AI/ML-ASSISTED REDESIGN OF NATIVE PROTEINS

Computational Redesign Of Monomeric Native Protein Using Non-canonical Amino Acids

Doo Nam Kim Pacific Northwest National Laboratory **John Cort** Pacific Northwest National Laboratory **Richard Overstreet** Pacific Northwest National Laboratory **Katherine Graham** Pacific Northwest National Laboratory **Doug Renfrew** Flatiron Institute

Non-canonical amino acids (ncAAs) offer enhanced control over polypeptide backbone and sidechain configurations and interactions than the canonical amino acids. Therefore, we previously semi-automated the generation of ncAA rotamers and design input files and output analysis using the Bio Chemical Library (BCL) and Python Data Analysis Library (pandas), in the structural modeling software, Rosetta.

Recent breakthroughs in AI/ML-assisted protein structure predictions and design have enabled the generation of numerous schemes for idealized protein structure and function optimization. However, most of these are unable to model ncAAs. The sole AI/ML program capable of modeling ncAAs can only handle sidechains without any connection to the backbone and lacks reproducibility. We plan to highlight these limitations in our benchmark results. Additionally, we will discuss the feasibility of providing AI/ML training sets for ncAAs using our physics-based sampled protein structures.

Most redesigns of native proteins using ncAAs have focused on modifying the interface between protein/peptide multimers or on developing new probes. Therefore, we came up new hypotheses as follows: 1) We hypothesize that geometrically constraining globular peptides by introducing ncAAs of connecting loops will vary their thermal stability. 2) We propose that designing nearby oxygen by ncAA can enhance the lesser-known fluorine-induced sigma hole effect in peptides. 3) We aim to demonstrate that redesigning native protein monomers for stronger electrostatic interactions with ncAA can enhance their thermal stability without altering their native tertiary structures.

To explore the effects of varying physical loop lengths while keeping the number of loop residues constant (achieved by mutating ncAAs), we redesigned a tryptophan-cage (trp-cage, the smallest globular folding peptide) variant (TC10b, PDB ID: 2jof) into four variants. To test our sigma hole effect hypothesis, we have redesigned our previous successful trp-cage variant (TC10b W6(5-fluoro-Trp), validated by Non-Canonical Amino Acid Parameterization Engine for CHARMM Potential:NCAP) into two variants. We synthesized all six peptides, and began experimental assessment of the extent and effectiveness of these AA substitutions in maintaining the peptides' stable tertiary structure and to test our hypotheses. To evaluate their thermal stability and determine structure, we are using circular dichroism and NMR. Finally, we have redesigned our previous successful trp-cage variant (TC10b P12W, validated by NCAP) to enhance its electrostatic interactions. We ordered the synthesis of this peptide from a commercial company and are awaiting its delivery for experimental validation.

The range of present and emerging threats from chemical and biological sources is vast. As a result, it is crucial to broaden the chemical space within our repertoire. This expansion will facilitate the development of targeted solutions that are specific and high-affinity for various Chemical and Biological Defense (CBD) applications (e.g., antidotes, biomarkers, vaccines). Additionally, demonstrating that classic physics based ncAA rotamer exploration in semi-automated fashion can generate the same protein structure with desired stability will illuminate possibilities and foster collaboration among protein engineers.

We received advices from Benjamin Brown (Jens Meiler group in Vanderbilt University) for BCL application. Dr. Thomas Lane in Collaborations Pharmaceuticals introduced sigma hole effect. This work is supported by Defense Threat Reduction Agency.