

# Fluctuations and predictability in protein evolution

**Saverio Rossi**<sup>1</sup>, Leonardo Di Bari, Martin Weigt, Francesco Zamponi

<sup>1</sup>Sapienza University of Rome, Rome, Italy

Protein evolution is a fundamental driver of biological diversity, shaping the structure and function of proteins over billions of years through the accumulation of mutations under functional constraints. A central challenge is to understand to what extent evolutionary trajectories are predictable, given the pervasive role of epistasis - context-dependent interactions between mutations - that couples changes across sites.

In this talk, I will present a quantitative framework to investigate how such interactions generate dynamical heterogeneity and constrain predictability in protein evolution. Building on recent work using data-driven generative models of protein sequence landscapes (e.g., generative Potts models) to simulate protein evolution, I will first show how epistasis gives rise to emergent evolutionary time scales and heterogeneous mutational behaviors across sites [1].

I will then focus on recent results characterizing fluctuations in protein evolution using tools inspired by disordered systems in statistical physics [2]. By modeling evolution on data-driven fitness landscapes and analyzing spatio-temporal correlations, we disentangle distinct sources of variability: those inherited from the ancestral sequence and those arising from stochastic mutational dynamics. We find that, at short to intermediate times, evolutionary trajectories retain a strong memory of their ancestor, defining a characteristic persistence time scale over which ancestral information can be reliably reconstructed. This persistence is tightly linked to the strength of epistatic constraints encoded in the ancestral sequence. At longer times, however, this memory fades as evolution becomes dominated by collective, epistatically constrained dynamics, setting a fundamental limit to predictability.

Overall, this work highlights the emergence of distinct evolutionary time scales, both for individual sites and entire proteins, as a direct consequence of epistatic constraints. These results provide a quantitative framework to interpret how correlations build up over time in protein evolution, and have practical implications for evolutionary inference, including improving evolutionary predictions and assessing the reliability of ancestral sequence reconstruction methods.

## References

[1] L. Di Bari, M. Bisardi, S. Cotogno, M. Weigt, and F. Zamponi, "Emergent time scales of epistasis in protein evolution", *Proc. Natl. Acad. Sci.* (2024)

[2] S. Rossi, L. Di Bari, M. Weigt, and F. Zamponi, "Fluctuations and the limit of predictability in protein evolution", *Rep. Prog. Phys.* (2025)