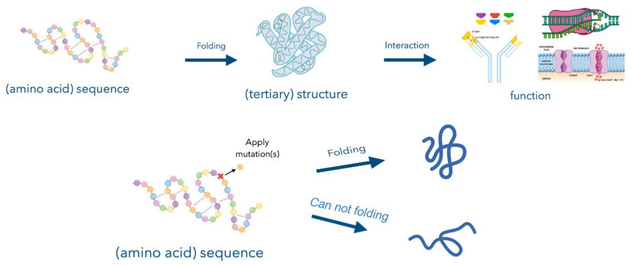
Machine Learning for Immune Receptor Discovery and Design

DEPARTMENT OF PHYSICS

Problem Statement

- Protein structure/function prediction is vital for learning of biological processes.
- Wet labs are time-consuming, labor-intensive, and expensive, they are not effective for high throughput screening.
- Current ML efforts focus on structure prediction from the sequence, but not mutational effects.
- There are high barriers to entry on using and implementing computation tools.

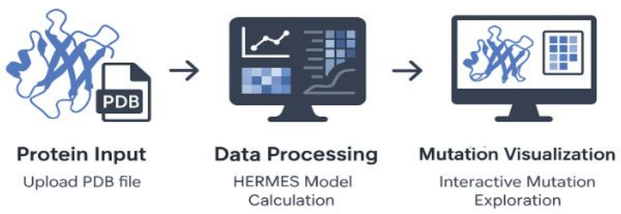


Our Solution

A platform built on the web that combines the HERMES ML model with a simple user interface:
On our Web-Based Platform

- Combines HERMES (High-accuracy Estimation of Residue Mutational Effects on Structure) with an easy to use interface for viewing of protein structures.
- Pre-computed database of common proteins.
- Make custom runs of predictions without knowing how to code.
- Interactive data exploration and visualization.

System Workflow

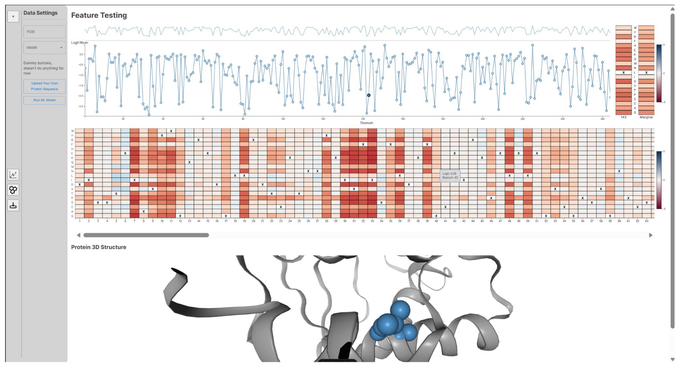


6-Month Development Milestones

- **February:** HERMES is running with data/run model on HYAK and configure the working environment.
- **March:** Website prototype/Design a database structure.
- **April-May:** Features for showing data/Create APIs/Adjus HERMES model/design of the web interface.
- **End of May:** Launched for testing/Performance optimization/User testing and user validation.

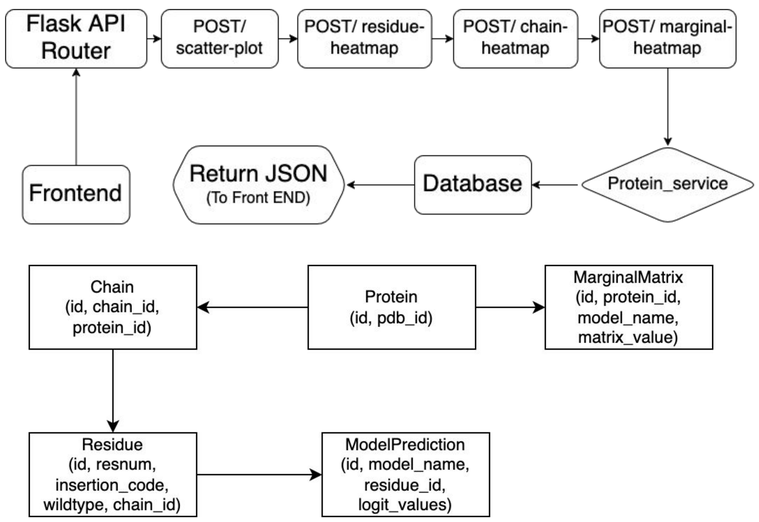
Frontend Implementation

- Interactive 3D visualization of protein structures.
- Scatter plots displaying predictive position-based scores from HERMES.
- Heatmaps for mutation analysis.
- D3.js for charts, NGL.js for molecular visualizations.



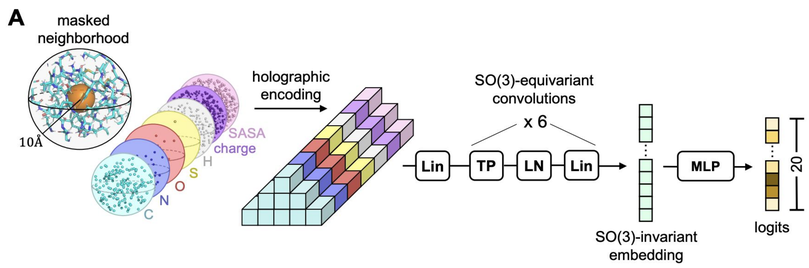
Backend Implementation

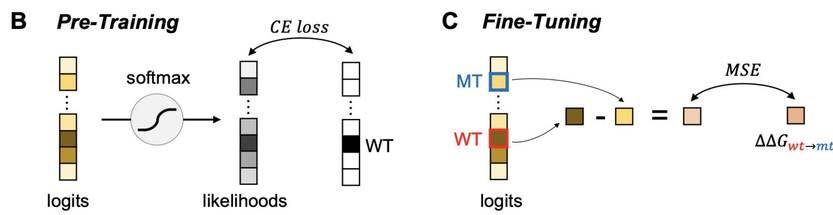
- RESTful API: Validates PDB ID and uploaded structures.
- HERMES Predictions: Fast SQLite retrieval by PDB, chain, residue.
- Structure Comparison: Uses AlphaFold to benchmark against experimental data.
- HYAK + AWS EC2: Scalable GPU compute for new structures.
- Covers all 20 possible amino acids.
- API outputs JSON with probabilities, caching, and compression.



Fine Tuning HERMES

- State of the art mutational prediction ML model
- SO(3)-equivariant convolutional networks that respect 3D rotational symmetries.
- Trained on huge protein structure datasets.
- Able to accurately predict stability shifts caused by a single mutation.
- Fine-tuning allows for generalizability of HERMES predictions





Results

- A platform built for functionality with an intuitive user experience.
- Ability to create 3D visualizations in real time.
- Lowers barriers of entry to require less non-domain specific knowledge.
- Enhances flexibility by providing unique models for specific use-cases and units.
- Allows quick visual identification of potential interesting protein sites and mutations.

Future work

- AlphaFold integration for sequence-to-structure-to-prediction pipeline.
- Expansion of pre-computed prediction database.
- Performance optimization for real-time analysis.
- Other modes of visualization for various research requirements.
- Supporting mobile device access.
- Live inference on unseen, user-provided proteins.

ELECTRICAL & COMPUTER ENGINEERING
UNIVERSITY of WASHINGTON

ADVISORS: Armita Nourmohammad, Jai Jaisimha, Gian Marco Visani, Kevin Borisiak, Ella Carlander

SPONSOR: Department of Physics