A new quantitative model for CRISPR/Cas activation

Andjela Rodic[1], Marko Tumbas[1], Jane Kondev[2], Magdalena Djordjevic[3] and Marko Djordjevic[1]

[1]Quantitative biology group, Faculty of Biology, [2]Martin A. Fisher School of Physics, Brandeis University [3]Institute of Physics Belgrade, University of Belgrade

Science Fund of the Republic of Serbia

Background

A model system for CRISPR/Cas regulation (Type I-E in *Escherichia coli*) is silent under standard conditions, and it is unclear if and how the system can be activated rapidly enough to protect the cell. Two qualitative models were proposed based on the fact that cas genes and CRISPR array are cooperatively repressed by H-NS (a pleiotropic transcription regulator):

- The first model assumes that foreign DNA generally has higher AT content and that H-NS prefers binding to AT reach sequences. When foreign DNA enters the cell, the concentration of free repressor is lowered due to binding to the foreign DNA sequence, relieving CRISPR/Cas repression. However, it is unclear if the perturbation of H-NS is sufficient to efficiently activate the system.
- According to the second model, a transcription activator LeuO (H-NS antagonist) is activated upon virus entry, which relieves CRISPR/Cas repression. However, what exactly activates LeuO?

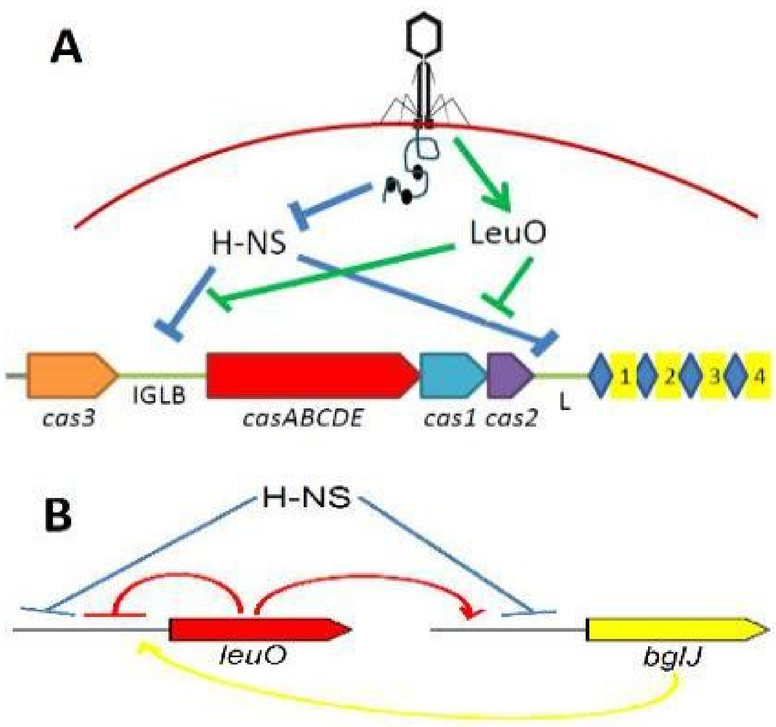


Figure: A) **Scheme of two CRISPR/Cas regulation models.** At the bottom of the Figure, a scheme of the Type I-E CRISPR/Cas system in *E. coli* is shown. Repeats and spacers are indicated in blue and yellow, respectively. Spacers are often complementary to phage and plasmid sequences, which confers immunity against foreign DNA invasion. The prominent CRISPR/Cas induction models are indicated by blue and green arrows, respectively. The promoters for cas genes (in the IGLB region) and CRISPR array (in the Leader region) are repressed by strong cooperative interactions by H-NS. This repression is abolished upon induction signal, according to one of the two models. Arrows and flat-ended lines indicate activation and repression, respectively. Note that LeuO abolishes repression by H-NS, as indicated in the Figure. B) **Scheme of LeuO regulation.** Red and yellow blocks represent *leuO* and *bglJ* genes, respectively, repressed by H-NS. LeuO activates the transcription of *bglJ* (red arrow) and represses its transcription (red bar-headed line), while BglJ activates *leuO* transcription (yellow arrow).

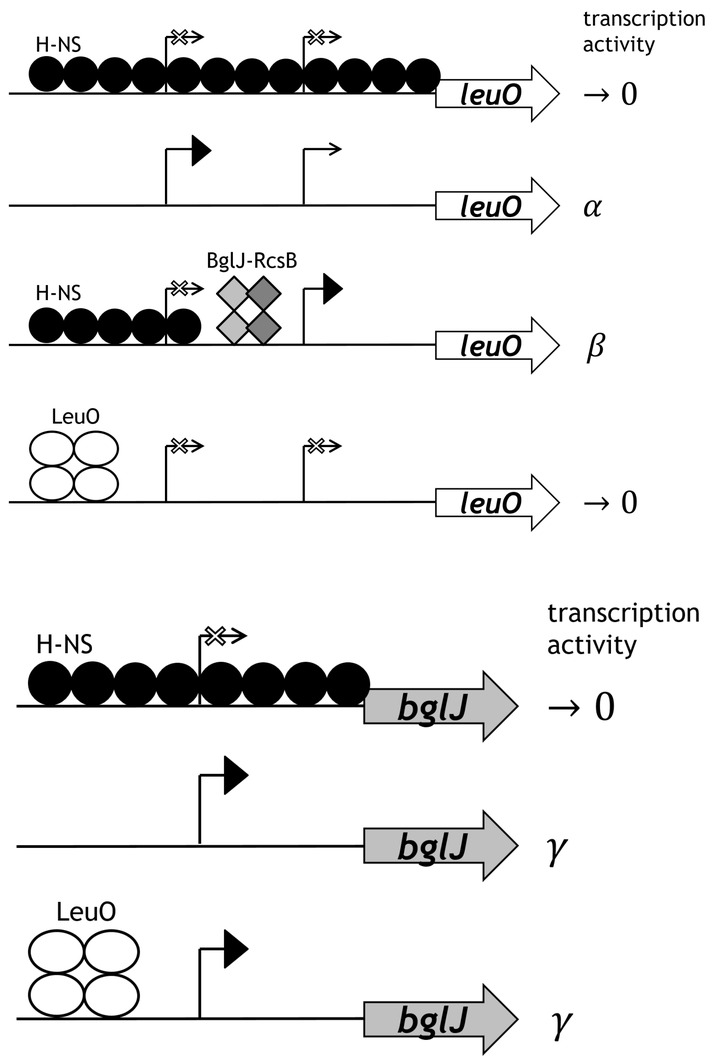
New model

We here propose a new model, which combines the mechanisms assumed by models i) and ii), as they both have clear deficiencies individually. Note that leuO is strongly activated by BglJ, that both *leuO* and *bglJ* are repressed by H-NS, while LeuO also activates *bglJ*. Consequently, lowering H-NS concentration through the mechanism in model i) can lead to strong LeuO activation by a double positive feedback loop: i.e., decreasing the amount of H-NS, would lead to initial derepression of *leuO*, activating BglJ, which would, in turn, lead to strong activation of LeuO. Once LeuO is activated, CRISPR/Cas can be further induced through mechanism ii).

Methods

Bioinformatics A large-scale analysis of CRISPR array spacers from 31845 complete bacterial genomes was performed. All bacterial and 16388 viral genomes were retrieved using NCBI datasets API. Newly developed CRISPRidentify and CRISPRcasIdentifier tools based on machine learning were used for CRISPR array, Cas genes detection, and subtyping. Viral genomes were mapped to their hosts using the latest version of the Virus-Host DB. Mapping was performed on the genus level of the host phylogenetic tree. GC content for bacterial and viral genomes was calculated for each genus.

Models



Statistical weights of promoter configurations are used to calculate promoter transcription activities of *leuO* and *bglJ* genes. The coupled dynamics, in the rescaled form, of LeuO (L) and BglJ (B) is given by the equations below:

$$\frac{dl}{d\tau} = \frac{\alpha(1+H^n B)}{1+H^n+H^m B+L} - l$$

$$\frac{dL}{d\tau} = kl - \lambda L$$

$$\frac{db}{d\tau} = \frac{\gamma(1+L)}{1+H^p+L} - b$$

$$\frac{dB}{d\tau} = kb - \lambda B$$

Regulation of Cascade genes, including Cas6e (denoted by E for protein and e for its transcript) that processes pre-crRNA to crRNA is equivalent to the regulation of *leuO* gene:

$$\frac{de}{d\tau} = \frac{\psi(1+L)}{1+H^n+L} - e$$

$$\frac{dE}{d\tau} = ke - \lambda E$$

Finally, the processing of pre-crRNA to crRNA is described by the following equations (adopted from PMID: 22849651), where pre-crRNA is denoted by p (rapidly degraded by λ_p), and crRNA is denoted by c (slowly degraded by λ_c):

$$\frac{dp}{dt} = \phi - (\lambda_p + Ek^*)p$$

$$\frac{dc}{dt} = Ek^* p - \lambda_c c$$

Results

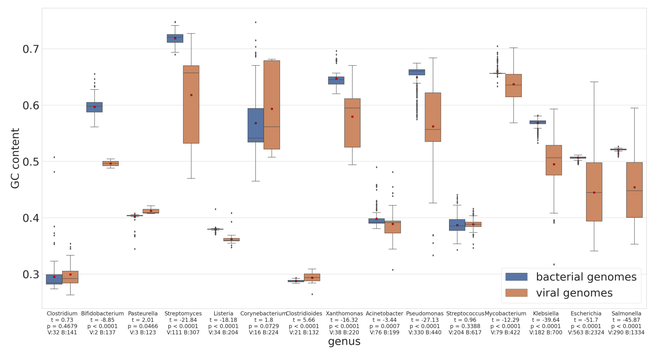


Figure: Consistently with model assumptions, bacteriophage GC content is significantly lower than host bacteria. The difference is, however, quantitatively not large ~10%, suggesting relatively small $\delta H/H_0$.

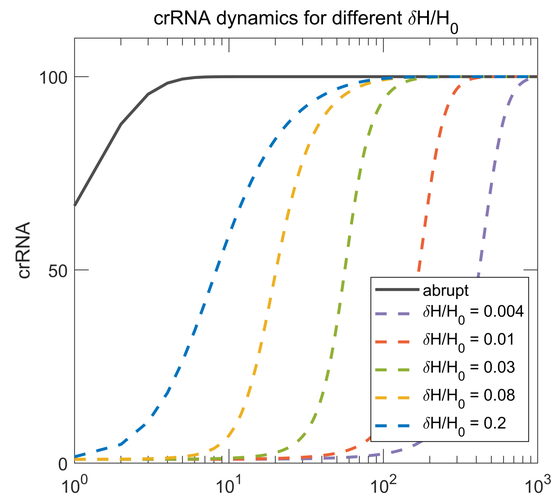


Figure: Figure shows that increasing $\delta H/H_0$ significantly changes (decreases) the delay in crRNA generation. The "abrupt" curve corresponds to an instantaneous processing rate increase.

	n ($\tilde{H}_0$)	30 (>1.328)	20 (>1.531)	13 (>1.925)	9 (>2.576)	6 (>4.135)	4 (>8.409)
	$\tilde{H}_0$	1.513	1.861	2.6	3.976	7.927	22.318
$\frac{\Delta\tilde{H}}{\tilde{H}_0}$	0.07	■	■				
	0.1	■	■				
	0.12	■	■				
	0.15	■	■	■			
	0.18	■	■	■			
	0.21	■	■	■	■		
	0.24	■	■	■	■		
	0.26		■	■	■		
	0.29		■	■	■	■	
	0.31		■	■	■	■	
	0.34		■	■	■	■	
	0.37			■	■	■	
	0.4			■	■	■	■
	0.44			■	■	■	■
	0.48			■	■	■	■
	0.51				■	■	■
	0.54				■	■	■
	0.58				■	■	■
	0.61				■	■	■
	0.68					■	■
	0.75					■	■
	0.81						■
	0.88						■

Figure: Shaded regions in the table show combinations of $\delta H/H_0$ and H_0 for which sufficient (experimentally observed) increases of pre-crRNA processing can be achieved. One can see that small $\delta H/H_0$ require H_0 values close to 1.

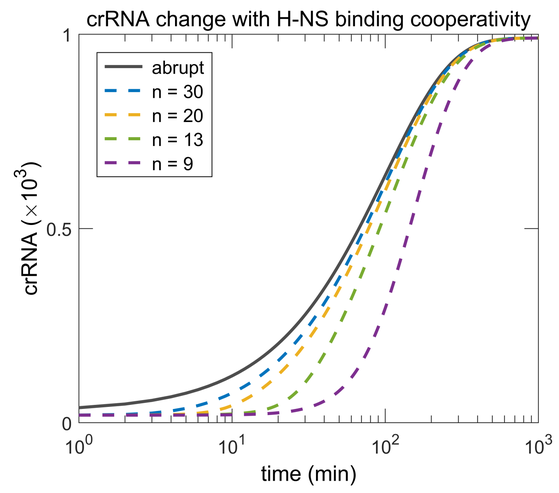


Figure: Increasing the repressor (H-NS) binding cooperativity (n), decreases the delay in crRNA generation

Conclusion

The delay with which crRNA is generated upon bacteriophage entry strongly depends on the decrease in free H-NS in solution (caused by H-NS binding to AT-reach phage genome). This decrease is likely, not large. Relatively small H-NS perturbation can, however, be compensated by i) H-NS large binding cooperativity to promoter ii) H-NS concentration being close to its binding dissociation constant so that in combination with large binding cooperativity, the system becomes maximally sensitive to changes in free H-NS concentration.

Acknowledgements

This work is supported by the Science Fund of the Republic of Serbia, projects no. 7750294, q-bioBDS and no. 6417603, CRISPR modelling.