COMBATTING EMERGING BIOLOGICAL THREATS – PREPARING FOR THE FUTURE TODAY

Machine Learning Predictions Of Antibody:antigen Interactions

Andrew Ellington The University of Texas at Austin **Danny Diaz** The University of Texas at Austin **James Loy** The University of Texas at Austin

The ability to engineer antibodies to bind novel antigens is of critical importance in both chemical and biological defense. The urgency of this task has been highlighted by the SARS pandemic, where the rate of change of the viral spike has largely outstripped the ability of biotechnologists to generate therapeutic antibodies. We suggest that it may prove possible to use machine learning approaches to both engineer antibodies that can bind to novel antigen variants, and to use the selfsame technologies to anticipate mutations that may arise in both natural and engineered diseases.

We will present results relevant to the machine learning program MutCompute (mutcompute.com), which relies on 3D convolutional neural networks (3D CNNs) and residual networks (ResNets) to accurately predict mutations that improve functionality in virtually any protein, and that have now been adapted to predict improved functionality in antibody:antigen interactions. In short, MutCompute queries the microenvironments of individual amino acids, based on the vast amount of structural data from the Protein DataBank (PDB). When amino acids are individually 'subtracted' from a protein structure, the remaining cavity can be characterized in terms of surrounding atom types (hydrogen, carbon, nitrogen, sulfur), partial charges, solvent accessibility, and even molecular dynamics. In doing so, MutCompute has an astounding 70% success rate for re-predicting wild-type amino acids across the PDB. However, what is more interesting is the predictions made for the remaining 30% of positions, where MutCompute will often predict improved fits to a given microenvironment, frequently leading to improved protein functionality. This approach has been applied to the stabilization and improved kinetic properties of enzymes as diverse as fluorescent proteins, polymerases, and plastic-degrading esterases.

We now show that these methods can be adapted not just within a protein, but between proteins. We will present results on the prediction and engineering of therapeutic antibodies for improved interactions with the viral spike from SARS-CoV-2, and will also speak to the ability to predict interactions and mutations between viral antigens and their receptors.