over the realisations of the process. This relation reads

$$Z(s) = \lambda \int_0^\infty W(s, t) e^{-\lambda t} dt \quad (1)$$

Using (1), we can relate the moments of Q to those of the current distribution.

We recently showed that the cumulants of the time dependent current in the TASEP have an interesting scaling form [6]. From that, one can then derive scaling forms for the N -dependence of the cumulants of Q . We find for example

$$\langle Q \rangle = J^* / \lambda + \sqrt{N} H(\lambda N^{3/2}) \quad (2)$$

where H is a scaling function and J^* is the steady-state current of the TASEP in the thermodynamic limit. In a mean-field approximation, it was found that $J^* = 1/(1 + \sqrt{l})^2$ in the maximal current phase of the TASEP [5]. The first term in (2) equals the one found in the standard model if we identify $J^* = k$. The second term in (2) determines deviations from the standard theory which are due to the finite size of mRNA and which can become important for small proteins. Similar expressions can be obtained for the variance of Q .

Our results indicate that mRNA-length may be an important quantity, which so far has been overlooked, in problems of stochastic gene expression.

REFERENCES

1. N. Maheshri, and E. K. O'Shea, *Annu. Rev. Biophys. Biomol. Struct.* **36**, 413-434 (2007).
2. O. Berg, *J. Theor. Biol.* **71** 587-603 (1978).
3. M. Dobrzyński, and F. Bruggeman, *Proc. Natl. Acad. Sci.* **106** 2583-2588 (2009).
4. C. MacDonald, J. Gibbs, and A. Pipkin, *Biopolymers* **6** 1-25 (1968).
5. L. B. Shaw, R. K. P. Zia, and K. H. Lee, *Phys. Rev. E* **68** 021910 (1-17) (2003).
6. M. Gorissen, and C. Vanderzande, preprint, [arXiv:1010.5139](https://arxiv.org/abs/1010.5139)