

## Understanding the evolution of SARS-CoV-2 non-structural protein 15 (Nsp15) in three dimensions (3D)

S Stagers<sup>1</sup>, C Zardecki<sup>2</sup>, S Khare<sup>3</sup>, S Burley<sup>4</sup>

<sup>1</sup>1999, Lanham, MD, <sup>2</sup>Rutgers Proteomics, RCSB Protein Data Bank, Piscataway, NJ, <sup>3</sup>Institute for Quantitative Biomedicine, <sup>4</sup>RCSB Protein Data Bank, Rutgers University, Piscataway, NJ  
srstagers0@frostburg.edu

SARS-CoV-2 is the coronavirus responsible for the global COVID-19 pandemic. There is currently little information on SARS-CoV-2 and the proteins encoded by its positive-strand RNA genome. Our research utilizes computational structural biology tools to explore the evolution of SARS-CoV-2 proteins in 3D. As part of a virtual summer research experience with the RCSB PDB, we studied how SARS-CoV-2 proteins evolved during the first six months of the COVID-19 pandemic by exploring amino acid sequence and 3D atomic-level structure using various structural bioinformatics tools, including Clustal Omega ([www.ebi.ac.uk/Tools/msa/clustalo/](http://www.ebi.ac.uk/Tools/msa/clustalo/)) for sequence alignments and phylogenetic trees; Mol\* ([molstar.org](http://molstar.org)) for 3D molecular visualization; and Foldit ([fold.it](http://fold.it)) for structural/energetic effects of sequence mutations. Non-structural protein 15 (Nsp15) occurs near the C-terminus of the SARS-CoV-2 polyprotein 1ab that is expressed during the early stages of viral infection. Following cleavage from this polyprotein by the Main Protease (Nsp5), Nsp15 assembles into a symmetric hexamer that functions as a uridylylate-specific endoribonuclease. While not essential for viral replication in the laboratory setting, conservation of Nsp15 amongst coronaviruses implies that it plays important roles in the virus life cycle. More than 300 unique mutant forms of Nsp15 detected during the first six months of the pandemic were analyzed to determine differences from the wild-type protein. Investigation of the mutation data could provide insights relevant to discovery and development of drugs to combat SARS-CoV-2. This work was supported by an NSF REU. 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.

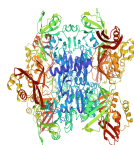


Figure 1