

Selecting codons not for proteins: codons and amino acids biases around proteins binding DNA sites are due to their electrostatics

A. Osypov^{1,2*}, G. Krutinin³, E. Krutinina³, P. Beskaravayny³, S. Kamzolova³

¹ Institute of Higher Nervous Activity and Neurophysiology RAS, Moscow, Russia

² Institute of Theoretical and Experimental Biophysics RAS, Pushchino, Moscow Region, Russia

³ Institute of Cell Biophysics RAS, Pushchino, Moscow Region, Russia

* e-mail: aosypov@gmail.com

Key words: DNA electrostatics, transcription regulation, genome evolution

Motivation and Aim: DNA is a highly charged molecule and its electrostatic and other physical properties define its interactions with different proteins, especially those regulating transcription in Prokaryotes. Electrostatic potential (EP) is distributed non-uniformly along DNA and correlates with GC content, depending on the sequence arrangement and its context [1]. Binding sites of transcription factors of different protein families in different taxa are located in wide areas of high electrostatic potential, multiple times of the protein size; EP distribution on transcription factors protein surface reflects that of their binding sites [2]. Promoters in average have high value of EP profile [1]. Hundreds of transcription factors binding sites lie in protein coding areas. Some protein coding areas also host promoters for the genes mainly located on the opposite strand.

Methods and Algorithms: DEPPDB and its tools [1, 2] were used to carry out the analysis.

Results: We found that in such diverse bacteria as *E. coli*, *B. subtilis* and *Corynebacterium glutamicum* (and other species) there are codons and even amino acids biases around promoters and transcription factors binding sites spanning for a hundred of codons and that these biases are due to physical properties of considered codons providing proper electrostatic attractors for transcription regulating proteins.

Conclusion: The data obtained demand serious rethinking of the concept of molecular evolution. The apparent non-synonymy of synonymous substitutions may lead to different wrong estimations of the sequences fate, including misuse of the molecular clock and mistaken evaluations of specific mutations biomedical importance. The amino acids bias leads to even more important shift in the conception of natural selection in proteins. It substantiates the view on the DNA as not only the text of the first step in realizing the Central Dogma, but a complex organ of heredity that fulfills different and sometimes contradictory demands.

Acknowledgements: Supported by the RFBR (16-04-01865).

References

1. Osypov A.A. et al. (2010) DEPPDB – DNA Electrostatic Potential Properties Database: Electrostatic properties of genome DNA. JBCB. 8(3):413-425.
2. Osypov A.A. et al (2012) DEPPDB – DNA Electrostatic Potential Properties Database: Electrostatic properties of genome DNA elements. JBCB. 10(2):1241004.