**Coarse Grained Modeling of Ribosome Availability In *E. Coli.***

Ayse Koruyucu1, Andreas Kremling1\*

*1 Systems biotechnology, Technical University of Munich, Garching b München*

*\*Corresponding author: a.kremling@tum.de*

**Highlights**

* Kinetic model to describe ribosome abundance for a broad range of growth rates
* Integration of length distribution of individual protein as basis for the determination of active ribosomes
* Estimation of kinetic parameters for protein synthesis in dependence of free available ribosomes

**1. Introduction**

The set-up of mathematical models is still a challenging task in systems biology. Although experimental data from “omics”-technologies are frequently available, the number of unknown parameters which must be estimated is still high. Therefore, in systems biology, the development of one model type, so called coarse-grained models which have only a minor number of parameters, has attracted attention in last years. This is mainly based on the observation that the relationship between large proteome fractions linearly correlate with the growth rate and allows researchers to set up and validate simple, empirical mathematical models to describe these relationships. In this way, general design principles of cellular resource allocation can be postulated.

**2. Methods**

A pioneering role plays the group of Terence Hwa, California, U.S.A. which developed a holistic framework for the description of cellular bio-chemical reaction networks [1]. Although simple, these models are used to make predictions for different experimental conditions like additional load in form of foreign protein expression and altered kinetic parameters, for example, for different values of the translation velocity. A general drawback of empirical models is, that they are not based on equations that guarantee mass conservation; this is typically realized by setting up a mass balance equation and deducing equations for intracellular concentrations (typically, units like mol/g dry weight or mol/ l cell are used for components of the network. Kinetic expressions for synthesis and degradation are based on the state variables of the model and therefore contain “lumped” information. In coarse-grained models, the number of state variables is low (smaller than 5) and (in general) they are analytically solvable for quasi steady state: if the dynamic of the system is not in the focus, a set of algebraic equations results from setting time derivatives to zero. Finally, the solution space is analyzed by a variation, for example, of the growth rate and calculation of the steady state values of the state variables.

In the presented study, we focus on the synthesis of the overall proteome of the cell taking into account the availability of ribosomes and neglecting the pure metabolic substrate flux in form of amino acids. The model takes into account geometric parameters like the length distribution of individual proteins, kinetic parameters for transcription/ translation, and the abundance distribution of proteins measured under different conditions.

**3. Results and discussion**

For the total protein *P* in the cell, a general differential equation in the form $\dot{P }= r\_{syn}- μ P$ with the total rate of protein synthesis $r\_{syn}(n\_{Rf})$ and the specific growth rate µ is used. The rate of synthesis depends on the free available ribosomes *nRf* in the cell. To count the active ribosome *nRa* for a distinct growth rate, we estimate the number of active RNA polymerases and, depending on the position of the RNA polymerase on the DNA strand, the number of ribosomes currently translating the mRNA. For a single transcription unit *j* one get according to [2]: $n\_{Rj}= m\_{j}\frac{\left(n\_{Pj}+1\right)}{2 d\_{R}} $, with the length *mj* of the transcription unit, *nPj,* the number of RNA polymerases, and *dR* the distance between two ribosomes. This term can be reformulated to take into account the steady state number of the respective protein that depend on the selected growth rate. In this way, integrating over all proteins with individual length and the relationship for the total number of ribosomes, that is, the sum of active and free ribosome (*nRt = nRf* + *nRa* ), we obtain a relationship for the total protein in steady state: $r\_{syn}(n\_{Rt}- n\_{Ra}(P))= μ P$. We analyzed several possible kinetic expressions for the rate of synthesis *rsyn* and fit the model to available experimental data for the total number of ribosomes. Figure 1 shows the length distribution for the protein and the final fit for the total number of ribosomes.



**Figure 1.** Left: (A) Protein length distribution with data from [3], (B) with steady state data weighted distribution. Right: Experimental data [4] and simulation comparison for the total number of ribosomes.

**4. Conclusions**

The model successfully describes experimental data for a broad range of the growth rate. Sensitive parameters of the model are the growth rate dependent velocities of the ribosomes and the RNA polymerase, as well as the total number of transcription units in the cell.

**References**

1. M. Scott et al., Molecular Systems Biology 10, 2014
2. A. Kremling, Biotechnology & Bioengineering 96, 2007
3. A. Schmidt et al., Nature Biotechnology 34, 2016
4. H. Bremer & P.P. Dennis, EcoSal Plus, 2008