A151 · Generalizable and scalable protein stability prediction with SPURS

Presenter: Ziang Li — Georgia Institute of Technology

Authors: Ziang Li, Yunan Luo

Session: Day 1 · Thu Oct 1, 2026 · 2:15–4:15 PM

Predicting how amino acid substitutions change protein thermostability is essential for understanding disease mechanisms and engineering robust proteins. Recent protein generative models achieve strong zero-shot performance across mutation-related tasks without task-specific supervision, yet this unsupervised capability remains underused for supervised stability learning. We present SPURS, a deep learning framework that rewires and integrates two complementary protein generative models, a protein language model and an inverse folding model, and reprograms the unified architecture for thermostability prediction using mega-scale $\Delta\Delta G$ data. SPURS injects structure-derived information into sequence-derived representations through lightweight cross-attention adapters and performs parameter-efficient fine-tuning, reducing trainable parameters by 98.5% relative to full language-model tuning. The model is also computationally scalable: it predicts $\Delta\Delta G$ for all single substitutions of a protein in one forward pass, enabling proteome-scale site-saturation analyses. As a concrete example, SPURS can scan the human proteome (19,652 proteins, about 10^9 variants) in about 30 minutes on one GPU. In benchmarking, SPURS improves over baselines on identity-controlled and external datasets, including better recovery of stabilizing mutations and generalization to unseen proteins and mutation contexts. Beyond stability prediction, SPURS supports broad protein-informatics applications, including unsupervised functional-site identification, improved low-N protein fitness prediction, and systematic analysis of stability-pathogenicity relationships in human variants. Together, these results establish SPURS as an accurate, efficient, and generalizable framework for protein stability modeling and as a tool for large-scale protein engineering and disease-variant interpretation.

Generalizable and scalable protein stability prediction with SPURS

Introduction. Predicting mutation effects on protein stability is central to protein engineering and human genetics, yet supervised predictors often generalize poorly because stability measurements are limited and biased. Protein foundation models provide complementary priors: sequence language models capture evolutionary constraints, whereas inverse-folding models encode 3D structural context. We developed SPURS to combine these modalities for accurate and transferable stability prediction.

Methods. SPURS (Stability Prediction Using a Rewired Strategy) integrates ESM2 and ProteinMPNN through lightweight cross-attention rewiring. Structure-derived representations from ProteinMPNN are injected into frozen ESM2 representations, while only ProteinMPNN, the adapter, and prediction head are optimized, reducing trainable parameters by 98.5% relative to full ESM2 fine-tuning. Given a wild-type sequence and AlphaFold-predicted structure, SPURS predicts $\Delta\Delta G$ for all single substitutions in one forward pass.

Results. We train SPURS on Megascale and evaluate on an identity-controlled test split ($< 25\%$ sequence similarity) and ten independent datasets spanning different assays and curation protocols. Baselines include FoldX, ThermoMPNN, and sequence- and structure-based generative models. SPURS improves generalization to unseen proteins and mutation contexts, outperforming ThermoMPNN on 24 of 28 Megascale test proteins and increasing median Spearman correlation from 0.77 to 0.83. On Domainome, correlation improves from 0.49 to 0.54. Its one-pass formulation enables proteome-scale scanning: approximately 10^9 single substitutions across 19,652 human proteins can be evaluated in about 30 minutes on one GPU.

Statement of significance. Beyond stability prediction, SPURS supports broader protein-informatics applications. Combining SPURS stability scores with language-model fitness signals enables unsupervised functional-site prioritization and recovers annotated catalytic and binding residues in human domains. SPURS also provides a mechanistically meaningful prior for low-N fitness prediction: on ProteinGym assays excluding stability-focused tasks, adding SPURS improves mutation-effect ranking, including an average 7% Spearman gain in the original low-N benchmark. Proteome-scale analyses further show stronger predicted destabilization among pathogenic than benign human variants. Together, these results establish SPURS as an accurate, scalable, and generalizable framework for protein stability modeling, protein engineering, and variant interpretation.

Publication availability: Published in *Nature Communications*: `https://doi.org/10.1038/s41467-025-67609-4`.

Code availability: `https://github.com/luo-group/SPURS`

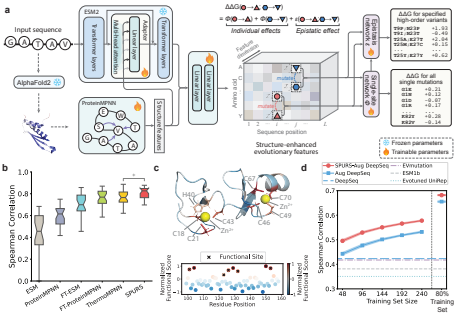

Figure 1: **Overview of SPURS. a**, sequence–structure model architecture; **b**, stability prediction on Megascale; **c**, functional-site identification; **d**, low-N fitness prediction.