

Bioinformatics and 3D Structural Analysis of the Coronavirus Main Protease Active Site

Amy Wu Wu¹, Mickayla Bacorn², MaryAgnes Balogun³, Cassandra Olivas⁴, Christine Zardecki⁵, Sagar Khare⁶, Stephen Burley⁷, Joseph Lubin⁸

¹University of Puerto Rico, Mayagüez Campus ²NA, ³Morgan State University, ⁴California State University, Stanislaus, ⁵Rutgers Proteomics, RCSB Protein Data Bank, ⁶Institute for Quantitative Biomedicine, ⁷RCSB Protein Data Bank, Rutgers University

amy.wu@upr.edu

Coronaviruses (Coronaviridae) such as SARS-CoV-2 (severe acute respiratory syndrome coronavirus) and MERS-CoV (Middle East respiratory syndrome coronavirus) have been the source of recent outbreaks and global health concerns. While vaccines have been essential for controlling the SARS-CoV-2 (COVID-19) pandemic, it is uncertain whether they will be effective against future coronavirus strains. Therefore, identification or design of a broad-spectrum drug that targets highly conserved regions of the main protease of multiple coronavirus strains is essential in the long term. As part of a virtual summer research experience with the RCSB PDB, bioinformatics tools were employed to predict and construct 3D models of the coronavirus main protease (MPro) using SARS-CoV-2 as the template, with a focus on mutational trends and active sites. This study focused on the active sites of MPro, a cysteine protease essential for viral assembly and replication. Sequence alignments and structure modeling of MPro structures has identified conserved regions across multiple coronavirus strains. Inhibition of MPro halts coronavirus replication, making it an ideal drug target, and studies of MPro may foster and accelerate the discovery of high affinity broad-spectrum drugs.

This work was supported by an NSF REU to RCSB PDB. RCSB PDB is funded by the National Science Foundation (DBI-1832184), the US Department of Energy (DE-SC0019749), and the National Cancer Institute, National Institute of Allergy and Infectious Diseases, and National Institute of General Medical Sciences of the National Institutes of Health under grant R01GM133198.