MUNI

COMMENTARY TO HABILITATION THESIS[1]

Advanced Data Modelling for Protein Engineering

Stanislav Mazurenko

This habilitation thesis is a compilation of publications from the domain of protein engineering authored or co-authored by Stanislav Mazurenko. The methods introduced in the publications were developed between the years 2015 and 2025 mainly at Masaryk University, Brno. The main motivation for the methods was to design computational tools for the analysis and modelling of complex biological data. Such a toolset enables protein engineers to select protein targets, plan experiments, analyse the collected data, and formulate a hypothesis about biological phenomena of interest in a more informed way, drastically enlarging the space of possibilities while reducing the experimental effort required to explore this space. The developed methods focus on two main approaches: (1) low-parameter modelling of the data, primarily based on physical principles, and (2) machine learning-based modelling, which leverages existing data sets to identify useful patterns in the data.

The thesis is divided into two parts. The first part provides a commentary on the contributions the thesis is based on. In Chapter 1, we motivate the need for data modelling and computational tools to study proteins. Then Chapter 2 discusses the bottom-up low-parameter modelling, which leverages biophysical principles. The alternative, top-down modelling that is based on machine learning methods, is presented in Chapter 3. Finally, Chapter 4 summarises the presented methods and proposes directions for future research. The second part consists of 11 publications in which these contributions were introduced. In the first two papers, the author was responsible for the design of the study, developing the mathematical framework, and programming the algorithms. In the remaining articles, the author was designing and overseeing data modelling and machine learning contributions to the corresponding studies.

The list of publications is as follows (the asterisk indicates corresponding authors):

[1][2] **Mazurenko, Stanislav**; Kunka, Antonin; Beerens, Koen; Johnson, Christopher M.; Damborsky, Jiri; Prokop, Zbynek. Exploration of Protein Unfolding by Modelling Calorimetry Data from Reheating. SCIENTIFIC REPORTS. 2017, vol. 7. DOI: 10.1038/s41598-017-16360-y

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
50	0	75	50

[2] **Mazurenko, Stanislav**; Stourac, Jan; Kunka, Antonin; Nedeljkovic, Sava; Bednar, David; Prokop, Zbynek; Damborsky, Jiri. CalFitter: a web server for analysis of protein thermal denaturation data. NUCLEIC ACIDS RESEARCH. 2018, vol. 46. DOI: 10.1093/nar/gky358

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
75	0	75	50

---

[1] The commentary must correspond to standard expectations in the field and must include a brief characteristic of the investigated matter, objectives of the work, employed methodologies, obtained results and, in case of co-authored works, a passage characterising the applicant's contribution in terms of both quality and content.

[2] Bibliographic record of a published scientific result, which is part of the habilitation thesis.

[3] Stourac, Jan; Dubrava, Juraj; Musil, Milos; Horackova, Jana; Damborsky, Jiri; **Mazurenko, Stanislav*;** Bednar, David*. FireProtDB: database of manually curated protein stability data. NUCLEIC ACIDS RESEARCH. 2021, vol. 49. DOI: 10.1093/nar/gkaa981

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
10	25	25	25

[4] Kunka, Antonin; Lacko, David; Stourac, Jan; Damborsky, Jiri; Prokop, Zbynek*; **Mazurenko, Stanislav*.** CalFitter 2.0: Leveraging the power of singular value decomposition to analyse protein thermostability. NUCLEIC ACIDS RESEARCH. 2022, vol. 50. DOI: 10.1093/nar/gkac378

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
10	75	25	75

[5] Velecky, Jan; Hamsikova, Marie; Stourac, Jan; Musil, Milos; Damborsky, Jiri; Bednar, David*; **Mazurenko, Stanislav*.** SoluProtMutDB: A manually curated database of protein solubility changes upon mutations. COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL. 2022, vol. 20. DOI: 10.1016/j.csbj.2022.11.009

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
35	75	10	75

[6] Khan, Rayyan Tariq; Pokorna, Petra; Stourac, Jan; Borko, Simeon; Arefiev, Ihor; Planas-Iglesias, Joan; Dobias, Adam; Pinto, Gaspar; Szotkowska, Veronika; Sterba, Jaroslav; Slaby, Ondrej; Damborsky, Jiri; **Mazurenko, Stanislav*;** Bednar, David*. A computational workflow for analysis of missense mutations in precision oncology. JOURNAL OF CHEMINFORMATICS. 2024, vol. 16. DOI: 10.1186/s13321-024-00876-3

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
25	25	25	25

[7] Marques, Sergio M.; Kouba, Petr; Legrand, Anthony; Sedlar, Jiri; Disson, Lucas; Planas-Iglesias, Joan; Sanusi, Zainab; Kunka, Antonin; Damborsky, Jiri; Pajdla, Tomas; Prokop, Zbynek; **Mazurenko, Stanislav*;** Sivic, Josef*; Bednar, David*. CoVAMPnet: Comparative Markov State Analysis for Studying Effects of Drug Candidates on Disordered Biomolecules. JACS AU. 2024, vol. 4. DOI: 10.1021/jacsau.4c00182

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
0	35	10	35

[8] Velecky, Jan; Berezny, Matej; Musil, Milos; Damborsky, Jiri; Bednar, David*; **Mazurenko, Stanislav*.** BenchStab: a tool for automated querying of web-based stability predictors. BIOINFORMATICS. 2024, vol. 40. DOI: 10.1093/bioinformatics/btae553

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
0	35	10	50

[9] Vavra, Ondrej; Tyzack, Jonathan; Haddadi, Faraneh; Stourac, Jan; Damborsky, Jiri; **Mazurenko, Stanislav*;** Thornton, Janet*; Bednar, Dednar*. Large-scale annotation of biochemically relevant pockets and tunnels in cognate enzyme-ligand complexes. JOURNAL OF CHEMINFORMATICS. 2024, vol. 16. DOI: 10.1186/s13321-024-00907-z

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
0	35	10	25

[10] Kohout, Pavel; Vasina, Michal; Majerova, Marika; Novakova, Veronika; Damborsky, Jiri; Bednar, David; Marek, Martin; Prokop, Zbynek*; **Mazurenko, Stanislav***. Engineering Dehalogenase Enzymes Using Variational Autoencoder-Generated Latent Spaces and Microfluidics. JACS AU. 2025, vol. 5. DOI: 10.1021/jacsau.4c01101

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
0	75	10	75

[11] Khan, Rayyan Tariq; Kohout, Pavel; Musil, Milos; Rosinska, Monika; Damborsky, Jiri; **Mazurenko, Stanislav***; Bednar, David*. Anticipating protein evolution with successor sequence predictor. JOURNAL OF CHEMINFORMATICS. 2025, vol. 17. DOI: 10.1186/s13321-025-00971-z

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
0	50	10	50