

Review on thesis of

Mgr. Vyacheslav Tretyachenko

The effect of amino acid repertoire on protein structure and function

Dissertation thesis of Mgr. Vyacheslav Tretyachenko submitted for defense is focused on the analysis of the effect of the amino acid repertoire on the structure and protein function. This is an unresolved key issue. The function of proteins depends on their spatial structure, and the secondary and tertiary structure of proteins derived from their primary structure tend to be ambiguous. Understanding this relationship is important not only for basic research, but also has a great potential for biotechnology.

The work itself is 69 pages long and it includes three publications as a appendix. Vyacheslav Tretyachenko is the first author on two publications and the publications were published in good peer-reviewed journals. One publication was already cited 18 times. The author states that he participated in three other publications, but did not include them as a part of his dissertation. The work has a standard chapters, the focus is on the chapter Results containing the mentioned publications. I liked the nicely conceived introduction to the evolution of the protein alphabet, a topic that has been discussed in the literature for a long time, and the author made a nice summary.

The presented work is a proof of the author's ability to work on complex problems of contemporary science. The methodology used includes not only the common bioinformatics procedures and software, but also the development of own software. A strong element of the whole work is the connection of computational methods with methods experimental by which theoretical models are verified. In many articles these two roles are often divided among several authors, here they are both represented by one author and the understanding of both sides, computational and experimental, had positive impact on the work. The presented results are of international quality. The defined goals of the dissertation were met. While reading, I did not come across any signs of plagiarism.

I am pleased to recommend the submitted dissertation for defense.

V Praze 11. 6. 202

A handwritten signature in black ink, appearing to read 'Jan Pačes', written in a cursive style.

Jan Pačes, Ph.D.

Ústav molekulární genetiky AV ČR

Vídeňská 1083, Praha

Questions:

Statistics of the occurrence of individual amino acids were calculated across entire databases. Do these statistics really apply across the whole domain? Wouldn't it be interesting to separate e.g. eukaryotes and prokaryotes? Or plants from animals? Has the alternative genetic code been taken into account in any way?

This work deals with reduced amino acid alphabets. However, there are articles that go in the opposite direction by extending the amino acid alphabet, either by alternating the genetic code or even extending the DNA alphabet by another pair of bases. I wonder if the author has any ideas in this direction and what perspectives he sees here.