

Supporting Information

Allosteric Control of

ACE2 Peptidase Domain Dynamics

Francesco Trozzi, Nischal Karki, Zilin Song, Niraj Verma,
Elfi Kraka, Brian Zoltowski, and Peng Tao*

Department of Chemistry, Center for Research Computing, Center for Drug Discovery, Design,
and Delivery (CD4), Southern Methodist University, Dallas, US

*Correspondence and Lead-contact: Peng Tao (ptao@smu.edu)

Table S1. Naming convention for ACE2 PD.

Residue	Feature key	Name	Residue	Feature key	Name
21 - 52	Helix	$\alpha 1$	331 - 339	Turn	$\alpha 20$ - $\beta 21$ loop
56 - 77	Helix	$\alpha 2$	347 - 352	Beta strand	$\beta 21$
78 - 90	Turn	$\alpha 2$ - $\alpha 3$ loop	355 - 359	Beta strand	$\beta 22$
91 - 102	Helix	$\alpha 3$	366 - 384	Helix	$\alpha 23$
103-109	Turn	$\alpha 3$ - $\alpha 4$ loop	385 - 387	Turn	$\alpha 23$ - $\alpha 24$ loop
110 - 129	Helix	$\alpha 4$	390 - 392	Helix	$\alpha 24$
131 - 134	Beta strand	$\beta 5$	396 - 399	Beta strand	$\beta 25$
137 - 143	Beta strand	$\beta 6$	400 - 413	Helix	$\alpha 26$
144 - 146	Turn	$\beta 6$ - $\alpha 7$ loop	415 - 420	Helix	$\alpha 27$
148 - 154	Helix	$\alpha 7$	422 - 426	Turn	$\alpha 27$ - $\alpha 28$ loop
158 - 171	Helix	$\alpha 8$	432 - 446	Helix	$\alpha 28$
173 - 193	Helix	$\alpha 9$	449 - 465	Helix	$\alpha 29$
196- 198	Beta strand	$\beta 10$	466 - 468	Beta strand	$\beta 30$
199 - 204	Helix	$\alpha 11$	470 - 472	Helix	$\alpha 31$
205 - 207	Turn	$\alpha 11$ - $\alpha 12$ loop	473 - 483	Helix	$\alpha 32$
213 - 215	Turn	$\alpha 11$ - $\alpha 12$ loop	486 - 488	Beta strand	$\beta 33$
220 - 251	Helix	$\alpha 12$	499 - 502	Helix	$\alpha 34$
253 - 255	Turn	$\alpha 11$ - $\beta 13$ loop	504 - 507	Helix	$\alpha 35$
258 - 260	Beta strand	$\beta 13$	514 - 531	Helix	$\alpha 36$
264 - 266	Helix	$\alpha 14$	532 - 534	Turn	$\alpha 36$ - $\alpha 37$ loop
267 - 271	Beta strand	$\beta 15$	539 - 541	Helix	$\alpha 37$
276 - 278	Helix	$\alpha 16$	548 - 558	Helix	$\alpha 38$
279 - 282	Helix	$\alpha 17$	559 - 562	Turn	$\alpha 38$ - $\alpha 39$ loop
284 - 287	Turn	$\alpha 17$ - $\alpha 18$ loop	566 - 574	Helix	$\alpha 39$
294 - 297	Turn	$\alpha 17$ - $\alpha 18$ loop	575 - 578	Beta strand	$\beta 40$
298 - 300	Helix	$\alpha 18$	582 - 598	Helix	$\alpha 41$
304 - 316	Helix	$\alpha 19$	600 - 602	Beta strand	$\beta 42$
317 - 319	Turn	$\alpha 219$ - $\alpha 20$ loop	607 - 611	Beta strand	$\beta 43$
327 - 330	Helix	$\alpha 20$			

Table S2. CV-CNN layers details.

Layer Property	Conv2D	Conv2D	Dense	Classification	Regression
Kernel initializer	He uniform	He uniform	He uniform	Glorot uniform	Glorot uniform
Kernel regularizer	L1, 0.00001	L1, 0.00001	L1, 0.00001	0	0
Kernel size	(4, 4)	(4, 4)	N/A	N/A	N/A
Activation	ReLU	ReLU	ReLU	SoftMax	Linear
Dropout	0	0	0.4	0	0
Filters	32	32	N/A	N/A	N/A
Data format	Channel last	Channel last	N/A	N/A	N/A
Padding	Same	Same	N/A	N/A	N/A
Neurons	N/A	N/A	64	2	1

Table S3. Un-refined REDAN paths with respective costs.

Source	Cost	Residues
59	149.2524	59, 50, 54, 51, 349, 379, 375, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503, 127
60	116.4176	60, 64, 61, 65, 62, 47, 349, 379, 375, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503, 127
61	70.3623	61, 65, 62, 47, 349, 379, 375, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503, 127
62	50.9962	62, 47, 349, 379, 375, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503, 127
63	88.1584	63, 66, 62, 47, 349, 379, 375, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503, 127
64	89.9224	64, 61, 65, 62, 47, 349, 379, 375, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503, 127
65	60.2454	65, 62, 47, 349, 379, 375, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503, 127
66	82.4471	66, 62, 47, 349, 379, 375, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503, 127
67	100.0196	67, 71, 74, 78, 75, 79, 76, 32, 100, 96, 99, 102, 202, 190, 187, 509, 506, 503, 127
68	60.6018	68, 72, 32, 100, 96, 99, 102, 202, 190, 187, 509, 506, 503, 127
69	62.2136	69, 72, 32, 100, 96, 99, 102, 202, 190, 187, 509, 506, 503, 127
70	69.4227	70, 74, 78, 75, 79, 76, 32, 100, 96, 99, 102, 202, 190, 187, 509, 506, 503, 127
71	60.9140	71, 74, 78, 75, 79, 76, 32, 100, 96, 99, 102, 202, 190, 187, 509, 506, 503, 127
72	47.4559	72, 32, 100, 96, 99, 102, 202, 190, 187, 509, 506, 503, 127
73	58.6718	73, 77, 80, 76, 32, 100, 96, 99, 102, 202, 190, 187, 509, 506, 503, 127
74	55.7830	74, 78, 75, 79, 76, 32, 100, 96, 99, 102, 202, 190, 187, 509, 506, 503, 127
75	48.9291	75, 79, 76, 32, 100, 96, 99, 102, 202, 190, 187, 509, 506, 503, 127
76	45.9068	76, 32, 100, 96, 99, 102, 202, 190, 187, 509, 506, 503, 127
77	54.1053	77, 80, 76, 32, 100, 96, 99, 102, 202, 190, 187, 509, 506, 503, 127
344	57.5303	344, 368, 346, 359, 332, 360, 307, 310, 314, 373, 409, 374, 405, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503, 127
345	57.1416	345, 368, 346, 359, 332, 360, 307, 310, 314, 373, 409, 374, 405, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503, 127
346	46.9889	346, 359, 332, 360, 307, 310, 314, 373, 409, 374, 405, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503, 127
347	35.7958	347, 375, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503, 127
348	38.0116	348, 375, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503, 127
349	37.9870	349, 379, 375, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503, 127
350	31.3724	350, 382, 385, 401, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503, 127
351	61.8150	351, 357, 348, 375, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503,

		127
352	61.1138	352, 41, 38, 35, 32, 100, 96, 99, 102, 202, 190, 187, 509, 506, 503, 127
353	61.1542	353, 41, 38, 35, 32, 100, 96, 99, 102, 202, 190, 187, 509, 506, 503, 127
355	62.2056	355, 41, 38, 35, 32, 100, 96, 99, 102, 202, 190, 187, 509, 506, 503, 127
356	37.8468	356, 382, 385, 401, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503, 127
357	39.6737	357, 348, 375, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503, 127
358	43.5352	358, 48, 357, 348, 375, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503, 127
359	41.8726	359, 332, 360, 307, 310, 314, 373, 409, 374, 405, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503, 127
360	37.5957	360, 307, 310, 314, 373, 409, 374, 405, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503, 127
361	71.4523	361, 344, 368, 346, 359, 332, 360, 307, 310, 314, 373, 409, 374, 405, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503, 127
363	60.5551	363, 344, 368, 346, 359, 332, 360, 307, 310, 314, 373, 409, 374, 405, 378, 402, 399, 517, 513, 516, 453, 512, 505, 509, 506, 503, 127

Table S4. Cost comparison for best path between REDAN and ML-weighted REDAN.

Path	Cost
REDAN Path (Raw)	31.3724
REDAN Path (State 2, ML weighted)	31.1855
REDAN Path (State 3, ML weighted)	29.2688

Machine learning model regularization

Dropout avoids the co-adaptation of neurons in feature detection by randomly dropping a percentage of neurons during the training process. This procedure ensures the lack of overfitting by forcing the network to not rely only on specific neurons.

On the other hand, L1 regularizer, or lasso regression, adds a penalty term to the layers loss functions that will force the least important features to zero. The L1 regularizer has the form:

$$\sum_{i=1}^N (y_i - \sum_j x_{ij} \beta_j)^2 + \lambda \sum_{j=1}^p |\beta_j|$$

In the above equation i are the instances, j the number of features, y is the real target value, $x_{ij} \beta_j$ forms the predicted y with x being the input data and β the model parameters. While the left side of the formula represent a standard linear regression, the right side represents the penalization term, with λ being shrinkage parameter. This implementation, on top of functioning as a data regularization to avoid overfitting, is expected to help selecting important bonding pattern within the network as it forces the values of unimportant features to zero.

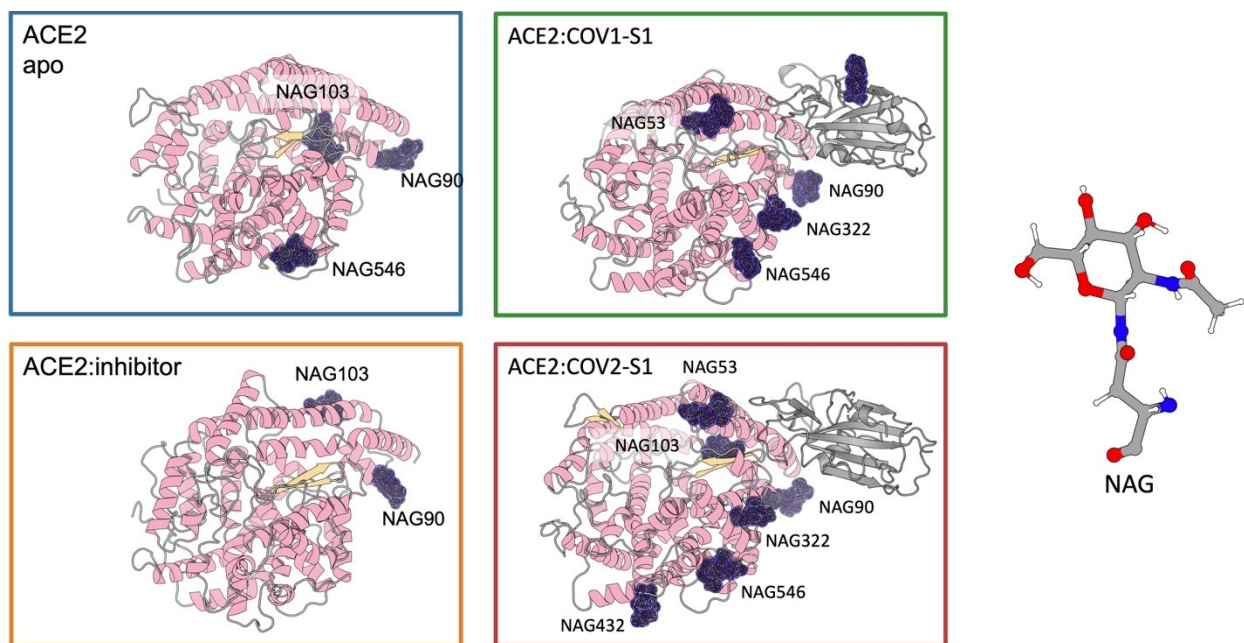


Fig. S1. Glycosylation profile of the systems analyzed.

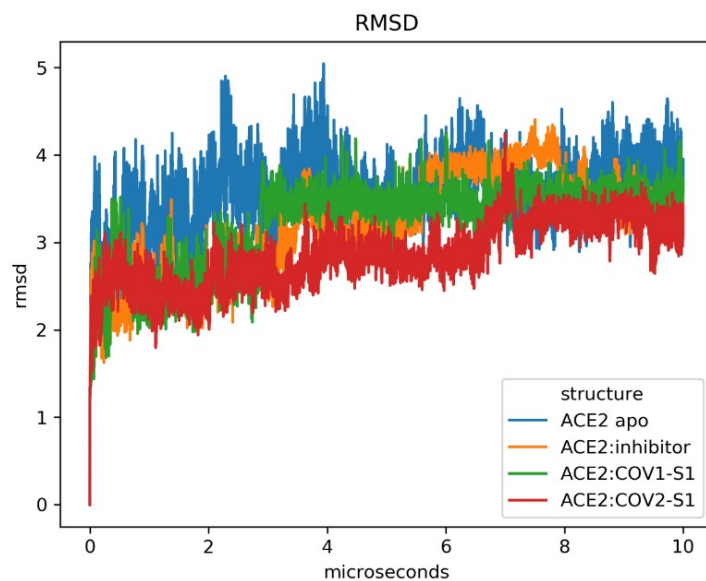


Fig. S2. Root Mean Square Deviation (RMSD) analysis of ACE2 PD in different systems.

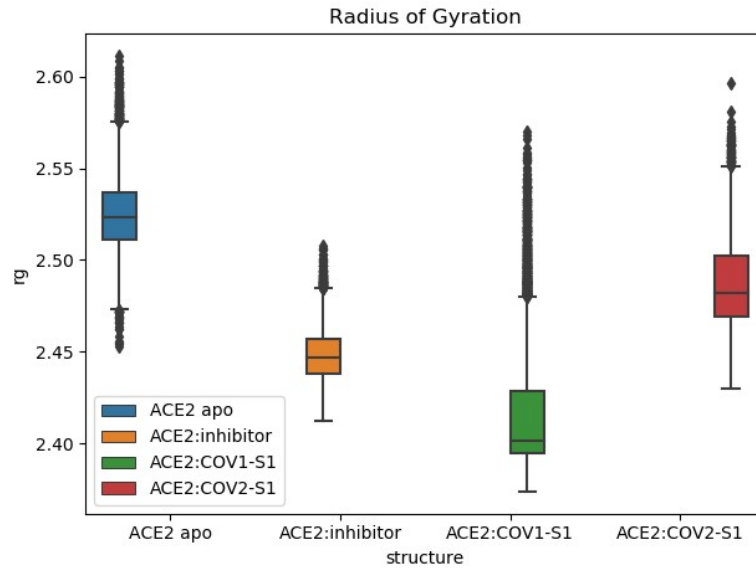


Fig. S3. Radius of Gyration analysis for ACE2 PD in multiple systems.

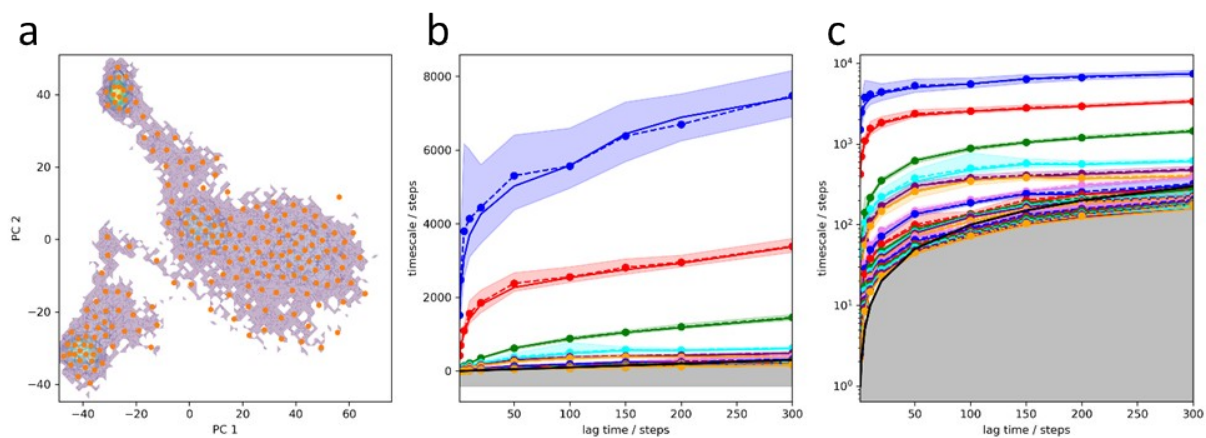


Fig. S4. Kinetic clustering of ACE2 PD conformational space. a) K-means microclustering of ACE2 PD conformational space. b) Relaxation timescales of ACE2 PD at different lag times. c) Log values of relaxation timescales of ACE2 PD at different lag times.

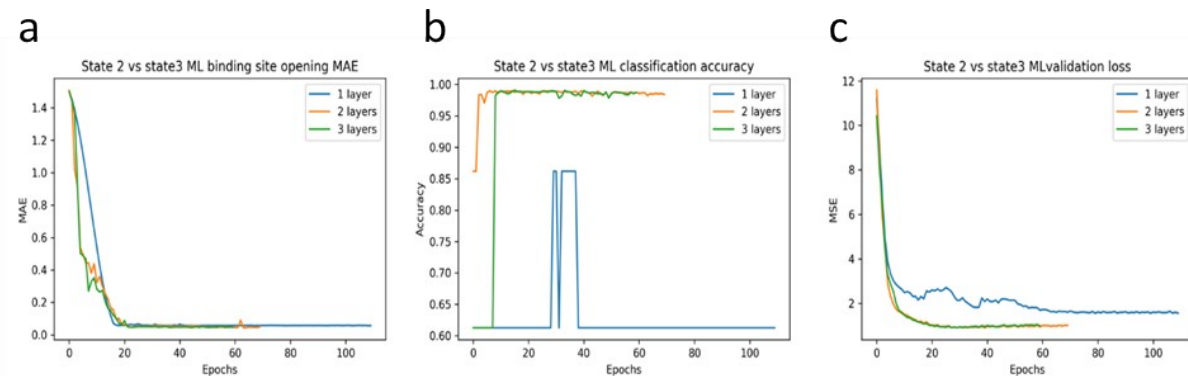


Fig. S5. Validation of CV-CNN model. a) Mean Average Error of binding site opening regression task. b) Classification accuracy; c) Model validation loss.

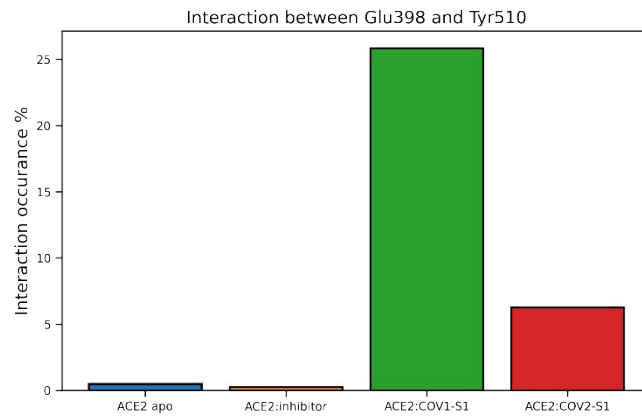


Fig. S6. Interaction between Glu398 and Tyr510. This interaction is measured as percentage of frames where these two residues are found at a distance lower than 4 Å.

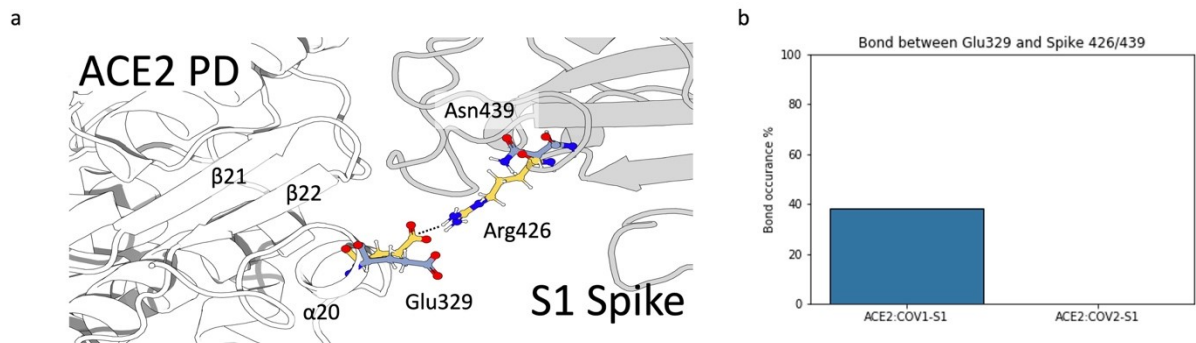


Fig. S7. a) Interaction between ACE2 PD Glu329 with SARS-CoV-1 spike Arg426 (yellow) vs SARS-CoV-2 spike Asn439 (blue). b) Percentage of bond occurrence between ACE2 PD Glu329 with SARS-CoV-1 spike Arg426 (ACE2:COV1-S1) vs SARS-CoV-2 spike Asn439 (ACE2:COV2-S1). The percentage is measured based on the number of frames where the pair of residues was at a distance smaller than 4 Å.