INSTITUTE OF BIOCHEMISTRY AND BIOPHYSICS
POLISH ACADEMY OF SCIENCES

1. **Research Unit:** Laboratory of Biological Chemistry of Metal Ions
2. **Supervisor:** dr hab. Marcin Grynberg
3. **Supervisor (email):** greenb@ibb.waw.pl
4. **Co-supervisor:** dr hab. inż. Aleksandra Gruca (Silesian University of Technology)
5. **Project title (English):** Creating a benchmark and comparing the performance (correctness, number of detected LCR pairs) of LCR comparison methods.
6. **Project title (Polish):** Stworzenie benchmarku oraz porównanie wydajności (poprawność, liczba wykrytych par LCRów) metod porównywania LCRów.
7. **Description of the project (up to 500 words):**

Low complexity regions (LCRs) are protein fragments with a low diversity of amino acids. Although about 14% of proteins contain such fragments, for many years low complexity regions were ignored by the scientific community and treated as non-functional parts of a proteome. While the function of most LCRs is still a mystery, recent evidences suggest that LCRs often play important roles in structure stability preservation, protein-RNA interactions, phase separation, adhesion, transduction of conformational information, membrane interactions, DNA binding, the binding of metals by cysteine, histidine, or charge clusters, in driving the formation of membraneless organelles through phase separation and some other processes. LCRs are also directly involved in the development of various diseases including neurodegenerative diseases and cancer.

While presently there are several tools to identify LCRs, their subsequent analyses is only an emerging field. In particular, there are no tools yet that can compare or exploit sequence similarities across LCRs for a reliable prediction of likely LCR functions. Recently we developed the first programmes able to compare LCRs (GBSC, LCR-BLAST, MotifLCR, EvansP). However there is a lack of benchmarks that will allow the comparison of efficacy, coverage and sensitivity of LCR space discovery using these methods. **The aims of this work would be to semi-manually design such a benchmark and to test in-house programmes using this benchmark**.

8. **References related to conducted /planned research (maximum 3)**:

   - Gruca A, Ziemska-Legiecka J, Jarnot P, Sarnowska E, Sarnowski TJ, Grynberg M. Common low complexity regions for SARS-CoV-2 and human proteomes as potential multidirectional risk factor in vaccine development. BMC Bioinformatics. 2021 Apr 8;22(1):182. doi: 10.1186/s12859-021-04017-7. PMID: 33832440; PMCID: PMC8027979.
   - Tørresen OK, Star B, Mier P, Andrade-Navarro MA, Bateman A, Jarnot P, Gruca A, Grynberg M, Kajava AV, Promponas VJ, Anisimova M, Jakobsen KS, Linke D. Tandem repeats lead to sequence assembly errors and impose multi-level challenges for genome and protein databases. Nucleic Acids Res. 2019 Dec 2;47(21):10994-11006. doi: 10.1093/nar/gkz841. PMID: 31584084; PMCID: PMC6868369.
   - Kubáň V, Srb P, Štégnerová H, Padrta P, Zachrdla M, Jaseňáková Z, Šanderová H, Vítovská D, Krásný L, Koval' T, Dohnálek J, Ziemska-Legiecka J, Grynberg M, Jarnot P, Gruca A, Jensen MR, Blackledge M, Žídek L. Quantitative Conformational Analysis of Functionally Important Electrostatic Interactions in the Intrinsically Disordered Region of Delta Subunit of Bacterial RNA Polymerase. J Am Chem Soc. 2019 Oct 23;141(42):16817-16828. doi: 10.1021/jacs.9b07837. Epub 2019 Oct 10. PMID: 31550880.

9. **Scholarship amount (net):** 3000 PLN for mid-term evaluation, after mid-term evaluation, change to 57% professor's remuneration (currently it would be 3242 PLN net).