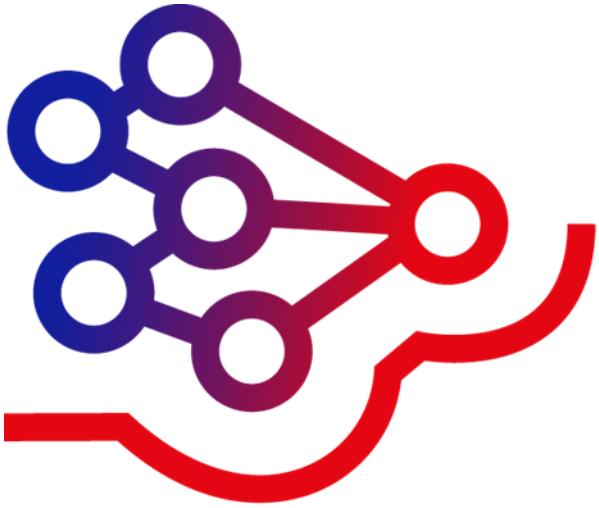

ML4NGP

MACHINE LEARNING FOR NON GLOBULAR PROTEINS

ABSTRACT BOOK

1st MEETING on MACHINE LEARNING AND NON-GLOBULAR PROTEINS

5 – 7, July 2023

Hotel Clarion, Bratislava, Slovakia

cost EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

Funded by the European Union

S10

LEA4 proteins: How disordered are they?

Marija Vidović [1], Ana Pantelić [1], Milan Senćanski [1], Ivana Prodić [1]

[1]Institute of Molecular Genetics and Genetic Engineering, University of Belgrade, Serbia

Resurrection plants (such as Ramonda serbica) can survive a long desiccation period and fully resume their metabolic functions rapidly upon watering. The hallmark of desiccation tolerance (DT) is the accumulation of protective, intrinsically disordered proteins, called late embryogenesis abundant proteins (LEAPs). Their high structural plasticity allows them to interact with various ligands and partners and stabilise membranes and enzymes during drying. However, no specific cellular targets of LEAPs have been identified so far.

To propose their function in DT, we structurally characterised and classified LEAPs in hydrated and desiccated R. serbica leaves. Even 17 proteins belonging to the LEA4 protein family group were induced by desiccation. They exhibit high disorder propensity (82 %), as assessed by four disorder estimators (DisProt, FELLS, Espritz-X, Iupred). At the same time, however, LEA4 proteins show an exceptionally high tendency to form α-helices (>80% of the sequence length) as showed by five secondary structure predictors: FELLS, JPred, SOPMA, PsiPred, and Phyre2. Moreover, motifs corresponding to the R. serbica LEA4 protein family group are predicted to fold into A-type α-helices containing positive, negative, and hydrophobic surfaces that could allow them to immerse laterally into the membrane and reinforce it during drying. To predict the 3D structure of the representative LEA4 protein family member, we performed molecular dynamics simulations (MDS) under hydration, desiccation and in the presence of a membrane.

Alfafold modelling and MDS confirmed that almost the entire length of the LEA4 protein adopted α-helical structure, with a single intramolecular salt bridge orienting polar and charged amino acid residues towards the water medium. Moreover, the conserved M4.1 domain folds into an α-helix that is stabilised by intramolecular electrostatic/hydrophilic interactions when embedded in the lipid bilayer, with hydrophobic side residues interacting with fatty acid tails.

In summary, although the LEA4 protein family member contains a high percentage of amino acids related to disorder, it shows a high propensity to form different amphipathic α-helices depending on the microenvironment. Therefore, the notion of disorder needs to be re-evaluated in terms of random coil content on the one hand and promiscuity to form different helical structures on the other. Nevertheless, our in silico findings will be experimentally validated in ongoing studies using the recombinant LEA4 protein. Our study is an important starting point for future efforts to elucidate LEAP's mechanism of action at the cellular level, especially regarding their structural flexibility, which is not yet known.

Keywords: α-helix, desiccation, intrinsically disordered proteins, late embryogenesis abundant proteins, Ramonda serbica, resurrection plants

Acknowledgements: This work was funded by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia (Contract No. 451-03-47/2023-

01/200042) and by the Science Fund of the Republic of Serbia-RS (PROMIS project LEAPSyn-SCI, grant no. 6039663).