Computational Efforts in Development of COVID-19 Therapeutics

Research Synopsis

Center for Advanced Study, Aspiring Scholars Directed Research Program
46307 Warm Springs Blvd., Fremont, CA 94539

Contributing High School Researchers (by seniority and alphabetically):
Stephanie S.[1,16], Alex L.[2,15], Andrew S.[3,16], Anvi S.[4,16], Atri S.[4,16], Audrey K.[5,16], Ayush B.[4,16], Bhavesh A.[6,16], Brian F.[7,16], Ishani A.[2,16], Kavya A.[8,16], Karankumar M.[2,16], Pranjal V.[8,16], Tyler S.[2,16], Varsha B.[4,16], Allen C.[2,16], Aishwarya Y.[7,16], Andrew L.[9,16], Ansh R.[2,16], Ayeeshi P.[2,16], Jeslyn W.[2,16], Karthikha S. I.[2,16], Krithikaa P.[8,16], Ria K.[2,16], Sania K.[2,14], Simrun S.[9,16], Sophia F.[8,15], Aashi S.[6,16], Ananya V.[8,15], Anushka R.[6,16], Harsha R.[10,16], Ishya M.[6,15], Kara T.[5,16], Nambita S.[11,15], Rohan A.[12,14], Saahil D.[12,15], Saira H.[2,16], Sarah S.[13,16], Shamita B.[12,16], Shloka R.[6,16], Tanish S.[12,16]

Corresponding Investigators:
Robert A. Downing[14] robert.downing@fremontstem.com
Peter Le[15] peter.le@fremontstem.com
Soumya Suresh[15] soumya.suresh@fremontstem.com
Edward Njoo[16] edward.njoo@fremontstem.com

[1]Basis Independent High School Silicon Valley, *San Jose, CA*
[2]Mission San Jose High School, *Fremont, CA*
[3]Foothill High School, *Pleasanton, CA*
[4]The Quarry Lane School, *Pleasanton, CA*
[5]Dublin High School, *Dublin, CA*
[6]Amador Valley High School, *Pleasanton, CA*
[7]American High School, *Fremont, CA*
[8]Dougherty Valley High School, *San Ramon, CA*
[9]The College Preparatory School, *Oakland, CA*
[10]Fremont Christian High School, *Fremont, CA*
[11]Crystal Springs Uplands School, *Hillsborough, CA*
[12]Irvington High School, *Fremont, CA*
[13]Los Altos High School, *Los Altos, CA*
[14]Department of Computer Science & Engineering, *Center for Advanced Study, Aspiring Scholars Directed Research Program*
[15]Department of Biological, Human, & Life Sciences, *Center for Advanced Study, Aspiring Scholars Directed Research Program*
[16]Department of Chemistry, Biochemistry & Physical Science, *Center for Advanced Study, Aspiring Scholars Directed Research Program*

Synopsis

On March 11, 2020, the World Health Organization officially identified the outbreak of SARS-CoV-2 as a global pandemic. Currently, over 3.6 million people have been infected with the coronavirus, and there have been over 250,000 deaths worldwide. As a result of its global

impacts, the coronavirus has become a focal point for research groups around the world. It is imperative that we focus our attention and resources to the research and development of potential therapeutic methods that counteract the rapid spread of COVID-19. Even as a group of high school students, we are able to use our resources to contribute to the resolution of this global crisis. Over the last several months, our groups have turned to a variety of computational means towards discovery and development of molecular therapeutics with potential impact in the treatment of COVID-19. Below are a list of some of the original research work we have been involved in over the last semester.

Research #1: High throughput virtual screening of known compound libraries identifies potential lead compounds as ligands to the ACE2-SarsCoV2 spike glycoprotein assembly

Currently, few treatments have shown significant efficacy to fight the urgent COVID-19 pandemic. Efforts towards drug repurposing have gained attention and become increasingly popular due to lower costs, faster pace of development, and decreased risk from undesired side effects compared to the process required for novel drugs. We are working towards studies on repurposing current drugs—whose capabilities are already well understood and show promise through computational molecular docking as an efficient means to find possible drugs that target SARS-CoV-2.

Our study focuses on identifying possible inhibitors through *in-silico* screening of previously developed drugs, such as hydroxychloroquine, emodin and remdesivir, and similar compounds. Ligands that are similar in biochemical structure to known drugs were selected and screened for relative binding affinities with the spike glycoprotein. These studies can provide insight into potential small molecules that can prevent the binding of the spike protein with the ACE2 receptor.

Another way that many have attempted to identify inhibitors of SARS-CoV-2 is by screening large libraries of compounds from ZINC in hopes of identifying possible drug therapies. While it may produce effective results, this method is very time-consuming and may prove to be overwhelming given the myriad of data involved. To avoid this problem, we utilized the principles of chemical reactivity to more effectively design molecular compounds that may potentially inhibit the activity of SARS-CoV-2.

Research #2: Reactivity-guided design and in silico screening of targeted compounds towards strategic therapeutic approaches in treating SARS-CoV-2

Here, we employ a background in synthetic organic chemistry and reactivity of small molecules to design and screen chemical inhibitors of two distinct targets in SARS-CoV-2: the main protease and the spike glycoprotein. The main protease is essential for the replication of viral RNA, which is what allows the virus to replicate inside a host cell. The spike glycoprotein allows the virus to enter human cells, where it uses the cell's resources for viral replication. The molecular design of our small molecule inhibitors was based on known reactions and driven by the interactions within the active site of key proteins in SARS-CoV-2.

In our design of covalent inhibitors of the coronavirus protease, we modeled a library of 384 peptidomimetic Michael acceptor small molecules, which are designed to engage the nucleophilic cysteine residue (Cys145) in the active site of the protease in an irreversible 1,4-conjugate addition. We then employed a variety of computational tools to determine binding affinity of our designed compounds when bound to the protease active site.

Entry of the coronavirus into human cells is mediated through interaction of the spike glycoprotein with the human ACE2 receptor. Our approach in targeting the spike glycoprotein involves using a core scaffold with threefold symmetry, designed to engage the trimeric spike glycoprotein through noncovalent interactions, with the aim of "locking" the spike glycoprotein in the closed conformation, thereby potentially inhibiting viral interaction with the human ACE2 receptor, and, by extension, the virus' ability to enter host cells.

Designed compounds are screened through *in-silico* methods—and results acquired from density functional theory, molecular docking, and molecular dynamics simulations not only provide valuable insight into the biological activity of these compounds but can also guide the design of future small molecule inhibitors of SARS-CoV-2. We will soon be reporting structure-activity relationships (SAR) of our designed inhibitors to both the SARS-CoV2 protease and spike glycoprotein trimer.

Research #3: Development of a versatile machine learning platform for rapid identification of and computer-guided design of novel chemical entities with potential antiviral activity

Here, we employ machine learning and spatial binding analysis to identify potential chemical structures to inhibit a key protein in SARS-CoV-2. A training data set is developed by analyzing the molecular descriptors of compounds and the subsequent trends that are identified. The algorithm would then compare the molecular descriptors of potential compounds with the training set and predict whether the molecule and the accuracy of these predictions would be confirmed by molecular docking. The goal for the machine learning algorithm is to be able to correctly identify possible inhibitors from an intentionally-chosen, unbiased training dataset, from a learning dataset of SARS-CoV-2 inhibitors, demonstrating its ability to correctly identify the requisite three-dimensional structure of a molecule under question.

Further, we have generated a 3-D map of the key residues on the SARS-CoV-2 S-glycoprotein receptor-binding domain (RBD), which is the active site responsible for binding to the human ACE2 receptor. This will allow for potential screening of molecules simply by their 3-D conformations. Molecules will be scored by their binding conformations and potential hydrogen bonding or pi stacking interactions. This 3-D mapping function will allow for the visualization of potential interactions between the ligand and RBD without having to spend resources on DFT and docking calculations.

Research #4: Identification of polynucleotide sequences for potential application in ddRNAi (DNA-directed RNA interference) gene-silencing approaches for Sars-CoV2 infected cells

ddRNAi, or DNA-directed RNA interference, is a gene-silencing method that uses DNA constructs to hijack the pre-existing animal cell's RNA interference pathways. Here, we developed an RNA fragment that may silence a gene in the SARS-CoV-2 genome, which inhibits viral replication, thereby inhibiting the virus' spread. We designed a novel ddRNAi strand that inhibits gene expression in SARS-CoV-2 infected cells. Our siRNA strand does not share any complete similarities with the known human genome and corresponds to the SARS-CoV-2 virus genome at 9841-9859 bp.

Conclusion

The end goal of our projects is to ultimately provide a more in-depth understanding of the SARS-CoV-2 virus and offer potential methods that will effectively combat its spread. All groups are working towards a publishable unit and submitting manuscripts to high school journals. The first phase of this work is currently being wrapped up, with immediate plans for submission for publication of our findings in the coming weeks. The material above represents all information that is releasable, for the time being. More information will be made publicly available as the intellectual property described becomes released in published form.