A1

Nooriel Banayan Advisor(s): John Hunt

Computational Protein Engineering for Rapid Drug-Target Discovery

Protein crystallization is the gold-standard method for elucidating protein structures, specifically drug-target structures, in industry and academia, yet many proteins do not crystallize and there are currently no reproducible methods to crystallize recalcitrant proteins. Crystallization propensity, therefore, is an intrinsic property that depends on the structural and thermodynamic properties of proteins. To elucidate the structural features that give rise to high crystallization propensity, we data-mined the Protein Data Bank (PDB), analyzed 87,684 crystal structures and found arginine to be over-represented in crystal-packing interfaces. To this end, we developed a publicly available computational pipeline called "BulkArg" that uses protein homology analysis to introduce arginine substitutions at select sites to improve the crystallization propensity of recalcitrant proteins. We applied this method to three proteins--two of which crystallize readily and have structures available and one of which is recalcitrant with no publicly available structure—and in all cases, arginine substitutions significantly increased crystallization propensity leading to structure determination. Moreover, the greater the number of arginine substitutions at select sites, the more readily the proteins crystallized into larger and more robust crystals. Our observations indicate that "BulkArg" successfully introduces arginine substitutions at favorable sites which can make recalcitrant proteins crystallize while conserving the protein's structure. This method can significantly streamline the drug-target discovery process and have major impacts in medicine and science.