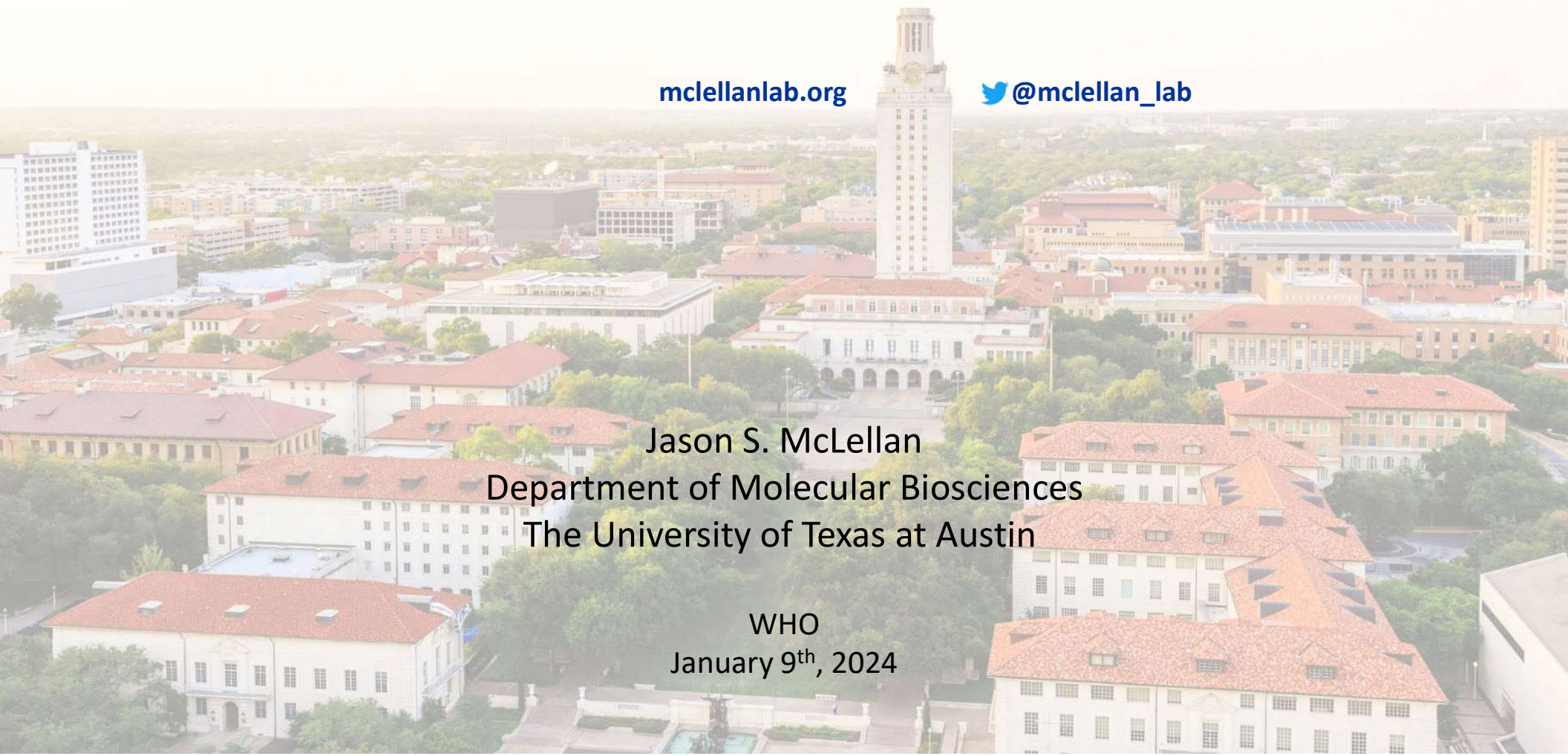
Using New Technologies to Define the Atomic-level Details of Surface Proteins Likely to be Vaccine Targets

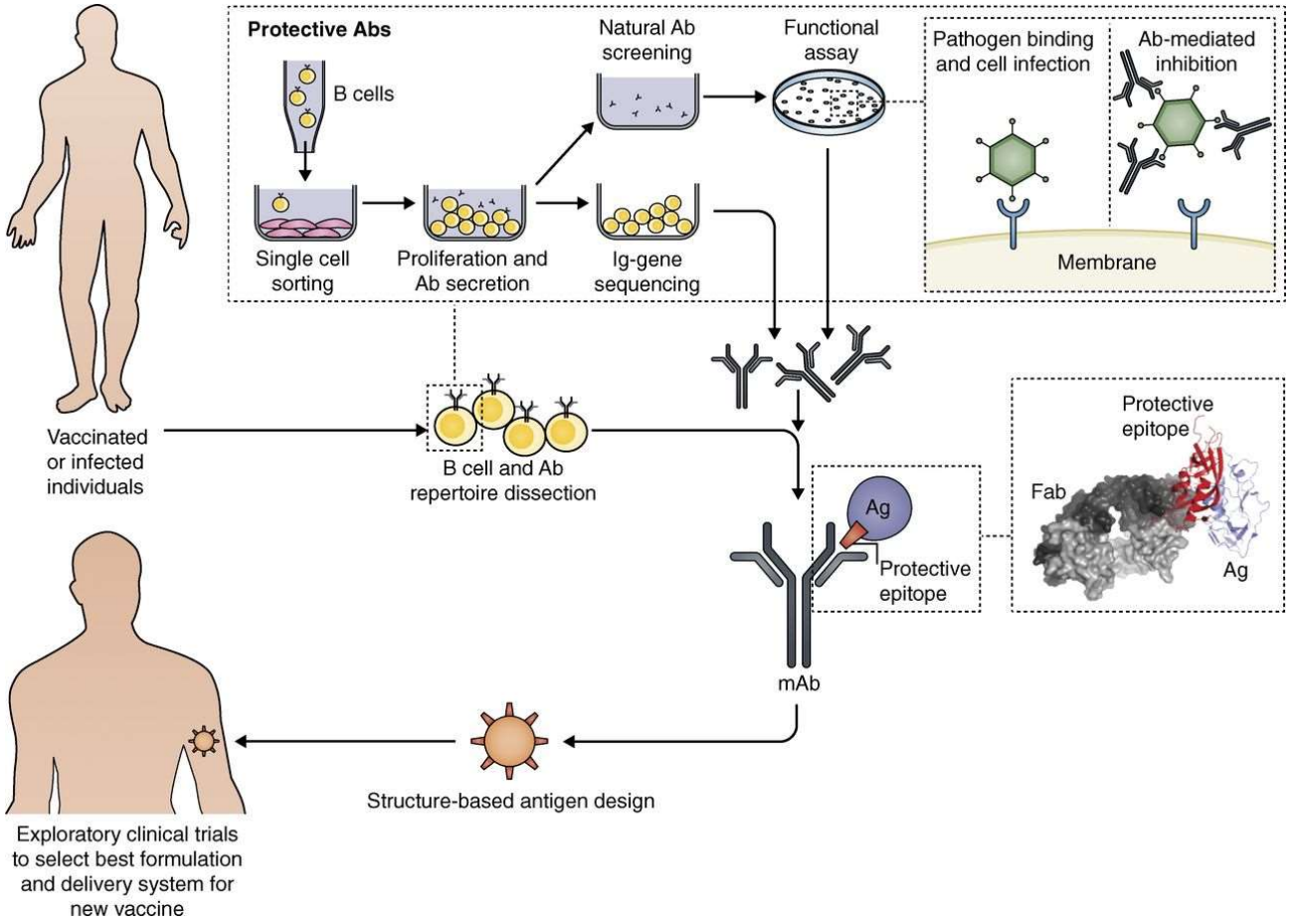


mclellanlab.org @mclellan_lab

Jason S. McLellan
Department of Molecular Biosciences
The University of Texas at Austin

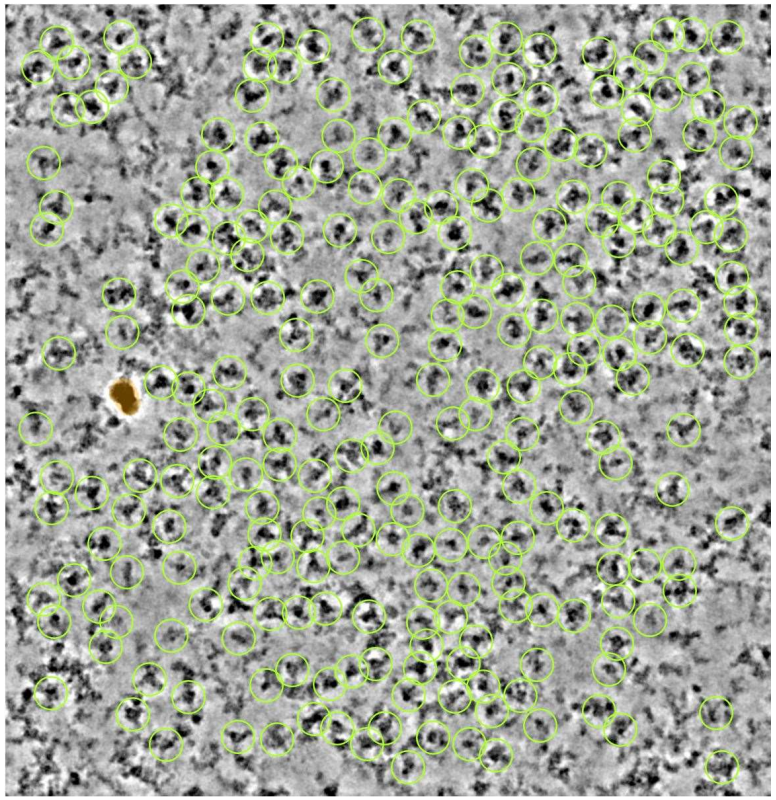
WHO
January 9th, 2024

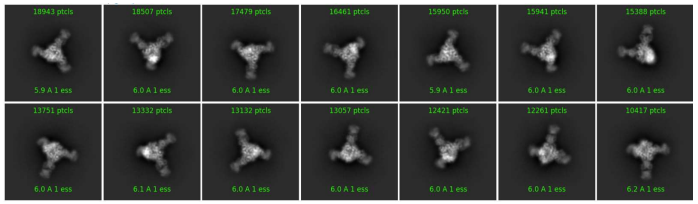
Structure-based Vaccine Antigen Design

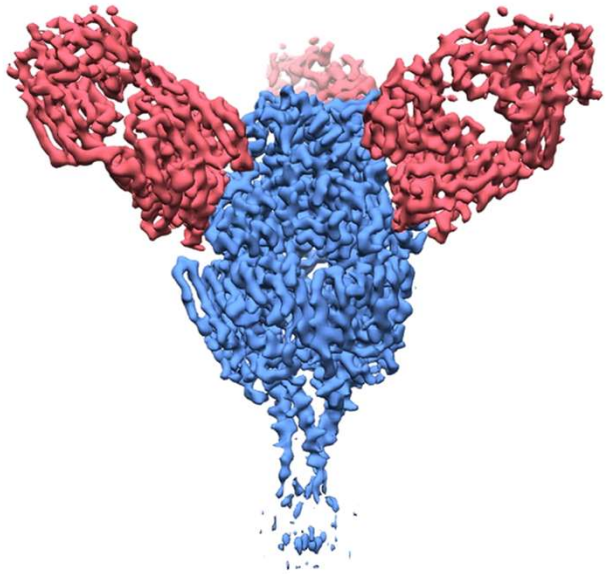


Rappuoli et al, *J Exp Med* 2016

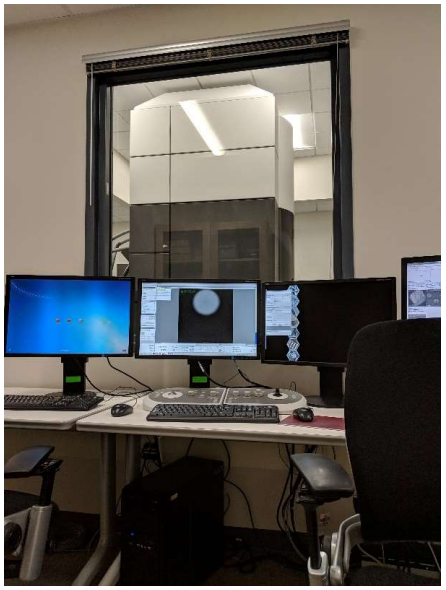
Structural Biology: Cryo-EM





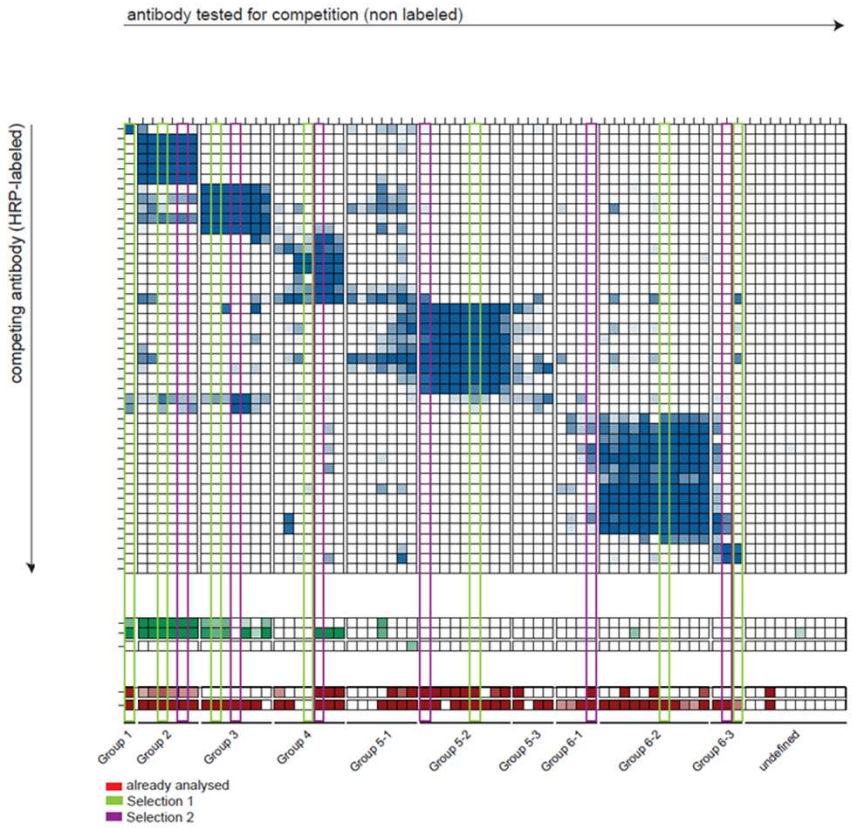


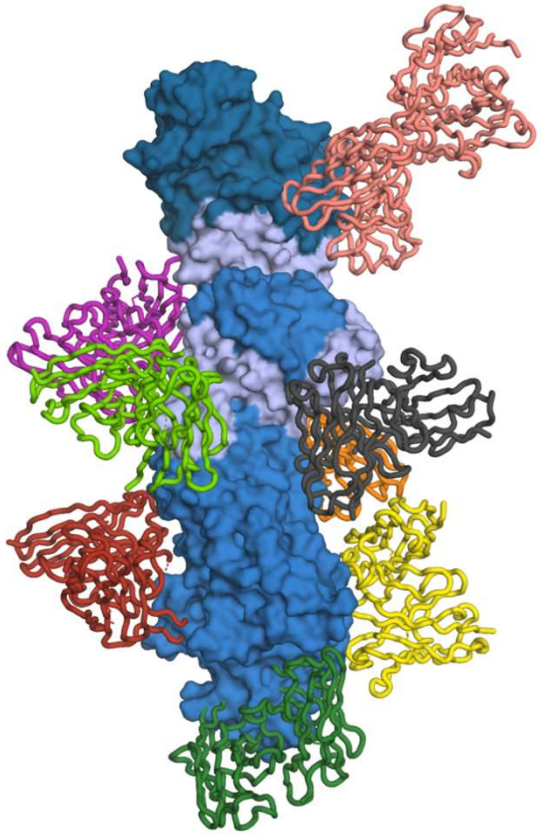
Structural Biology: Cryo-EM



Multiplexing to Increase Antigen-Antibody Structure Throughput

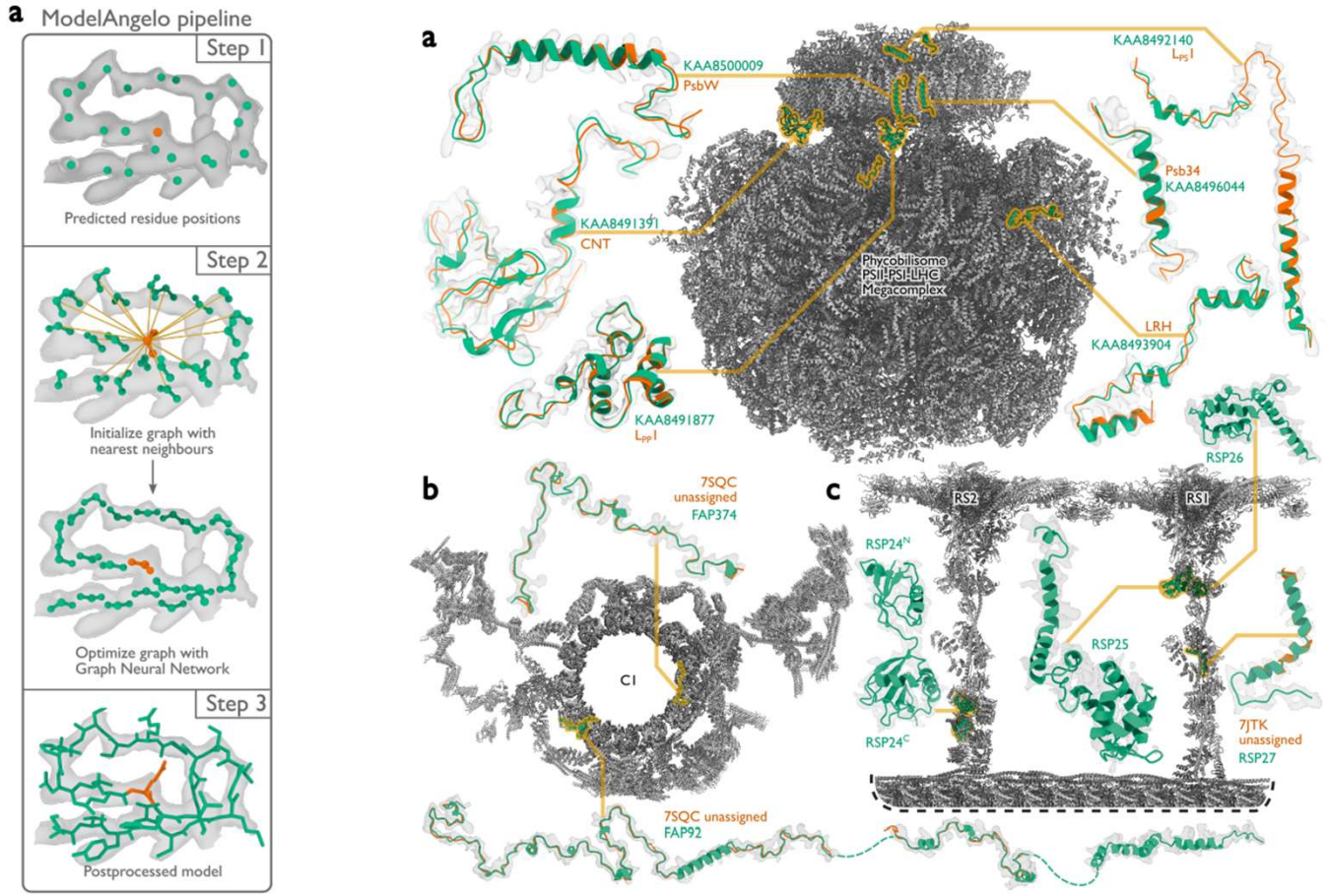
- High-throughput antibody isolation combined with precise binning of antibodies into competition groups allows multiple antibody interfaces to be structural determined simultaneously





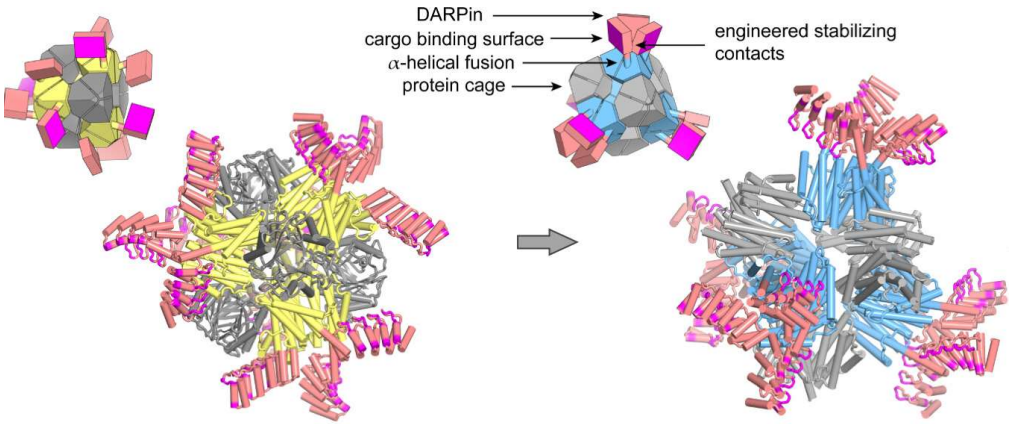
McLellan lab, unpublished

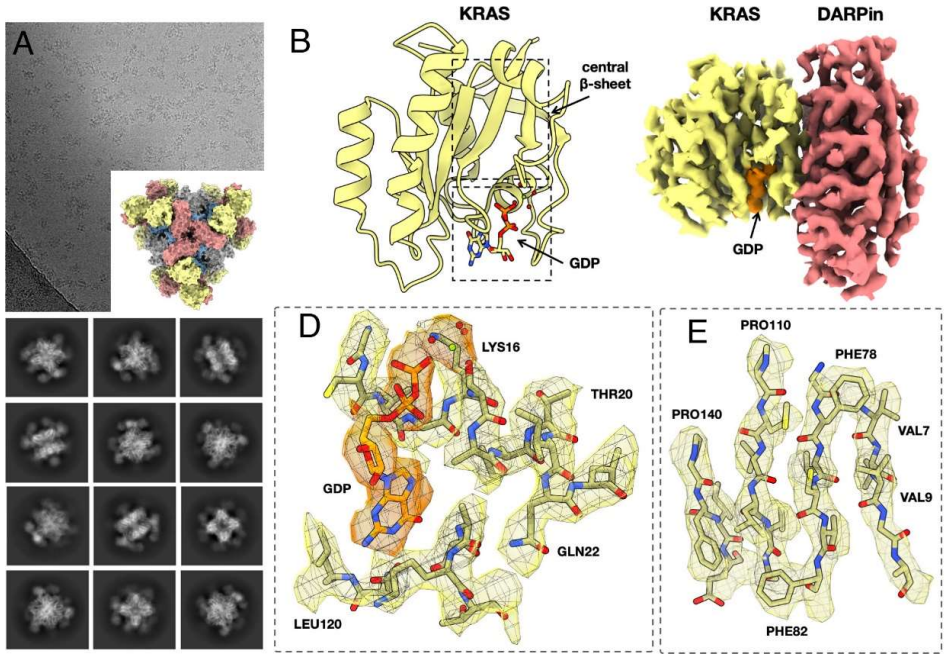
New Advances in Automated Model Building Accelerate Structure Determination Efforts



Jamali, …, Scheres. https://www.biorxiv.org/content/10.1101/2023.05.16.541002v2.full

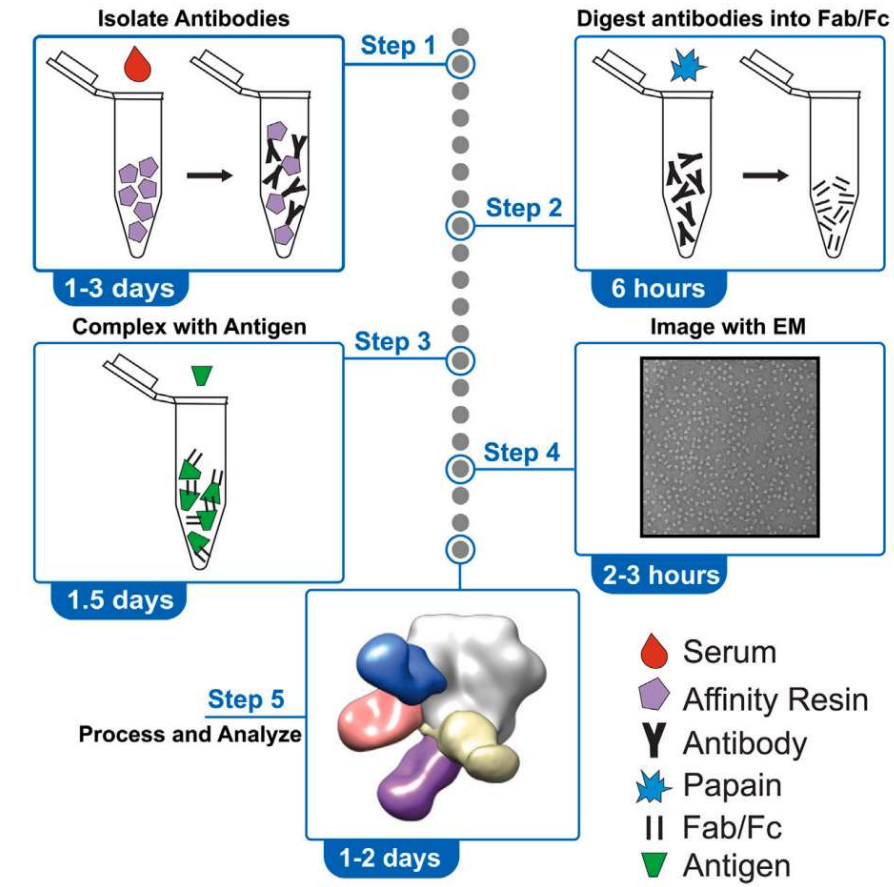
Cryo-EM Imaging Scaffolds to Enable Determination of Small Proteins





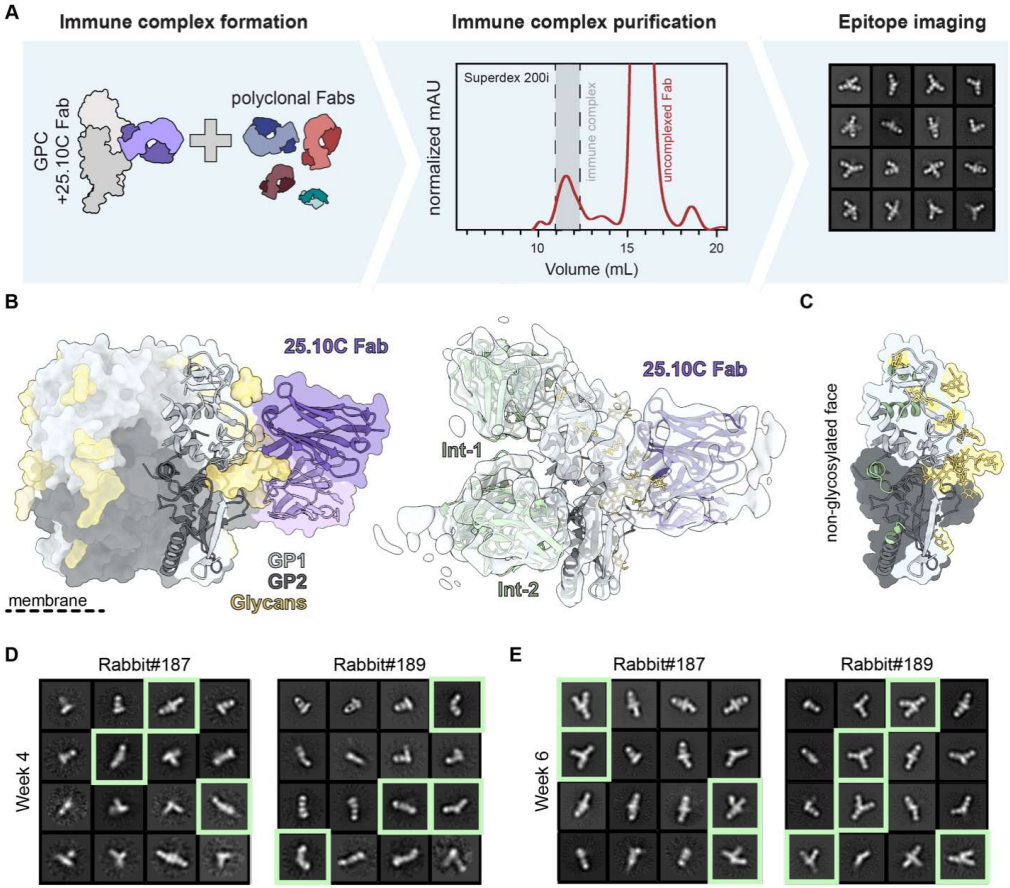
Castells-Graells, …, Yeates, *PNAS*, 2023

Electron-microscopy-based Polyclonal Epitope Mapping (EMPEM)



Turner, …, Ward, *STAR Protoc*, 2023

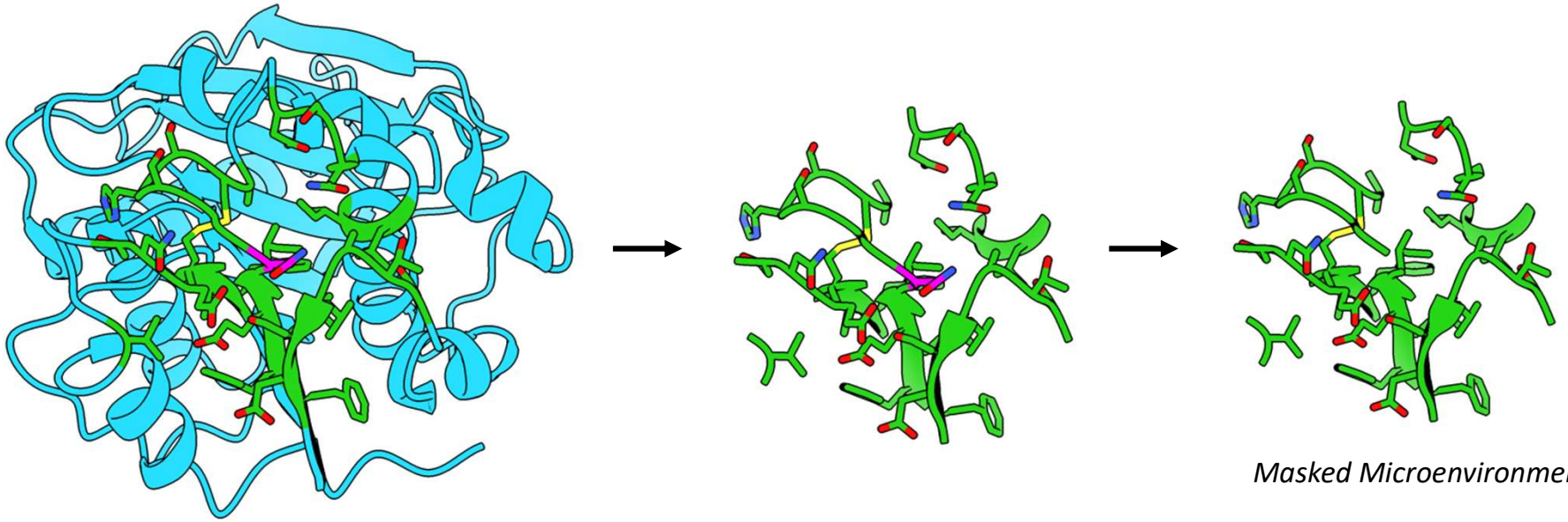
EMPEM Experiments with LASV GP Monomers Reveal Responses Targeting the Trimer Interior



Brouwer, …, Ward. https://www.biorxiv.org/content/10.1101/2023.12.21.572918v1.full.pdf

MutComputeX: Using Self-supervision to Learn What Proteins "Should" Look Like

- Vison-based approach: uses protein structures for input and training



Masked Microenvironment

Center **microenvironment** around an **amino acid**

Delete **remaining** protein atoms

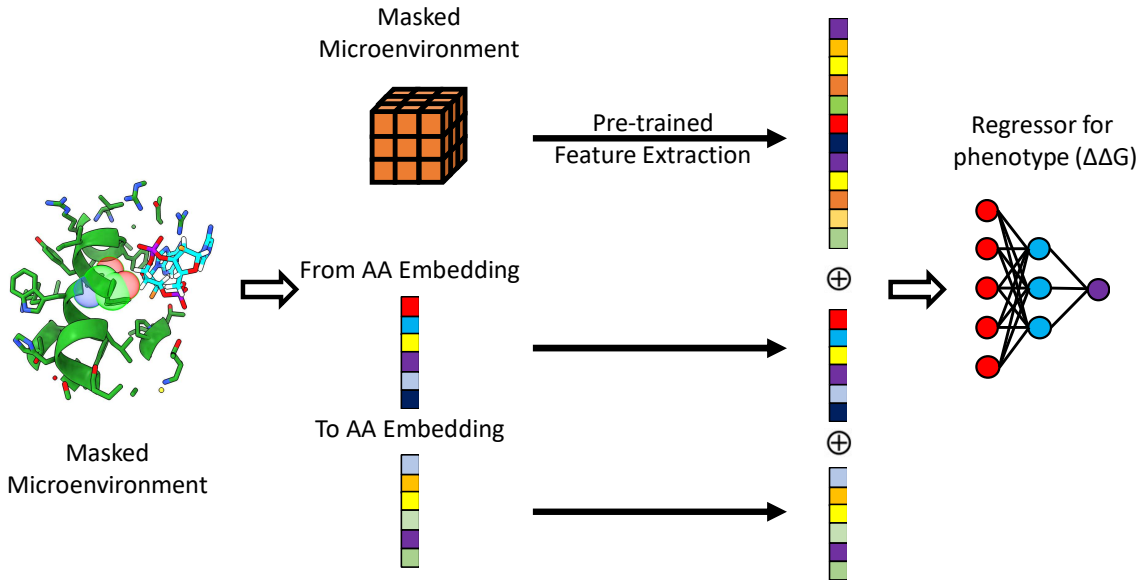
Delete centered **amino acid** and use as **label**
Classification into amino acid **likelihoods**

Adam Klivans
Danny Diaz

Institute for Foundations of MACHINE LEARNING

Torng and Altman, BMC Bioinformatics **2017**

Stability Oracle: Fine-Tuning MutComputeX on Thermodynamic Stability (ΔΔG)



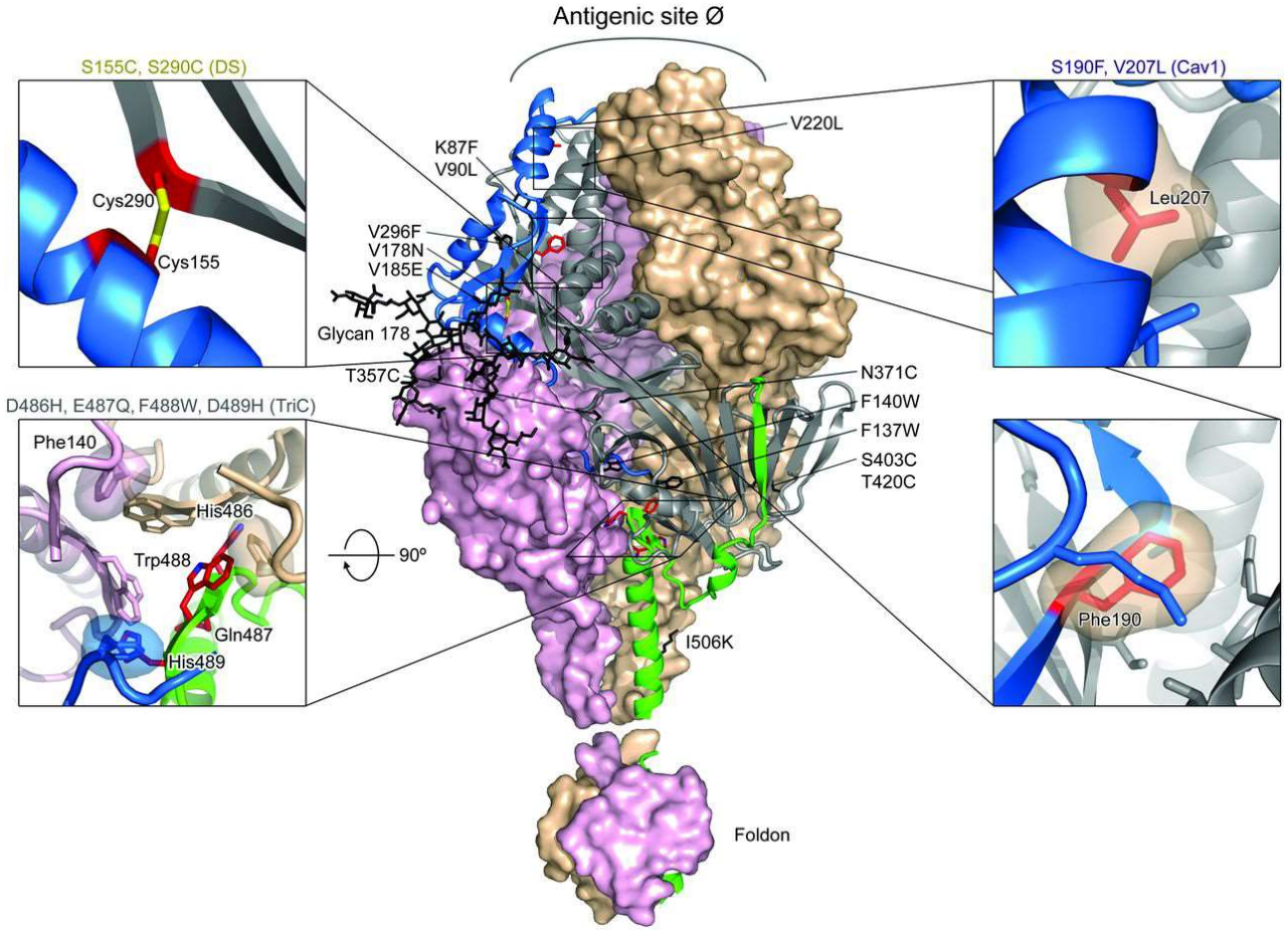
Input: Masked Microenvironment of Chemistry and vector representations of amino acids (embeddings)
Feature Extractor and ***Amino Acid embeddings*** are obtained from a pre-trained classification model (MutComputeX)

Adam Klivans
Danny Diaz

IFML Institute for Foundations of MACHINE LEARNING

Diaz et al. Biorxiv **2023**

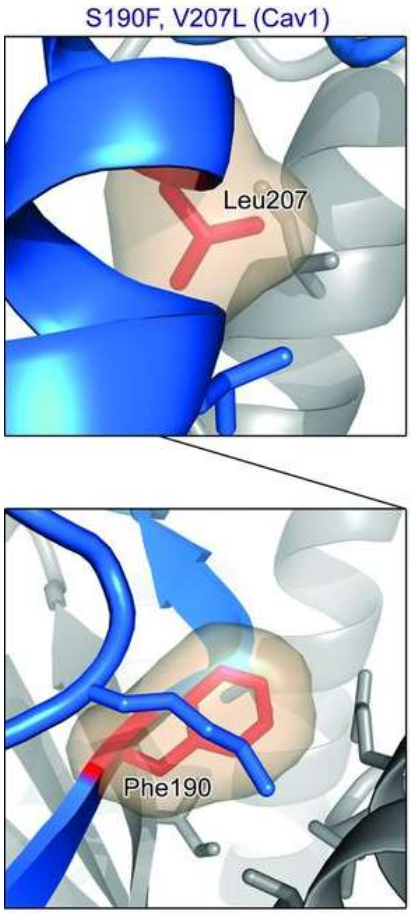
Stabilizing RSV Fusion (F) Glycoprotein Substitutions



Ajit Ramamohan

McLellan et al., *Science*. 2013

ML Model Predictions of Stabilizing RSV F Cavity-Filling Substitutions

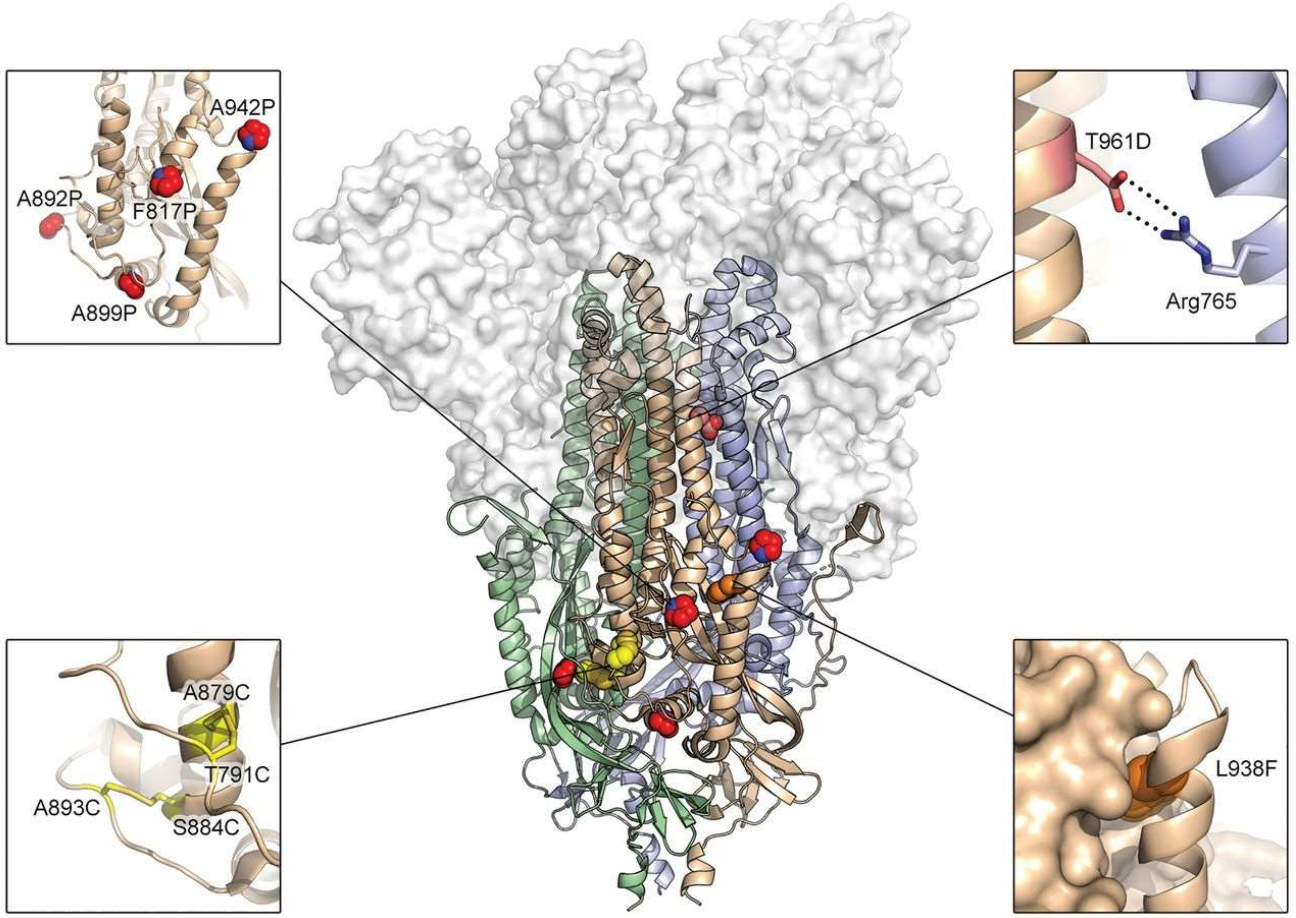


	S190F	V207L
Stability Oracle (ΔΔG)	-1.0	-0.1
Stability Fold (ΔΔG)	0.4	0.4
Prostata-IFML (ΔΔG)	0.2	0.0
MutComputeX (Log Probability)	1.4	-0.4
ProteinMPNN (Log Probability)	0.7	0.0

ΔΔG < 0 and Log Probability > 0 are predicted to be beneficial

Ajit Ramamohan

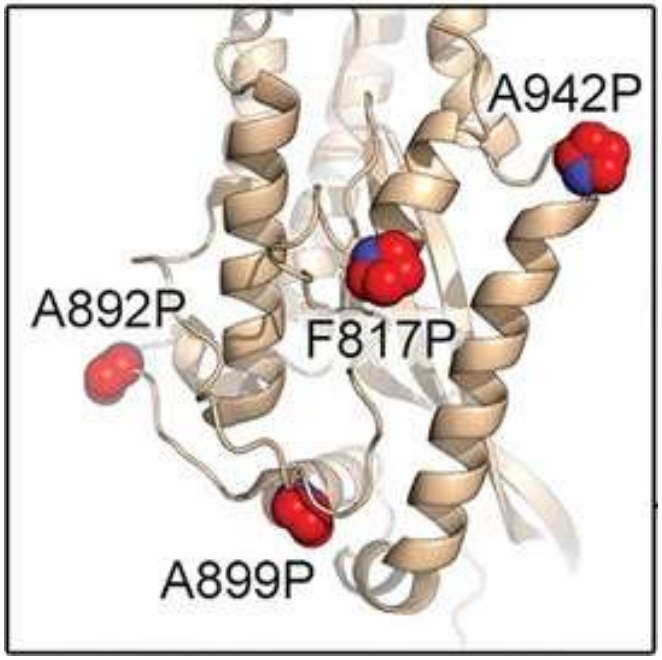
Exemplary Substitutions for SARS-CoV-2 Spike (S) Stabilization



Ajit Ramamohan

Hseih et al., *Science.* 2020

ML Model Predictions of Stabilizing SARS-CoV-2 S Proline Substitutions



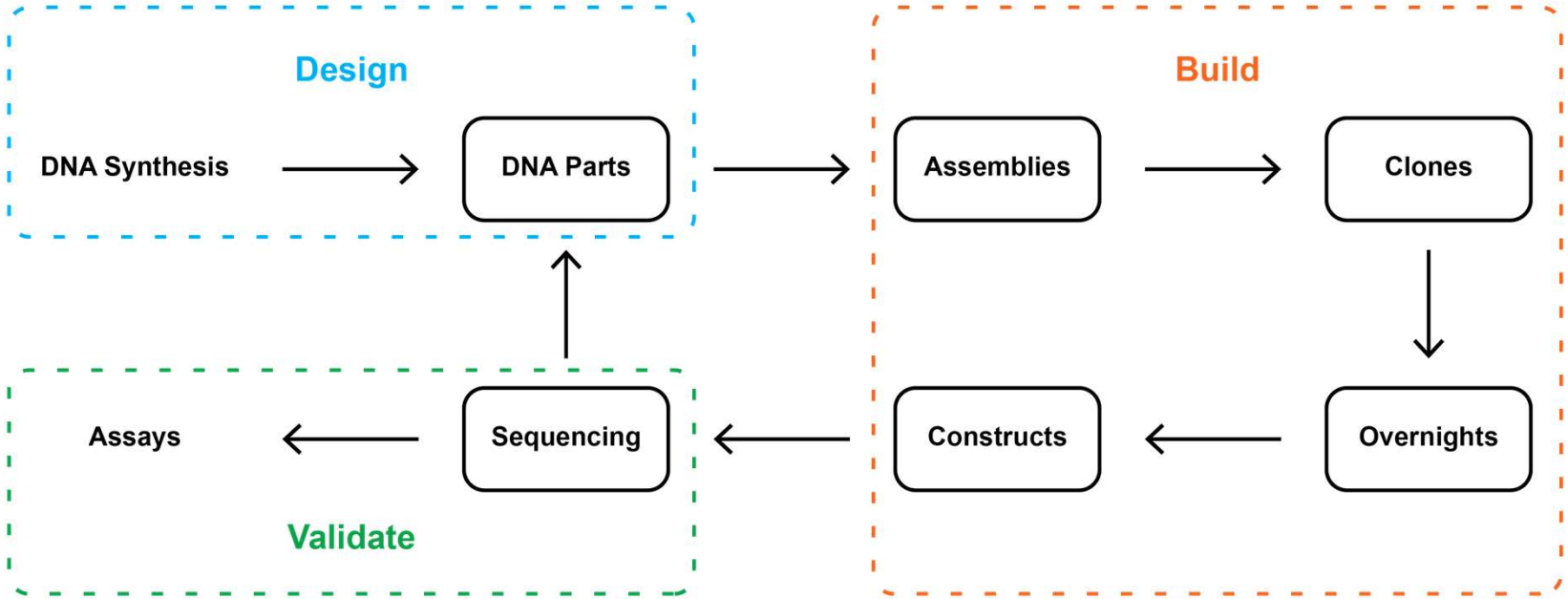
	F817P	A892P	A899P	A942P
Stability Oracle (ΔΔG)	-0.1	-0.9	1.6	0.1
Stability Fold (ΔΔG)	1.0	0.0	0.8	1.3
Prostata-IFML (ΔΔG)	1.1	0.8	1.0	1.7
MutComputeX (Log Probability)	0.9	2.3	0.4	-0.1
ProteinMPNN (Log Probability)	1.6	-0.1	-0.6	1.0

ΔΔG < 0 and Log Probability > 0 are predicted to be beneficial

Ajit Ramamohan

Synthetic Biology for Antigen Engineering

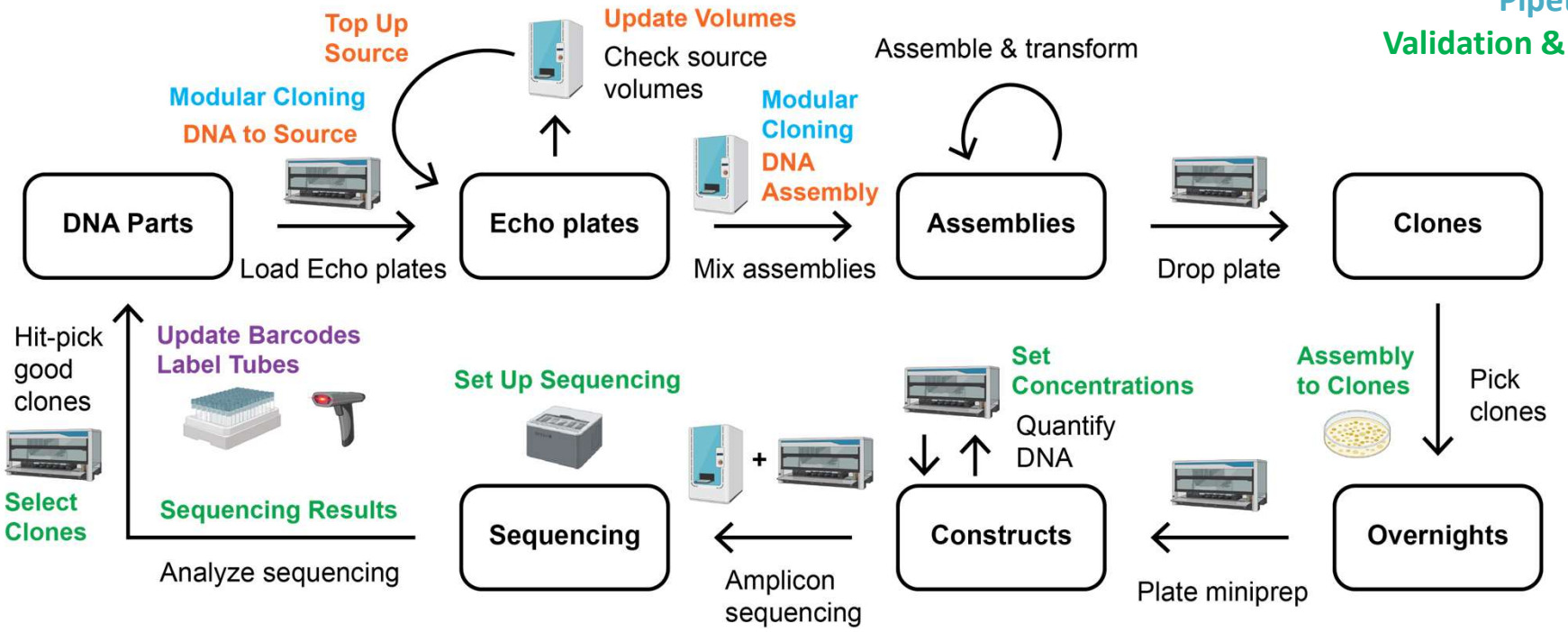
High-throughput synthetic biology accelerates antigen engineering by enabling rapid design-build-validate cycles for many new protein designs



Jimmy Gollihar
Houston Methodist

Scripts are Necessary for High-throughput Synthetic Biology

Sample handling & Barcode Management
Genetic Design
Pipetting & Echo
Validation & Sequencing



Jimmy Gollihar
Houston Methodist

High-throughput Variant Screening via Flow Cytometry

- Screened 10 proline substitutions and 8 disulfide substitutions for binding to 26 antibodies in a few hours
- Heat-shocked the cells to induce conformational change and assess retention of antibody binding

		prolines										L4	L6	disulfides							
Name	mAb																				
10.4B	1	973.03	289.88	256.53	168.71	345.04	338.79	440.5	297.47	254.97	1061.5	370.37	77.94	85.48	546.06	272.53	934.2	980.95	1343.91	1323.75	649.89
12.1F	2	3587.36	987.53	781.57	496.72	1000.08	1027.92	1674.35	1109.8	917.79	2137.7	1460.93	87.52	97.03	2321.55	1544.67	1019.6	3487.78	3697.8	3521.91	2177.84
18.5C	3	313.17	187.09	149.91	142.29	191.11	233.55	211.72	181.43	104.17	359.13	204.85	87.03	86.17	180.18	225.5	179.55	799.69	451.84	322.4	269.25
18.5C-M30	4	1461.93	557.68	496.72	614.75	627.91	1309.01	812.56	921.26	136.48	1208.26	719.11	104.17	98.61	789.39	674.2	578.37	1886.22	2009.97	1537.99	632.98
19.7E	5	2627.64	1083.98	761.62	637.94	1122.38	1192.04	1603.12	1261.92	837.59	1936.78	1467.34	84.53	86.78	2334.13	1847.71	1857.82	3226.6	3393.49	3295.73	1918.38
25.10C	7	3925.26	1400.1	1023.76	784.06	1506.93	1669.05	2469.95	1704.2	1276.62	2835.21	2251.39	83.44	81.53	3283.35	2663.4	3388.94	4311.92	6242.23	5965.83	3693.02
25.6A	8	402.88	209.1	174.15	155.94	207.61	262.67	234.36	210.1	126.15	261.7	197.87	89.44	83.81	263.91	221.77	192.58	681.69	536.27	454.66	313.17
36.1F	9	1186.07	503.84	396.06	325.46	571.89	454.66	646.72	403.44	380.54	809.32	558.25	106.5	96.22	697.34	585.67	740.23	2288.48	2878.22	3359.41	1695.01
36.9F	10	106.31	98.97	97.45	75.46	112.12	153.38	109.28	95.65	92.99	97.63	99.54	87.16	90.87	104.54	103.34	86.61	180.39	132.17	125.05	126.57
37.2D	11	995.01	434.98	352.58	316.38	442.86	566.76	644.19	501.26	304.88	1071.07	482.79	93.95	85.39	666.76	594.09	551.84	1819.17	2036.04	1293.24	953.11
37.2G	12	196.34	144.48	105.48	114.59	163.89	187.44	140.49	139.42	94.92	200.72	138.47	78.11	90.58	137.77	148.85	135.21	396.79	298.16	205.72	209.56
37.7H	13	818.81	382.95	273.25	282.63	445.03	535.81	539.67	533.2	85.29	703.67	444.02	89.45	88.02	688.05	514.01	581.39	1365.59	2062.7	1315.51	883.08
8.11G	14	943.88	365.61	292.45	255.53	387.69	370.26	482.79	360.47	293.8	567.09	375.49	85.75	80.77	507.76	442.61	526.15	1227.22	1206.12	1607.14	799.69
8.9F	15	1158.68	243.47	164.48	123.46	296.07	291.46	363.83	357.17	184.01	349.32	310.25	82	92.71	845.03	451.14	1178.3	1291.24	1576.14	1974.06	522.9
9.8A	16	703.67	271.82	245.38	204.6	260.26	329.28	364.13	327.31	146.71	601.77	320.66	77.42	92.46	418.97	348.37	258.56	911.75	1342.84	580.37	602
NE13	17	200.88	121.35	115.19	106.14	137.56	152.69	146.71	129.33	96.19	160.76	147.21	82	83.03	136.77	155.94	140.37	239.52	286.6	223.5	174.82
18.5D	18	235.74	215.02	212.09	147.08	193.24	221.52	264.35	205.6	236.52	278.6	237.84	200.36	248.14	230.63	275.6	283.57	331.01	329.26	329.89	329.93
LAVA01	19	207.84	98.16	104.17	74.32	111.85	106.99	122.51	111.76	101.96	125.73	109.75	76.9	88.02	117.74	111.5	83.36	202.57	351.07	365.9	263.5
10.4B-Li	20	3981.07	5524	4457.89	2658.26	2770.14	3354.57	4707.58	3493.44	4939.06	5401.99	5489.44	5844.94	7877.22	4763.36	7376.35	7094.59	5159.68	6515.21	6185.11	8330.17
12.1F-pdb	21	2322.9	687.82	555.43	396.35	775.81	765.31	1268.18	894.72	672.11	1509.28	1041.74	81.11	83.2	1839.46	1196.01	928.48	3094.72	3957.64	3566.68	2293.34
18.5C-pdb	22	314.69	200.11	159.54	158.73	223.61	252.54	213.36	204.67	95.51	395.85	186.57	83.68	88.56	204.25	216.61	185.98	502.82	502.41	360.56	316.81
25.10C-FNQI	24	4508.66	1579.96	1283.22	766.94	2013.38	1781.65	2766.58	1873.28	1301.13	2947.19	2059.48	73.64	77.3	3240.54	2590.72	3382.21	5262.07	6293.92	6372.83	3828.47
36.1F-pdb	25	1557.43	684.92	527.92	393.15	787.03	725.88	1018.56	661.04	504.64	1098.5	781.57	121.98	134.89	990.53	951.09	918.18	2648	3510.43	3450.23	1986.66
8.9F-pdb	26	505.27	176.75	124.53	109.28	217.02	212.81	263.71	263.91	121.85	215.81	186.3	74.09	92.91	477.35	250.27	802.09	700.55	893.82	958.34	338.48

												L4	L6								
Name	mAb																				
10.4B	1	2.63	0.78	0.69	0.46	0.93	0.91	1.19	0.80	0.69	2.87	1.00	0.21	0.23	1.47	0.74	2.52	2.65	3.63	3.57	1.75
12.1F	2	2.46	0.68	0.53	0.34	0.68	0.70	1.15	0.76	0.63	1.46	1.00	0.06	0.07	1.59	1.06	0.70	2.39	2.53	2.41	1.49
18.5C	3	1.53	0.91	0.73	0.69	0.93	1.14	1.03	0.89	0.51	1.75	1.00	0.42	0.42	0.88	1.10	0.88	3.90	2.21	1.57	1.31
18.5C-M30	4	2.03	0.78	0.69	0.85	0.87	1.82	1.13	1.28	0.19	1.68	1.00	0.14	0.14	1.10	0.94	0.80	2.62	2.80	2.14	0.88
19.7E	5	1.79	0.74	0.52	0.43	0.76	0.81	1.09	0.86	0.57	1.32	1.00	0.06	0.06	1.59	1.26	1.27	2.20	2.31	2.25	1.31
25.10C	7	1.74	0.62	0.45	0.35	0.67	0.74	1.10	0.76	0.57	1.26	1.00	0.04	0.04	1.46	1.18	1.51	1.92	2.77	2.65	1.64
25.6A	8	2.04	1.06	0.88	0.79	1.05	1.33	1.18	1.06	0.64	1.32	1.00	0.45	0.42	1.33	1.12	0.97	3.45	2.71	2.30	1.58
36.1F	9	2.12	0.90	0.71	0.58	1.02	0.81	1.16	0.72	0.68	1.45	1.00	0.19	0.17	1.25	1.05	1.33	4.10	5.16	6.02	3.04
36.9F	10	1.07	0.99	0.98	0.76	1.13	1.54	1.10	0.96	0.93	0.98	1.00	0.88	0.91	1.05	1.04	0.87	1.81	1.33	1.26	1.27
37.2D	11	2.06	0.90	0.73	0.66	0.92	1.17	1.33	1.04	0.63	2.22	1.00	0.19	0.18	1.38	1.23	1.14	3.77	4.22	2.68	1.97
37.2G	12	1.42	1.04	0.76	0.83	1.18	1.35	1.01	1.01	0.69	1.45	1.00	0.56	0.65	0.99	1.07	0.98	2.87	2.15	1.49	1.51
37.7H	13	1.84	0.86	0.62	0.64	1.00	1.21	1.22	1.20	0.19	1.58	1.00	0.20	0.20	1.55	1.16	1.31	3.08	4.65	2.96	1.99
8.11G	14	2.51	0.97	0.78	0.68	1.03	0.99	1.29	0.96	0.78	1.51	1.00	0.23	0.22	1.35	1.18	1.40	3.27	3.21	4.28	2.13
8.9F	15	3.73	0.78	0.53	0.40	0.95	0.94	1.17	1.15	0.59	1.13	1.00	0.26	0.30	2.72	1.45	3.80	4.16	5.08	6.36	1.69
9.8A	16	2.19	0.85	0.77	0.64	0.81	1.03	1.14	1.02	0.46	1.88	1.00	0.24	0.29	1.31	1.09	0.81	2.84	4.19	1.81	1.88
NE13	17	1.36	0.82	0.78	0.72	0.93	1.04	1.00	0.88	0.65	1.09	1.00	0.56	0.56	0.93	1.06	0.95	1.63	1.95	1.52	1.19
18.5D	18	0.99	0.90	0.89	0.62	0.81	0.93	1.11	0.86	0.99	1.17	1.00	0.84	1.04	0.97	1.16	1.19	1.39	1.38	1.39	1.39
LAVA01	19	1.89	0.89	0.95	0.68	1.02	0.97	1.12	1.02	0.93	1.15	1.00	0.70	0.80	1.07	1.02	0.76	1.85	3.20	3.33	2.40
10.4B-Li	20	0.73	1.01	0.81	0.48	0.50	0.61	0.86	0.64	0.90	0.98	1.00	1.06	1.43	0.87	1.34	1.29	0.94	1.19	1.13	1.52
12.1F-pdb	21	2.23	0.66	0.53	0.38	0.74	0.73	1.22	0.86	0.65	1.45	1.00	0.08	0.08	1.77	1.15	0.89	2.97	3.80	3.42	2.20
18.5C-pdb	22	1.69	1.07	0.86	0.85	1.20	1.35	1.14	1.10	0.51	2.12	1.00	0.45	0.47	1.09	1.16	1.00	2.70	2.69	1.93	1.70
25.10C-FNQI	24	2.19	0.77	0.62	0.37	0.98	0.87	1.34	0.91	0.63	1.43	1.00	0.04	0.04	1.57	1.26	1.64	2.56	3.06	3.09	1.86
36.1F-pdb	25	1.99	0.88	0.68	0.50	1.01	0.93	1.30	0.85	0.65	1.41	1.00	0.16	0.17	1.27	1.22	1.17	3.39	4.49	4.41	2.54
8.9F-pdb	26	2.71	0.95	0.67	0.59	1.16	1.14	1.42	1.42	0.65	1.16	1.00	0.40	0.50	2.56	1.34	4.31	3.76	4.80	5.14	1.82

Jimmy Gollihar
Houston Methodist

Summary

Summary

- New advances in cryo-EM have enabled higher resolution and higher throughput than ever before

- Techniques such as EMPEM provide valuable insights into polyclonal immune responses to vaccination and infection

- AI/ML combined with high-throughput screening is allowing accelerated development of vaccine antigens for important human pathogens

Acknowledgements

McLellan Laboratory
Jory Goldsmith
Ching-Lin Hsieh
Ling Zhou
Nianshuang Wang
Daniel Wrapp
Patrick Byrne
Christy Hjorth
Nicole Johnson
Ryan McCool
Cory Acreman

Institute for Machine Learning
Adam Klivans
Danny Diaz

Houston Methodist
Jimmy Gollihar