

## **Into the wild: expression and characterization of random protein libraries**

Vyacheslav Tretyachenko<sup>1</sup>, Jiří Vymětal<sup>2</sup>, Lucie Bednarová<sup>2</sup>, Kosuke Fujishima<sup>3</sup>,

Jiří Vondrášek<sup>2</sup>, Klara Hlouchová<sup>1,2</sup>

<sup>1</sup> Charles University in Prague, Faculty of Science, Department of Biochemistry, Hlavova 2030, 128 00, Praha 2, Czech Republic, klara.hlouchova@natur.cuni.cz

<sup>2</sup> Institute of Organic Chemistry and Biochemistry, Flemingovo nám. 2, 166 10, Praha 6, Czech Republic

<sup>3</sup> Earth-Life Science Institute, Tokyo Institute of Technology, Tokyo, 1528550, Japan

Modern protein science heavily relies and benefits from the data generated from experimental characterization of natural protein sequences. Our study moves beyond the natural world in attempt to construct and describe the behaviour of random protein libraries without any evolutionary background. In order to investigate the structure-forming potential of random proteins we designed and applied a novel random library construction and purification methodology. Our libraries, although being random in sequence, are customized in amino acid content and ratios. This approach allowed us to study the secondary structure content of (i) natural-like random proteins composed of all 20 amino acids, library of proteins built from amino acids present in (ii) prebiotic and (iii) early biotic world and protein library made from (iv) minimal set of amino acids from the rational protein design point of view. In addition to experimental characterization we performed bioinformatic screening of random libraries in order to unveil structural landmarks of different amino acid alphabets.