SARS-CoV-2 drug discovery based on intrinsically disordered regions

Anish Mudide
Phillips Exeter Academy

Mentor: Gil Alterovitz

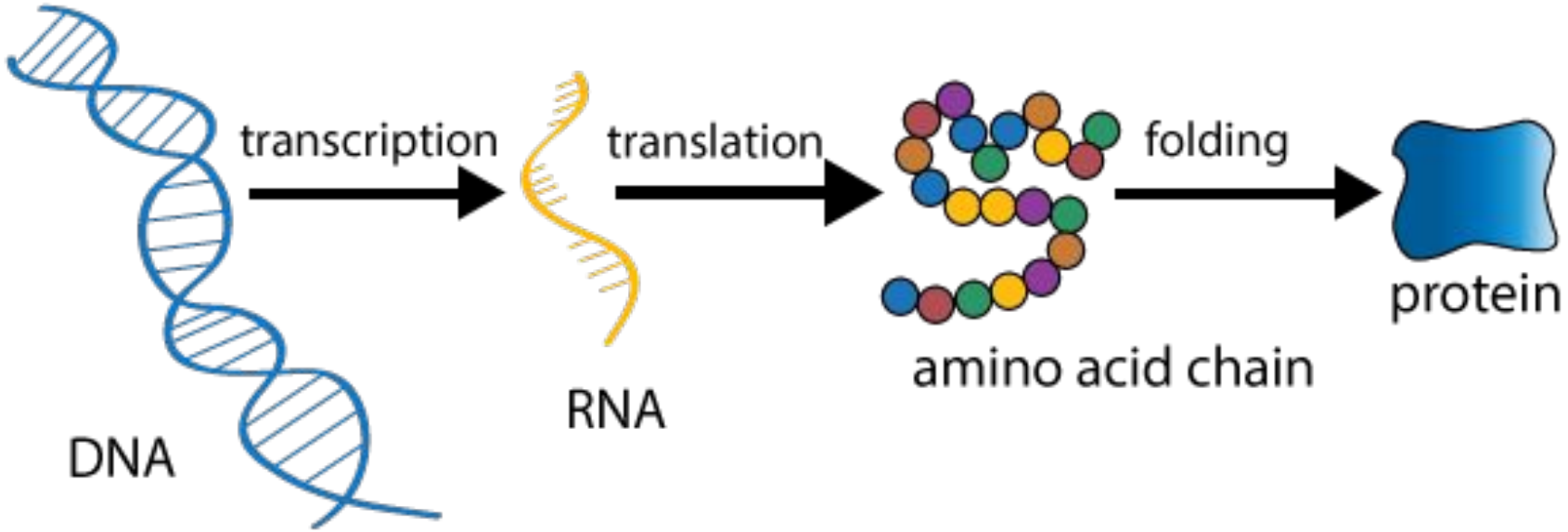
MIT PRIMES Fall Conference 2020

Agenda

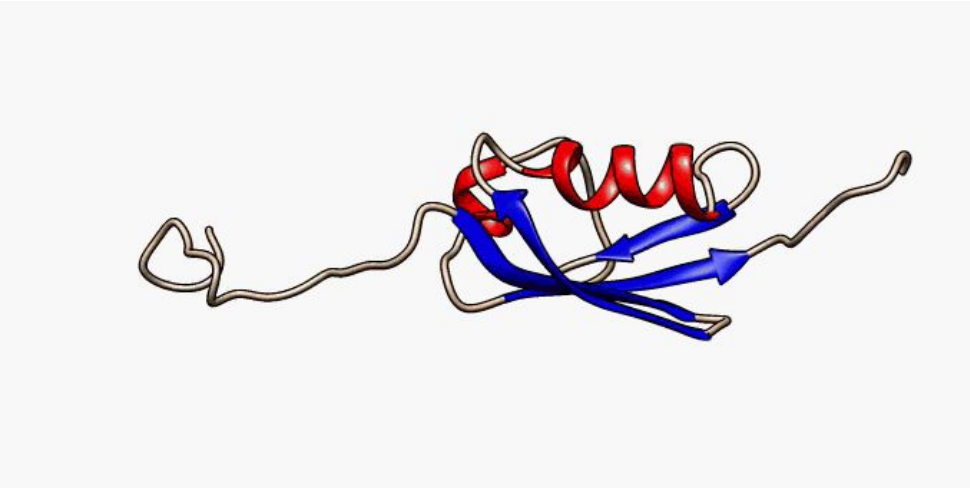
- Biological background
- SARS-CoV-2 therapeutic discovery
- Biomimicry-based drug discovery

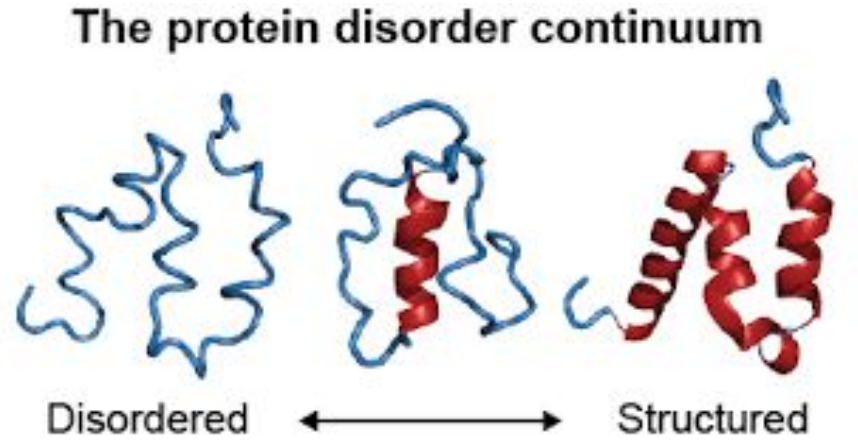
Background

Central dogma of molecular biology

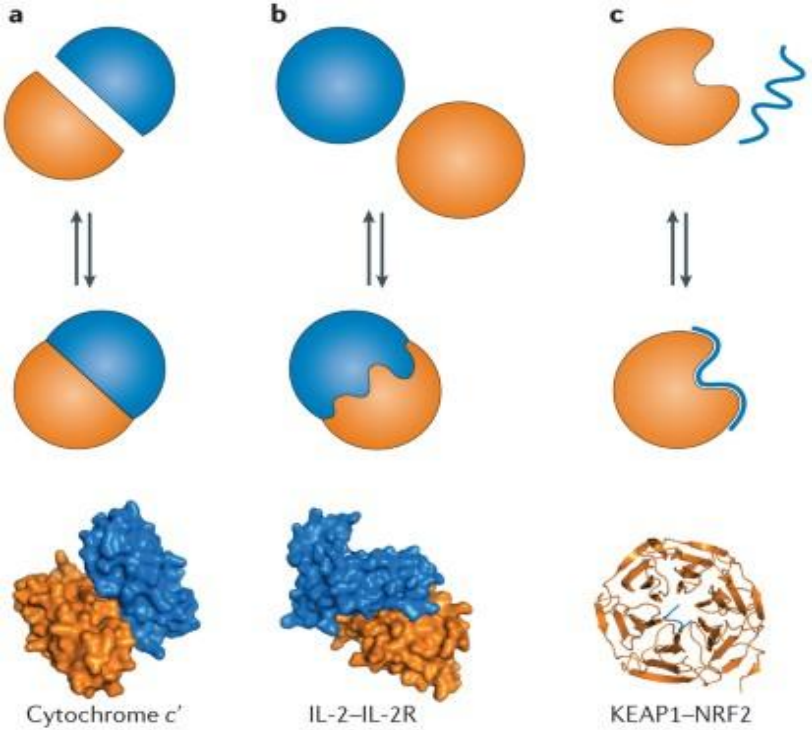


Intrinsically Disordered Proteins / Regions (IDPs / IDRs)





Drug Discovery



Why target IDRs?

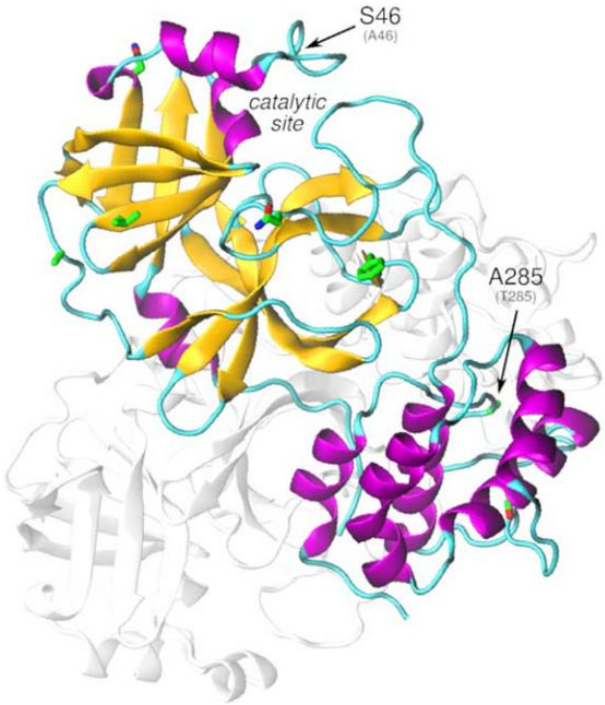
Common (33%)

Drug resistance

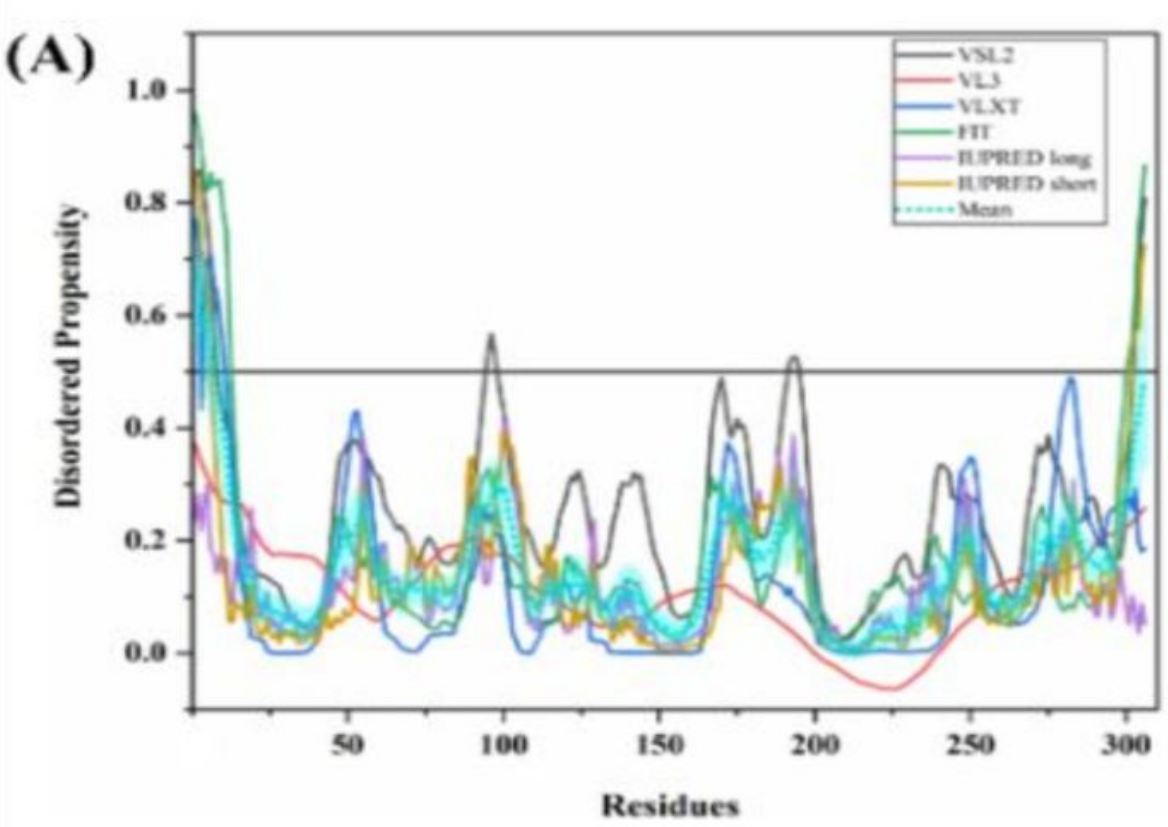
New opportunity

SARS-CoV-2 therapeutic discovery

The target: 3CL-protease



3CLpro disorder

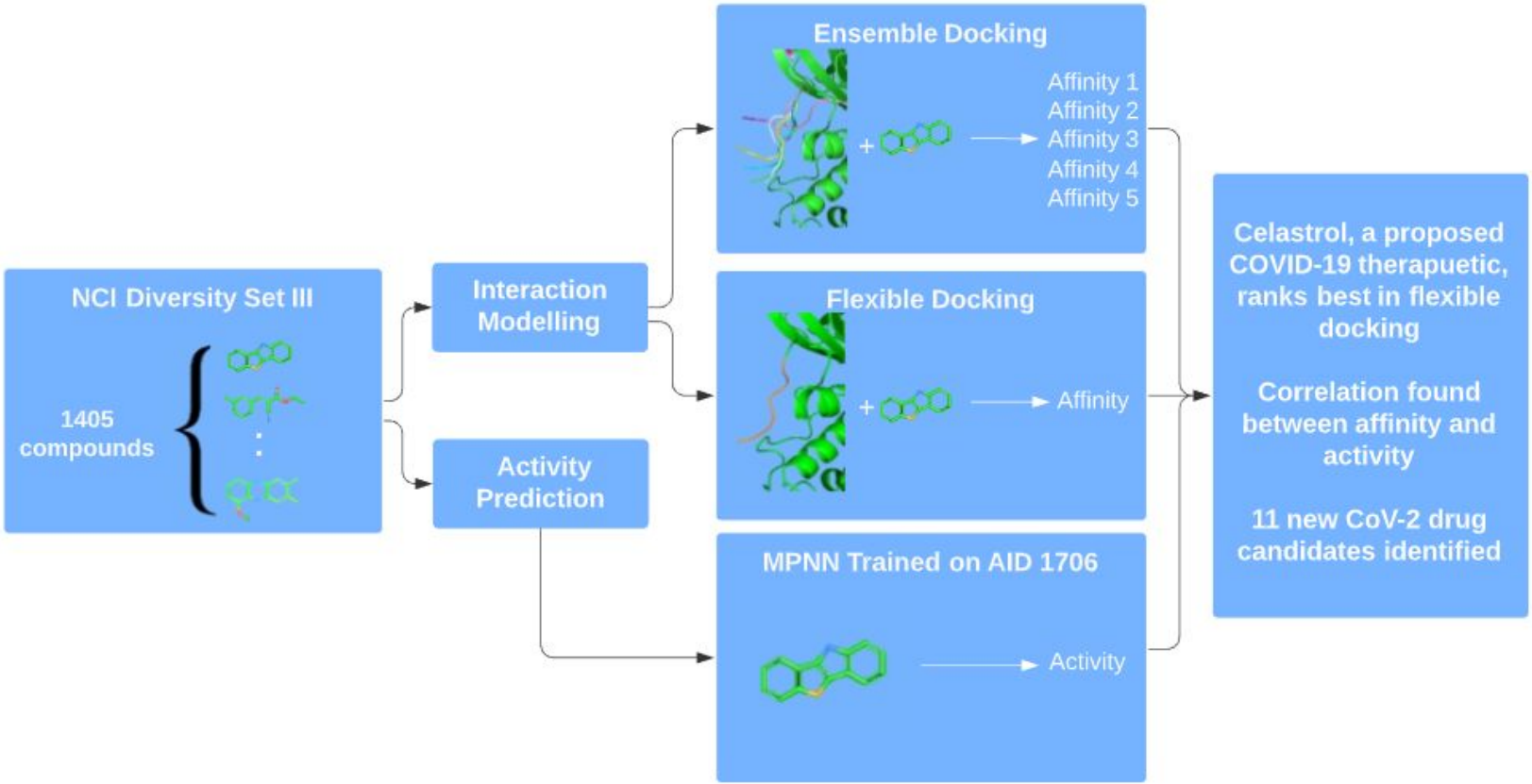


IDRs

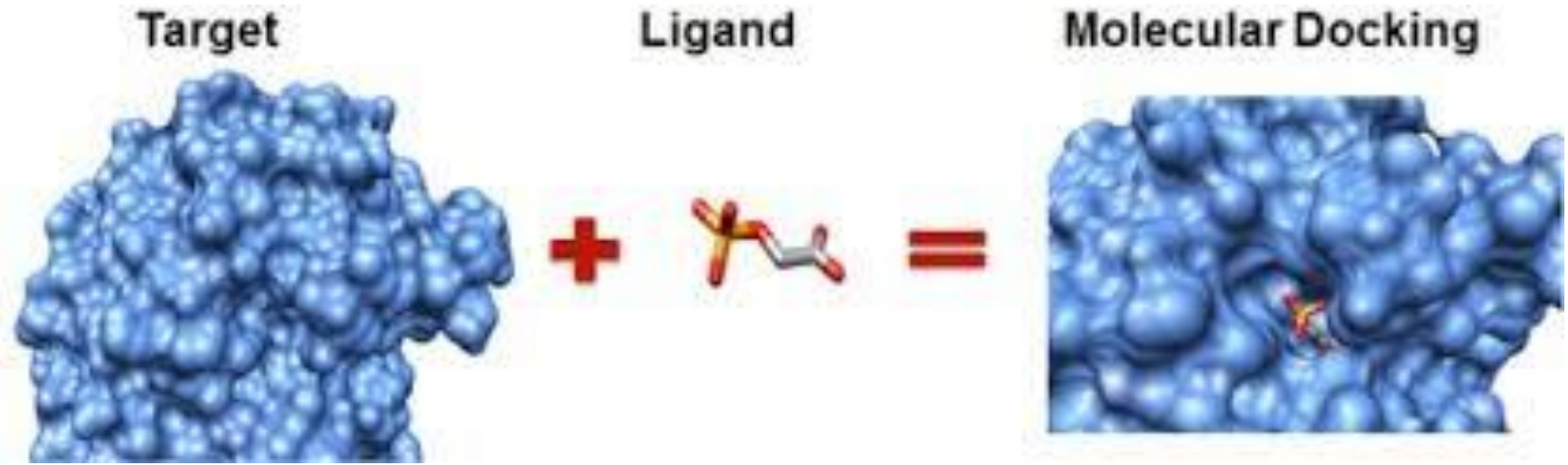
Residues 1-6*

Residues 302-306

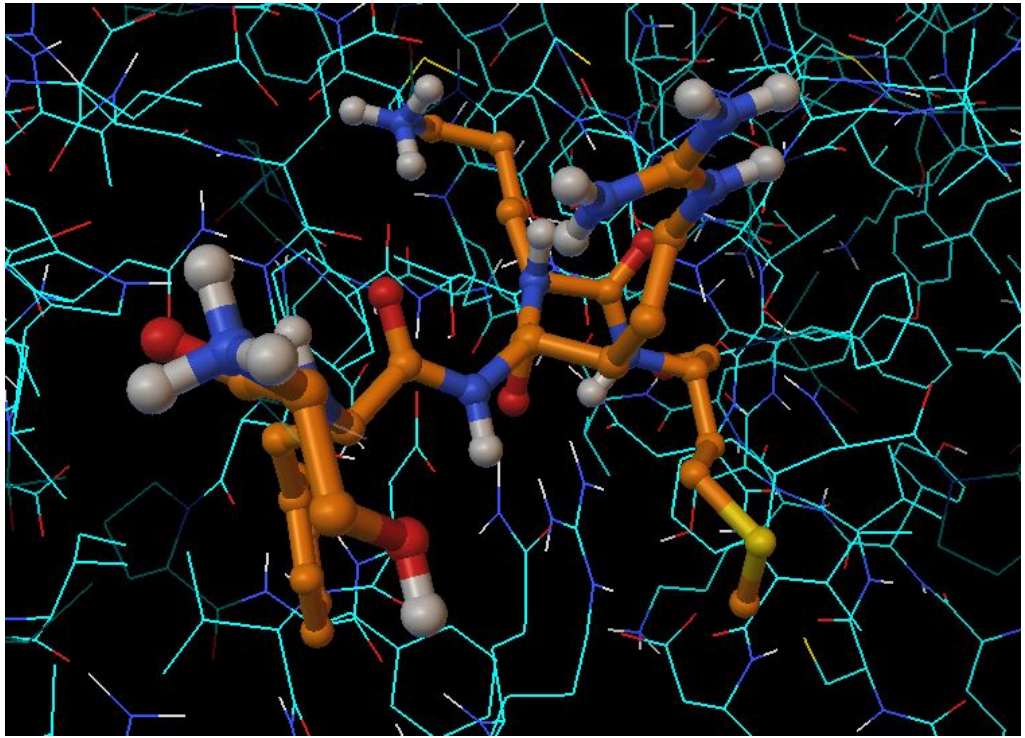
Our Approach



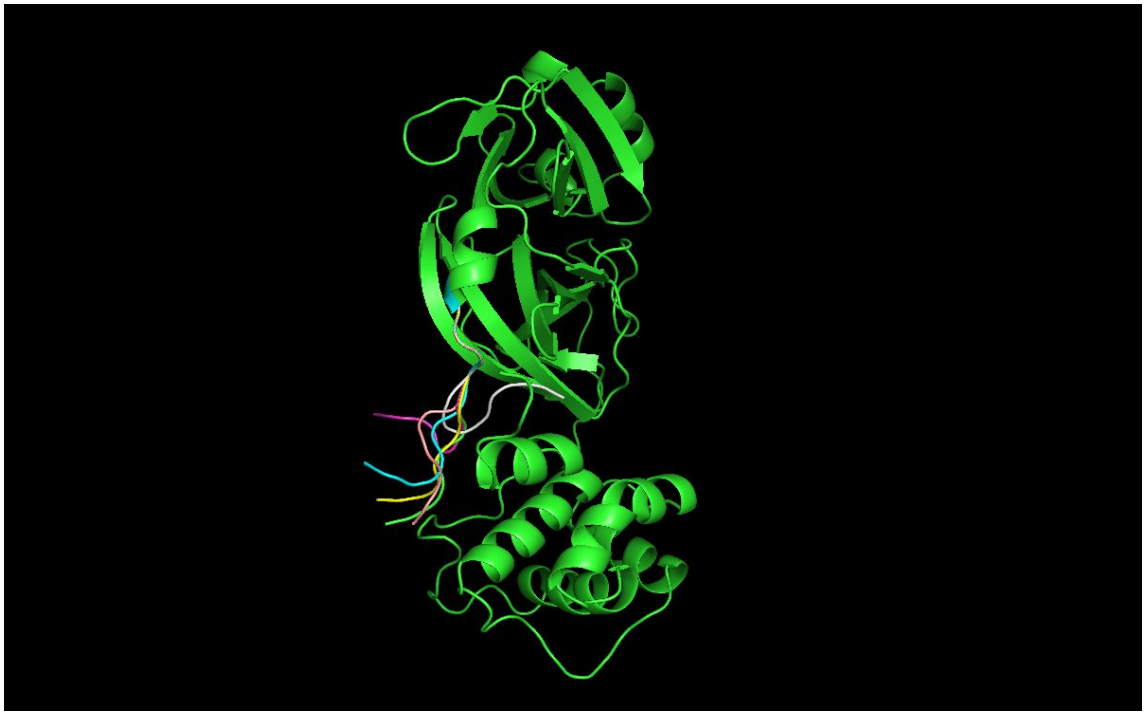
Interaction Modelling (Docking)



Independent Modification 1: Flexible side chains



Independent Modification 2: Ensemble docking



Results

Table 1. Binding affinity results from flexible docking (abridged)

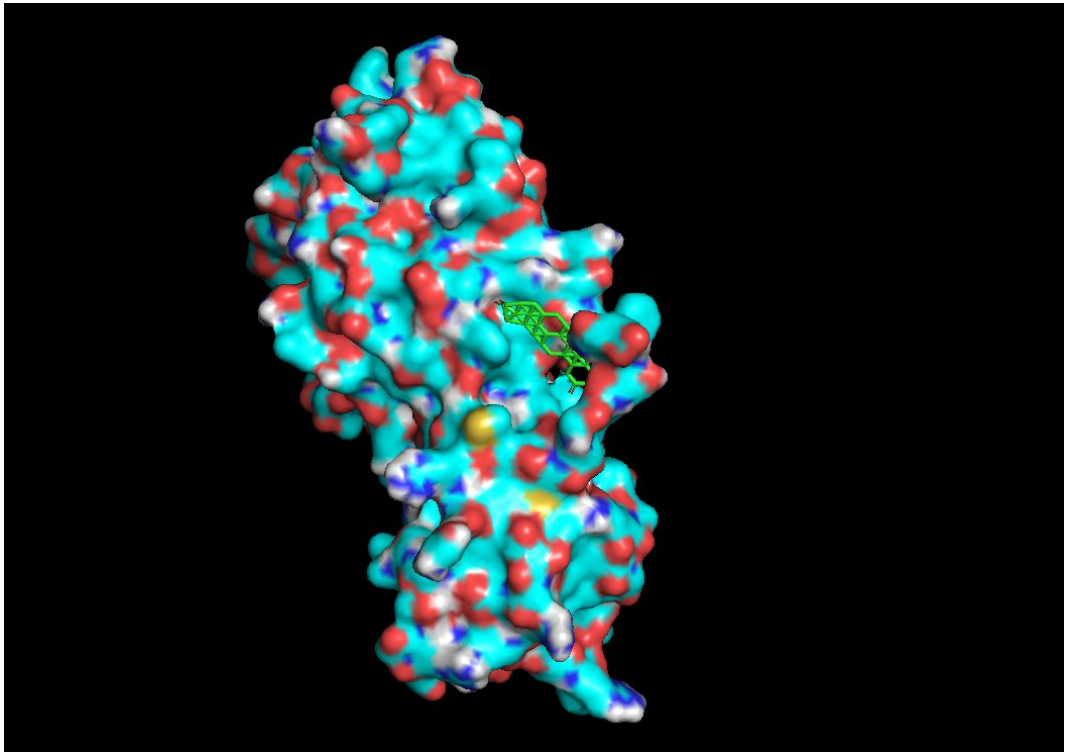
Molecule (NSC)	Binding Affinity (kcal/mol)
70931	-9.8
177862	-9.7
16437	-9.3
96541	-9.1
117987	-8.8
45527	-8.8
…	…

Results (cntd.)

Table 2. Binding affinity results from ensemble docking (abridged)

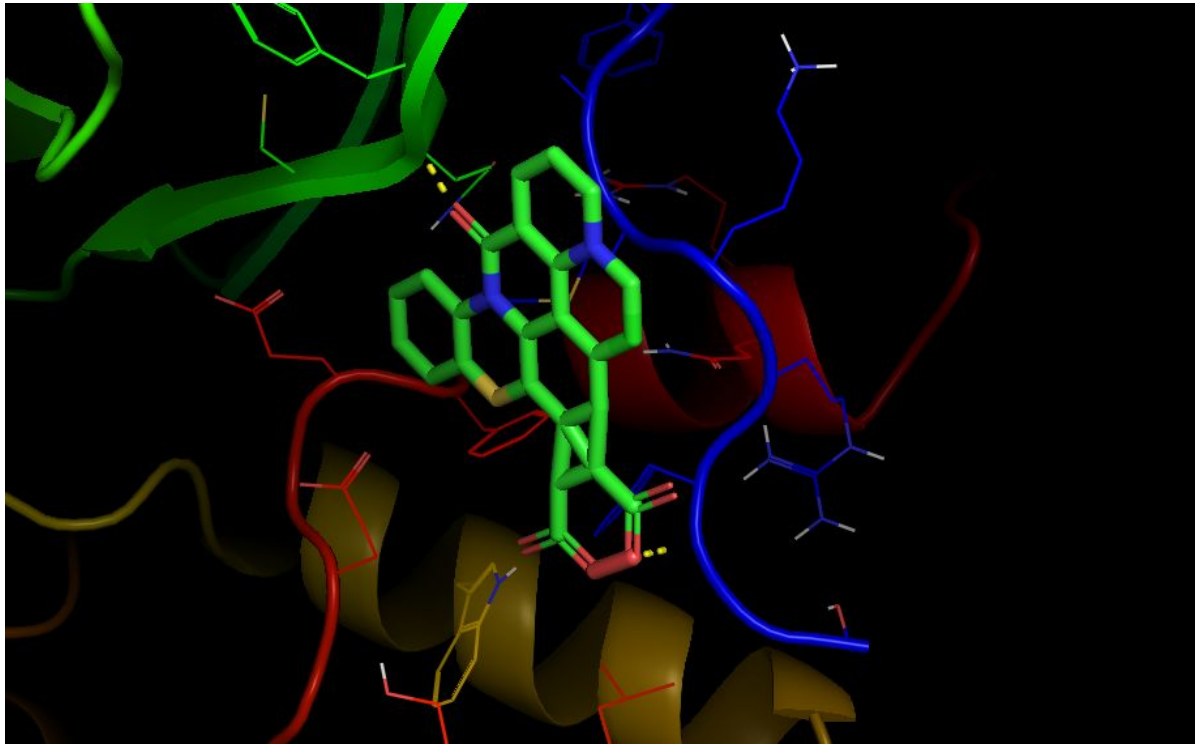
Molecule (NSC)	Best Binding Affinity (kcal/mol)
166259	-9.3
37641	-9.1
121868	-9.1
727038	-9.1
117987	-8.7
70931	-8.6
…	…

Visualizations



Molecule 70931 / Celastrol

Visualizations (cntd.)



Molecule 166259

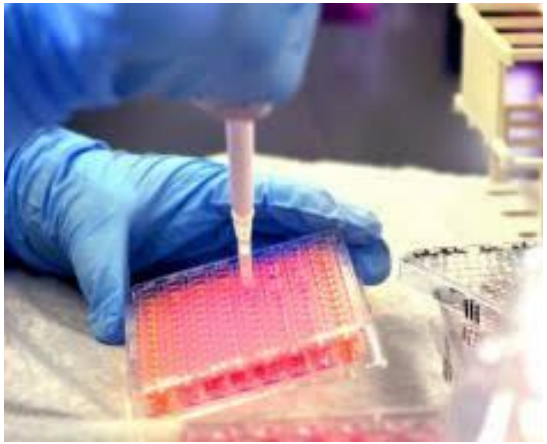
Cross-verification

In silico: docking

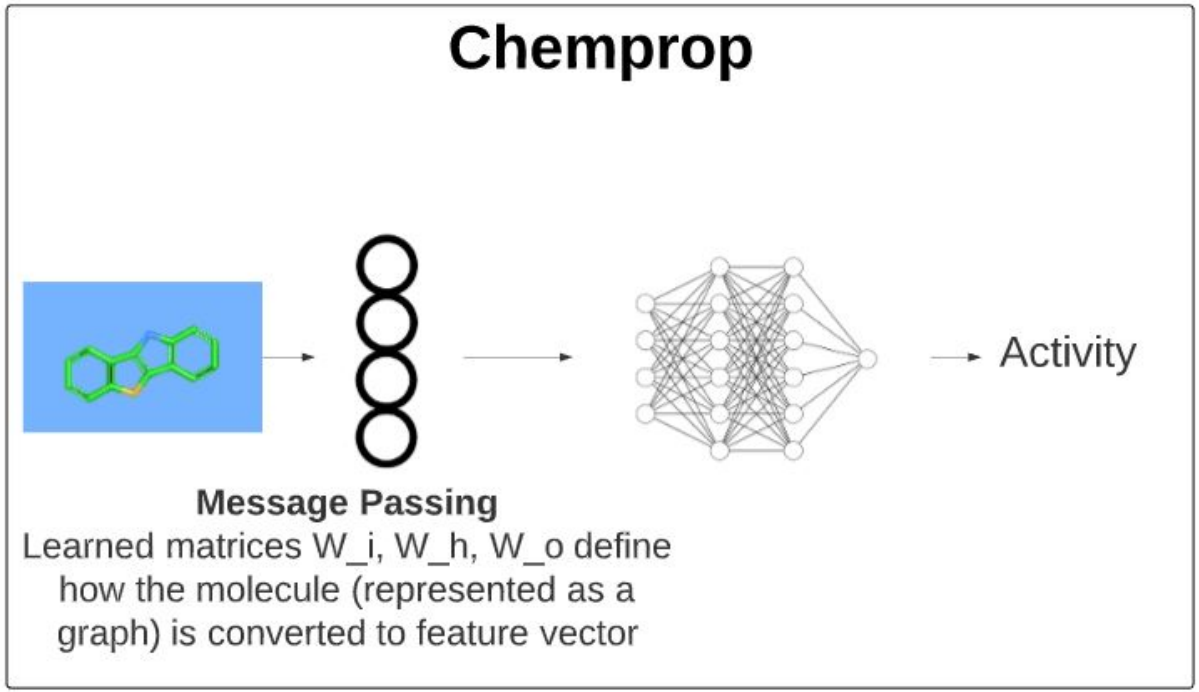
In vitro: AID 1706

In-vitro assay. Detects inhibition of SARS-CoV 3CL protease

N = 290,726



The model



Learns task-specific features, rather than fixed ones
Model is trained end-to-end

Results

Metrics + hyperparameters

80% train, 10% val, 10% test

Balanced training

Test AUC: .739

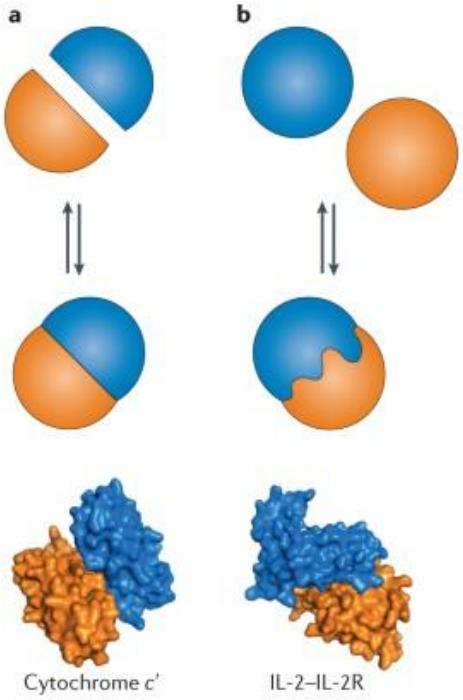
Results

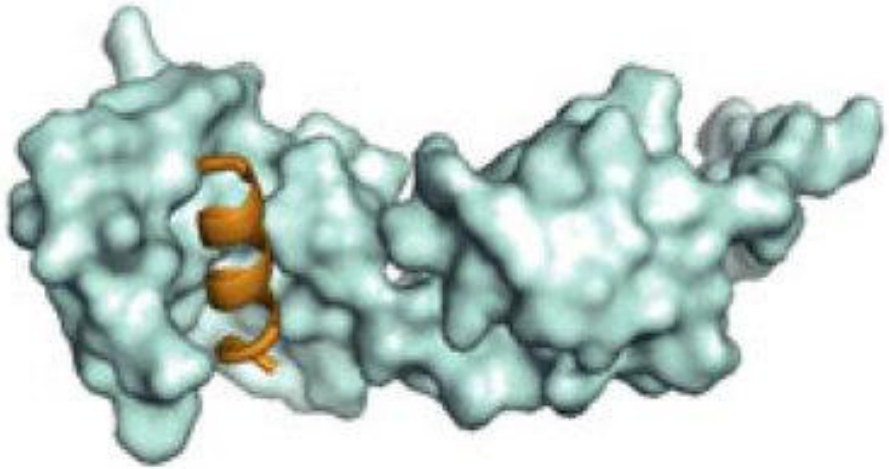
Table 4. Top 11 drug candidates in terms of affinity and activity.

Molecule (NSC)	Activity	Affinity	Active Bioassays
16437	.859	-9.3	Foot-and-mouth disease (FMD) virus
117987	.872	-8.8	
601359	.855	-8.4	Melanoma cell line, Malaria
13294	.825	-8.4	
127133	.908	-8.3	
61610	.823	-8.2	Malaria
107582	.877	-8.1	
128606	.920	-8.0	
211490	.808	-8.0	Hepatitis C virus, Human cytomegalovirus
679525	.894	-8.0	Orthopoxviruses, FMD virus
204232	.800	-7.9	DNA Polymerase Beta

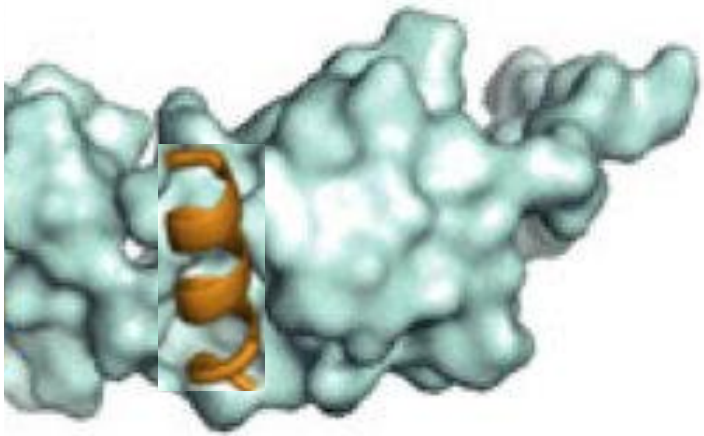
Biomimicry-based drug discovery

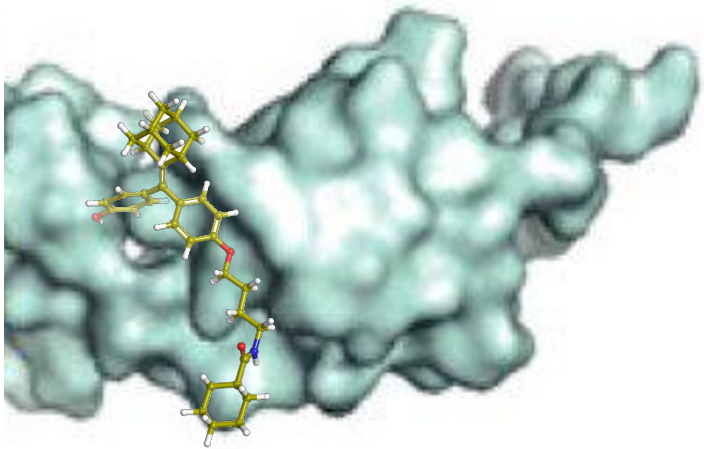
Molecular recognition features (MoRFs)



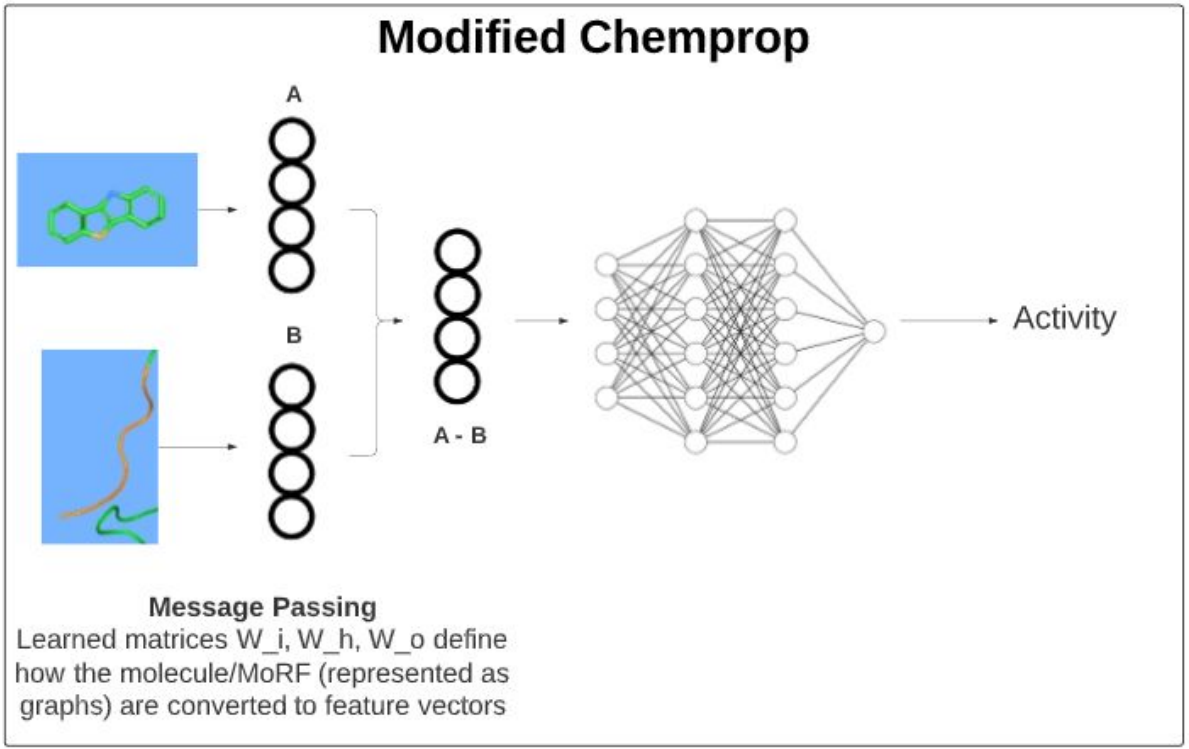


Biomimicry





The model


Modified Chemprop
A
B
A - B
Activity
Message Passing
Learned matrices W_i, W_h, W_o define how the molecule/MoRF (represented as graphs) are converted to feature vectors

Data

N = 7046

Each datapoint: ((ligand, MoRF), 0/1)

Bioassays 625185, 765185, 363

Total size = 7,045 | train size = 5,636 | val size = 704 | test size = 705

Results

Accuracy

Best validation AUC = 0.928 on epoch 29 (of 30).

Test AUC = 0.918

Previous work: .907

Pacific Symposium on Biocomputing

Mudide, Anish. et al. “SARS-CoV-2 Drug Discovery Based On Intrinsically Disordered Regions”, *Pac Symp Biocomput*, 2020.

Acknowledgements

Ning Xie and Ling Teng

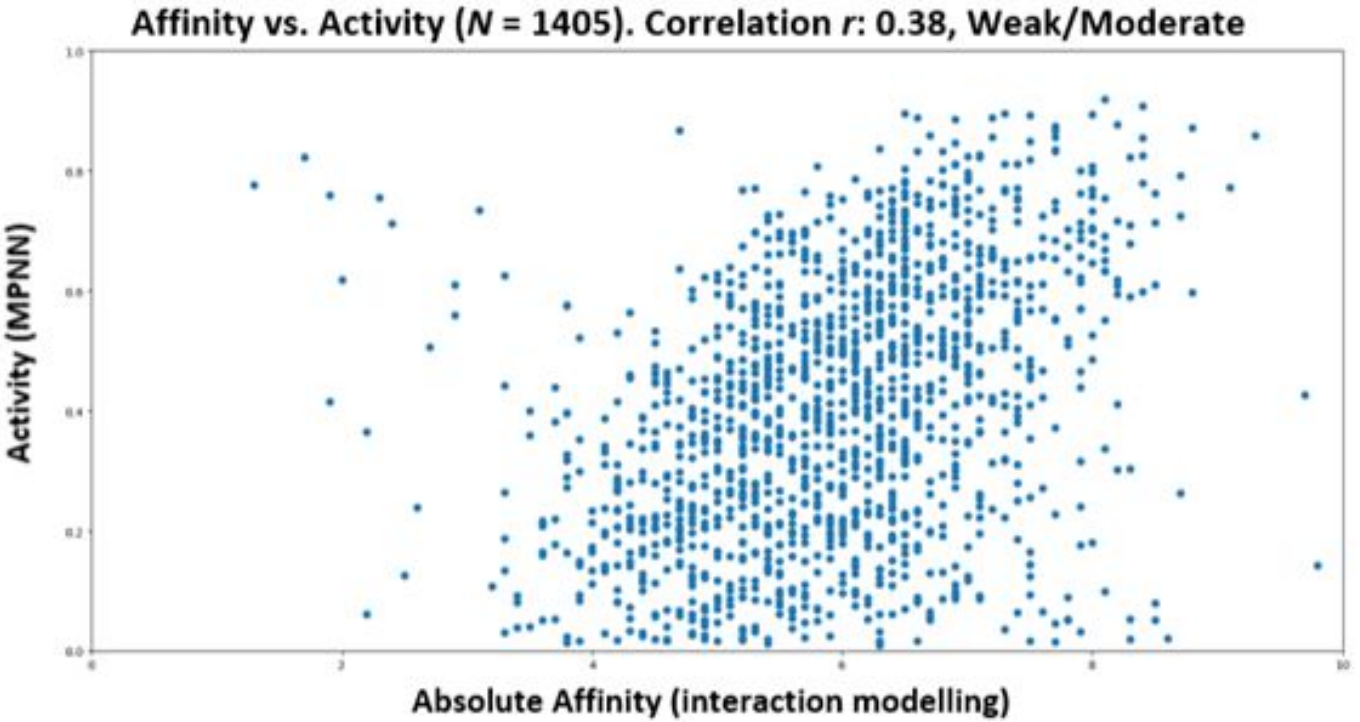
Prof. Gil Alterovitz

MIT PRIMES

My mom, dad and brother

Supplemental

Link between IDR affinity & 3CLpro inhibition score



Code

Molecular docking: https://github.com/Biomedical-Cybernetics-Lab2/IDR-SARS-CoV-2.

Modified chemprop for biomimicry: https://github.com/amudide/chemprop.

References

1. Wright, P. E., & Dyson, H. J. (2015). Intrinsically disordered proteins in cellular signalling and regulation. *Nature reviews. Molecular cell biology*, *16*(1), 18–29.
2. Wright, P. E., & Dyson, H. J. (1999). Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *Journal of molecular biology*, *293*(2), 321-331.
3. Rhoades, E. (2018). *Intrinsically Disordered Proteins*. Academic Press.
4. Mohan, A., Oldfield, C. J., Radivojac, P., Vacic, V., Cortese, M. S., Dunker, A. K., & Uversky, V. N. (2006). Analysis of molecular recognition features (MoRFs). *Journal of molecular biology*, *362*(5), 1043-1059.
5. Uversky, V. N., Oldfield, C. J., & Dunker, A. K. (2008). Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annu. Rev. Biophys.*, *37*, 215-246.
6. Chang, C. K., Hou, M. H., Chang, C. F., Hsiao, C. D., & Huang, T. H. (2014). The SARS coronavirus nucleocapsid protein--forms and functions. *Antiviral research*, *103*, 39–50.
7. Giri, R., Bhardwaj, T., Shegane, M., Gehi, B. R., Kumar, P., Gadhave, K., ... & Uversky, V. N. (2020). When Darkness Becomes a Ray of Light in the Dark Times: Understanding the COVID-19 via the Comparative Analysis of the Dark Proteomes of SARS-CoV-2, Human SARS and Bat SARS-Like Coronaviruses. *bioRxiv*.
8. Wang, R., Hozumi, Y., Yin, C., & Wei, G. W. (2020). Decoding SARS-CoV-2 Transmission and Evolution and Ramifications for COVID-19 Diagnosis, Vaccine, and Medicine. *Journal of chemical information and modeling*, acs.jcim.0c00501. Advance online publication.
9. Joshi, S., Joshi, M., & Degani, M. S. (2020). Tackling SARS-CoV-2: proposed targets and repurposed drugs. *Future medicinal chemistry*, 10.4155/fmc-2020-0147. Advance online publication.
10. Chen, Y. W., Yiu, C. B., & Wong, K. Y. (2020). Prediction of the SARS-CoV-2 (2019-nCoV) 3C-like protease ($3CL^{pro}$) structure: virtual screening reveals velpatasvir, ledipasvir, and other drug repurposing candidates. *F1000Research*, *9*, 129.

References (cntd.)

10. Antunes, D. A., Devaurs, D., & Kavraki, L. E. (2015). Understanding the challenges of protein flexibility in drug design. *Expert opinion on drug discovery*, *10*(12), 1301-1313.
11. Thafar, M., Raies, A. B., Albaradei, S., Essack, M., & Bajic, V. B. (2019). Comparison Study of Computational Prediction Tools for Drug-Target Binding Affinities. *Frontiers in Chemistry*, *7*.
12. Kairys, V., Baranauskiene, L., Kazlauskiene, M., Matulis, D., & Kazlauskas, E. (2019). Binding affinity in drug design: experimental and computational techniques. *Expert opinion on drug discovery*, *14*(8), 755–768.
13. National Center for Biotechnology Information (2020). PubChem Bioassay Record for AID 1706, Source: The Scripps Research Institute Molecular Screening Center.
14. Suárez, D., & Díaz, N. (2020). SARS-CoV-2 Main Protease: A Molecular Dynamics Study. Journal of chemical information and modeling.
15. Yang, K., Swanson, K., Jin, W., Coley, C., Eiden, P., Gao, H., ... & Palmer, A. (2019). Analyzing learned molecular representations for property prediction. *Journal of chemical information and modeling*, *59*(8), 3370-3388.
16. Khalili, N., Karimi, A., Moradi, M. T., & Shirzad, H. (2018). In vitro immunomodulatory activity of celastrol against influenza A virus infection. *Immunopharmacology and Immunotoxicology*, *40*(3), 250-255.
17. Yu, J. S., Tseng, C. K., Lin, C. K., Hsu, Y. C., Wu, Y. H., Hsieh, C. L., & Lee, J. C. (2017). Celastrol inhibits dengue virus replication via up-regulating type I interferon and downstream interferon-stimulated responses. *Antiviral research*, *137*, 49-57.
18. Habtemariam, S., Nabavi, S. F., Berindan‐Neagoe, I., Cismaru, C. A., Izadi, M., Sureda, A., & Nabavi, S. M. (2020). Should we try the antiinflammatory natural product, celastrol, for COVID‐19?. *Phytotherapy Research*.