## BROAD-SPECTRUM THERAPEUTICS FOR VIRAL DISEASES: A MEDICAL COUNTERMEASURE PLATFORM FOR EMERGING THREATS

# A Computational Pipeline To Identify Broad-spectrum Drug Targets And Interacting Chemotypes In Viral Pathogens

**Sarah Sandholtz** Lawrence Livermore National Laboratory **Jeffrey Drocco** Lawrence Livermore National Laboratory **Adam Zemla** Lawrence Livermore National Laboratory **Marisa Torres** Lawrence Livermore National Laboratory **Mary Silva** Lawrence Livermore National Laboratory **Jonathan Allen** Lawrence Livermore National Laboratory

Minimizing the threat of biological agents requires the ability to develop and deploy effective treatments for novel pathogens as soon as they emerge. Yet drug discovery is a notoriously slow process, and poor target selection often contributes to early failures in the pipeline. An accurate and efficient computational method for rigorously identifying drug target sites in emerging pathogens could reduce the time and cost of experimental screenings and serve a critical purpose in accelerating drug development.

We introduce a unique framework for the rapid identification of conserved binding sites in the proteins of novel viruses as well as the core chemical components with which these sites interact. Our fast, quantitative, and fully automated approach combines molecular-level structural modeling of proteins with data scientific and cheminformatic techniques and generates results at a fraction of the computational cost of more traditional simulation methods. We leverage the program PDBspheres, developed in-house and based on experimental crystal structures from PDB, to implement homology-based structural modeling of viral pathogen proteins and predict protein-ligand binding. Using the protein-ligand binding data from PDBspheres, we apply a clustering procedure to group protein residues based on their ligand contacts and select groups of residues that form relevant binding sites. A separate clustering procedure allows us to group the ligands that bind to each site based on their chemical structure and function, thus revealing the site-specific, key chemical features involved in binding.

Favorable comparisons with experimental data and other computational studies provide support for the effectiveness of our predictive framework. Using SARS-CoV-2 as a test case, we find that our results are not particularly sensitive to the amount of SARS-CoV-2 data in PDB, suggesting that our process can perform reasonably well even early in a pandemic when structural data on a new pathogen is limited. We also find that the composition of structural models for individual binding sites may be used as a simple heuristic for determining which sites are more conserved and therefore are more effective as broad-spectrum antiviral drug targets. While we present here a demonstration of our tool on SARS-CoV-2, our process is generalizable and can be applied to any new virus, as long as its genome sequence is available or sufficiently accurate homology models can be constructed.

Approved for public release; distribution is unlimited.