

Bojana Blagojevic¹, Andjela Rodic², Evgeny Zdobnov³, Magdalena Djordjevic¹, Marko Djordjevic²

¹Institute of Physics, University of Belgrade, Belgrade, Serbia

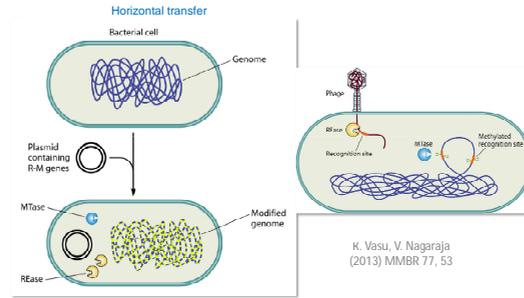
²Institute of Physiology and Biochemistry, Faculty of Biology, University of Belgrade, Belgrade, Serbia

³Department of Genetic Medicine and Development, University of Geneva and Swiss Institute of Bioinformatics, Geneva, Switzerland

BMC Systems Biology 2017 Feb 24; 11(Suppl 1):2

INTRODUCTION

Restriction-modification (R-M) systems (Type II) represent rudimentary bacterial immune systems since they protect the host bacterial cell against the infection by foreign DNA such as viruses (bacteriophage). RM genes and their promoters are usually plasmid-encoded, and they spread from one bacterial host to the other, through horizontal transfer. Their main ingredients are restriction endonuclease (R) and methyltransferase (M). Restriction enzyme recognizes and cuts the same unmethylated DNA sequences that are recognized and protected by M. This implies that methylation must precede the expression of R gene, in order to prevent the autoimmunity. Once the host genome is protected, the system should rapidly turn from OFF to ON state, so that the cell becomes 'immune' to the virus infection as fast as possible. Consequently, the expression of R and M genes has to be tightly regulated during RM system establishment in naive bacterial host, which is usually achieved through a dedicated transcription factor (control protein C).



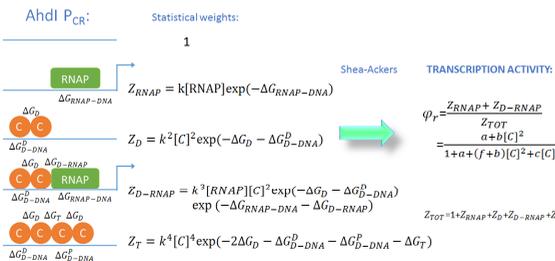
RM systems may exhibit different architectures (convergent and divergent) and different regulatory features, e.g. dimerization constants, binding cooperativities, translation rates. We here propose that all RM systems should meet the same dynamical properties that ensure the host genome protection and efficient system establishment:

- ❑ time delayed expression of R with respect to M,
- ❑ rapid transition of R from "OFF" to "ON" state
- ❑ increased stability of R steady state.

The relationship between these constraints on the dynamical response on one side, and regulatory features/architectures on the other, is still unclear. To address this question, we developed a biophysical model of gene expression regulation for convergent and divergent RM systems (we previously obtained experimental validation for convergent architecture). Then we computationally perturb characteristic regulatory features in both systems, and observe its effect on the dynamical properties. We here obtain that both systems exhibit the same aforementioned dynamical properties, while any perturbation of the systems features makes at least one dynamical property much less optimal, which may explain the existence/absence of the corresponding features.

METHODS

Our thermodynamical model of transcription regulation applies Shea-Ackers assumption of transcription activity being proportional to binding probability of RNAP to promoter, and is parameterized by available experimental data (for core promoters, C binding sites, binding affinities and transcription activity).

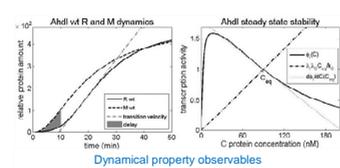


Thus obtained transcription activity, serves as an input for modeling the dynamics of transcript and protein expression:

$$\frac{d\varphi}{dt} = \varphi - \lambda_{\varphi} \varphi$$

$$\frac{dP}{dt} = k_{\varphi} \varphi - \lambda_P P$$

Next we quantify dynamical properties by defining the corresponding observables.



Finally, we apply the model to RM systems with convergent (AhdI) and divergent (EcoRV) organization of CR and M promoters, and ask if the systems exhibit the proposed three dynamical properties. Finally, to explain the systems features, we computationally perturb them in both systems.

I) Convergent system (AhdI)

- Regulatory features extracted from the experimental data:
 - ❖ Very high equilibrium C dimerization constant of disociation
 - ❖ C transcript is leaderless
 - ❖ Extremely large cooperativity of C dimer binding to distal and proximal binding site in CR promoter
- CR promoter transcription is regulated by C protein
- M promoter transcription is regulated through negative feed-back loop

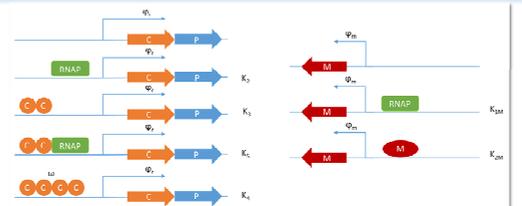


Figure 1. AhdI: Scheme of transcription configurations

II) Divergent system (EcoRV)

- Regulatory features extracted from the experimental data:
 - ❖ Partially overlapping CR and M core promoters
 - ❖ Absence of all AhdI regulatory features
- CR promoter transcription is regulated by C protein in the same manner as AhdI
- M promoter transcription is indirectly influenced by C (due to overlapping promoters)

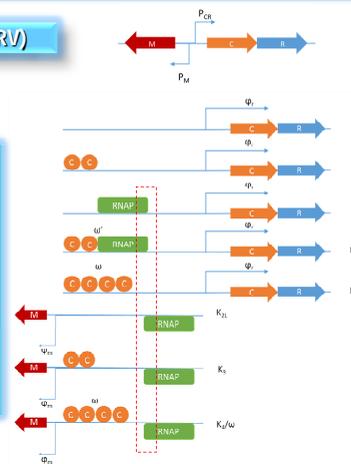
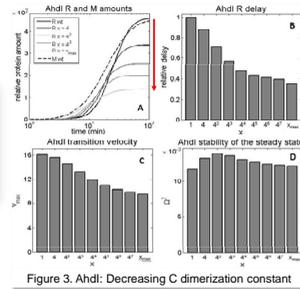


Figure 2. EcoRV: Scheme of transcription configurations

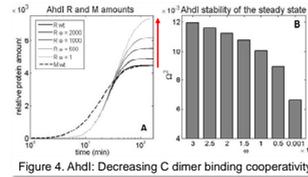
RESULTS

We obtained that, **despite different architectures and features, both RM systems satisfy the same dynamical properties.** Next we **computationally** perturbed systems either by abolishing the existing regulatory features (AhdI) or by introducing the new ones (EcoRV).

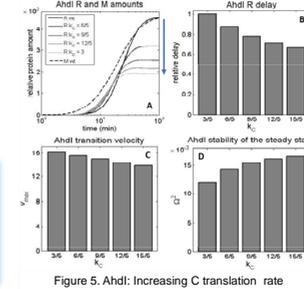
I) Perturbations of convergent system (AhdI) by abolishing the existing features



Significant adverse effect on two dynamical properties.



Significant adverse effect on one, while not affecting the other two dynamical properties.



Significant adverse effect on one dynamical property.

High binding cooperativity has an **opposite effect** from high dimerization constant on R steady state level!

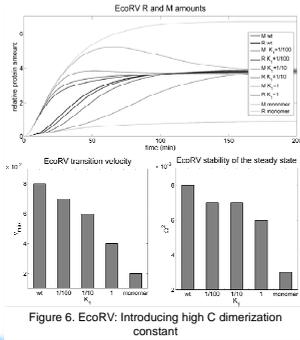
Permit balancing the amounts of R and M.

Are these two features always jointly present/absent? (AhdI, Esp1396I, EcoRV)

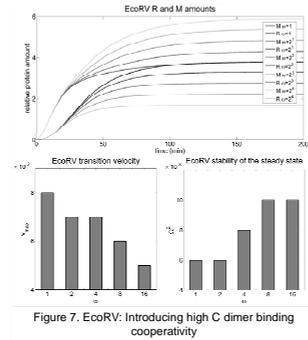
M to R ratio consistently increases for each perturbation!

Does not allow balancing the amounts of R and M.

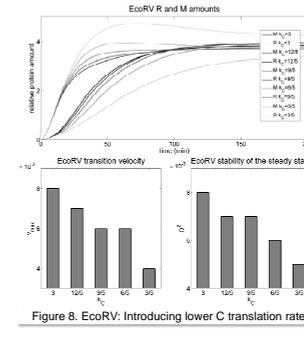
II) Perturbations of divergent system (EcoRV) by introducing AhdI features



Significant adverse effect on two dynamical properties.



Significant adverse effect on one dynamical property.



Significant adverse effect on two dynamical properties.

SUMMARY

I) For convergent system (AhdI) we obtained that:

the system meets all three dynamical properties;

- ❖ all three regulatory features, in general, have **the same effect** on the dynamical properties, i.e. abolishing them makes at least one dynamical property much less optimal, while not significantly affecting the other properties;
- ❖ on the other hand, extremely large binding cooperativity has **the opposite effect** from the other two features on the R steady state level, which allows balancing the amount of the toxic molecule.

May explain the existing AhdI regulatory features.

II) For divergent system (EcoRV) we obtained that:

system also meets all three dynamical properties;

- ❖ introducing AhdI regulatory features to EcoRV makes at least one dynamical property much less optimal.
- ❖ this introduction changes M to R ratio in the same direction, which does not allow balancing the amounts of R and M proteins in the host cell

May explain the absence of AhdI regulatory features from EcoRV system.

CONCLUSIONS

- ❖ Our theoretical model of the enzyme expression in RM systems, which is based on statistical thermodynamics and experimental data, leads to a very good agreement with experimentally measured transcription activity (AhdI) and protein expression dynamics (Esp1396I).
- ❖ RM systems exhibit different architectures and features, which motivated the question: can these features be explained by constraints imposed by few simple dynamical properties. To this end, we theoretically modeled convergent and divergent system expression regulation, and obtained that **both systems exhibit the same dynamical properties.**
- ❖ In order to explain the existence/absence of systems features, we **computationally perturbed** them in convergent and divergent system, **which led to diminishing of the dynamical properties.**
- ❖ As an outlook, it will be interesting to test if other RM systems, such as poorly studied linear, or even more advanced prokaryotic adaptive immune system, such as CRISPR/Cas, meet the proposed three dynamical properties. Also the progress of experimental techniques will allow testing our conclusions regarding the relationship between architectures/features and the constraints on systems dynamical response.

ACKNOWLEDGEMENTS

This work is funded by the Swiss National Science Foundation (I27320_152297), European Commission (GA-2010-276996) and Serbian Ministry of Science and Technology (ON173052).