IMPROVING RELIABILITY OF PROTEIN FOLDING RATE PREDICTIONS USING SEQUENCE AND STRUCTURAL INFORMATION

Bono Lučić,[a] Antonija Kraljević,[b,c] Viktor Bojović,[a] Jadranko Batista[d]

[a] Ruđer Bošković Institute, NMR Centre, Zagreb, Croatia
[b] University of Mostar, Faculty of Mechanical Engineering, Computing and Electrical Engineering, Mostar, Bosnia and Herzegovina
[c] University of Split, Postgraduate university study in biophysics, Split, Croatia
[d] University of Mostar, Faculty of Science and Education, Mostar, Bosnia and Herzegovina
lucic@irb.hr

Protein folding is regarded as an extremely important problem in the life sciences, chemistry, and biophysics, and has been intensively studied for many years both experimentally and through modelling.[1] It is assumed that all key information related to the process and rate of protein folding is contained in the primary structure of the protein. Over the past 30 years, the dependence of protein folding rate k_f (s^{-1}), defined as $k_f = 1/t_f$ (where t_f is the folding time) has been intensively modelled. Among the most important structural descriptors related to protein folding rate are protein length, the content of regular secondary structures, and the topology and average contact order distance between amino acids in the 3D structure have been identified as the most significant.[2] We conducted comparative studies of different methods for predicting protein folding rates, available as web servers and from the literature, and compared the results with those of a previous study.[3] We found that experimental data in literature databases and data available online are inconsistent and scattered. We also identified errors in the method evaluation in a previous, highly influential study.[3] After making the necessary corrections, we developed simple models for predicting protein folding rates based on the length of the protein chain. These simple models, in several comparisons, showed greater stability in prediction than almost all existing models from the literature. In addition, we proposed a physicochemical interpretation for these models with respect to transition state theory and polymer chemistry theory.[4]

Acknowledgements. This research is supported in part by the Croatian Ministry of Science and Education through basic grants given to their institutions, National Recovery and Resilience Plan (NPOO) [2258ZI726], and Croatian Government and European Regional Development Fund under the Competitiveness and Cohesion Operational Programme for the project Sustainable Bioprospecting of Organisms from the Adriatic Sea for Innovative Natural Products—BioProCro (PK.1.1.10.0001), Centre of Research Excellence.

REFERENCES

[1] D.N. Ivankov, A.V. Finkelstein, *Biomolecules* **2020**, 10, 250.

[2] A. Kraljević, J. Batista, V. Bojović, B. Lučić, *Int. J. Mol. Sci*. **2026**, 27, 2968.

[3] C. Chang, B. Tey, J. Song, R. Ramanan, *Brief. Bioinform*. **2015**, 16, 314–324.

[4] H. Eyring, *J. Chem. Phys*. **1935**, 3, 107–115.