

An insertion unique to SARS-CoV-2 exhibits superantigenic character strengthened by recent mutations

Mary Hongying Cheng¹, She Zhang¹, Rebecca A. Porritt^{2,3}, Moshe Ardit^{2,3*} and Ivet Bahar^{1*}

¹Department of Computational and Systems Biology, School of Medicine, University of Pittsburgh, Pittsburgh, PA 15261. Departments of ²Pediatrics, Division of Pediatric Infectious Diseases and Immunology, ³Biomedical Sciences, Infectious and Immunologic Diseases Research Center, Cedars-Sinai Medical Center, Los Angeles, CA 90048.

**Corresponding authors*

Moshe Ardit, MD

Executive-Vice Chair for Research, Department of Pediatrics, for Research

Prof of Pediatrics, Cedars-Sinai Medical Center

GUESS?/Fashion Industries Guild Chair in Community Child Health

Director, Infectious and Immunological Diseases Research Center (IIDRC), BMS

Department of Biomedical Sciences (BMS)

Director, Division of Pediatric Infectious Diseases and Immunology and

Kawasaki Disease Research Program

Moshe.arditi@cshs.org

Cedars-Sinai

Davis Building, D4018: Los Angeles, CA 90048

office 310-423-2593 : cedars-sinai.org

Dr. Ivet Bahar

Distinguished Professor and John K. Vries Chair

Department of Computational & Systems Biology

School of Medicine, University of Pittsburgh

3064 Biomedical Science Tower 3

3501 Fifth Avenue, Pittsburgh, PA 15213

Voice: 4126483332 - Fax: 4126483163

bahar@pitt.edu; <http://www.cccb.pitt.edu/Faculty/bahar/>

Supplemental Methods

Generation of a binary complex between SARS-CoV-2 spike and T cell receptor (TCR)

SARS-CoV-2 spike model in the prefusion state was generated using SwissModel¹ based on the resolved cryo-EM structure (Protein Data Bank (PDB): 6VSB²) for the spike glycoprotein where one of the receptor binding domains (RBDs) is in the up conformation. The structure of the T cell receptor (TCR) containing both TCR α and TCR β chains was taken from the crystal structure of the ternary complex between human TCR, staphylococcal enterotoxin H (SEH) and human major histocompatibility complex class II (MHCII) molecule³. Using protein-protein docking software ClusPro⁴, we constructed *in silico* a series of binary complexes for SARS-CoV-2 spike and TCR. We obtained 30 clusters of conformations for spike-TCR binary complexes, upon clustering ~1000 models generated by ClusPro. The clusters were rank-ordered by cluster size, as recommended⁴. We analyzed all models and found that the majority (>90%) showed that TCR bound to spike via its constant domain. Given that the constant domain is proximal to the cell membrane and TCR employs the variable domain for binding superantigens and/or antigen/MHC complexes³, we then added restraints to our docking simulations to prevent the binding of TCR constant domain and filter out those conformers where the variable domain would bind to the spike. This led to 27 clusters (based on a set of 666 models) from ClusPro. Interestingly, 45% of models showed the binding of TCR near the region of “PRRA” insert and 46% of models showed the binding of TCR within multiple RBDs. Thus, we identified two hot spots for TCR binding within SARS-CoV-2 spike: one is near “PRRA” insert and the other within the RBD. Representative members belonging to the top-ranking clusters are presented in [Fig S1](#).

Generation of a ternary complex between SARS-CoV-2 spike, TCR, and MHCII

Structure of the human MHCII was taken from the crystal structure of the ternary complex³ (PDB: 2XN9) between human TCR, SEH and MHCII. First, we performed docking simulations to generate binary complexes between MHCII and SARS-CoV-2 spike. Six representative MHCII-spike binary complexes were selected to explore further docking of TCR to form a ternary complex. We analyzed all predicted ternary complex models of MHCII-Spike-TCR. Potential tertiary MHCII-Spike-TCR complex models were selected following three filtering criteria: (1) TCR either binds near “PRRA” insert region or the RBD; (2) the binding regions involve homologous superantigen or toxin binding motifs predicted for SARS-CoV ([Fig. S2](#)); (3) MHCII and TCR are in close proximity. These filters led to the MHCII-Spike-TCR complex model illustrated in [Fig. S3A](#). Interestingly, the SARS-CoV-2 spike binding region harbors three residues that have been recently reported to have mutated in new strains from Europe and USA^{5,6} ([Fig. S3B](#)): D614G, A831V and D839Y/N/E). While we do not exclude the possible occurrence of other potential tertiary complexes, especially those involving the RBDs, we focus here on the complex shown in [Fig. S3](#), which uniquely satisfied all three aforementioned criteria.

***In silico* mutagenesis of D839 of SARS-CoV-2 spike**

We mutated D839 of the SARS-CoV-2 spike *in silico* to asparagine, glutamic acid and tyrosine in line with the aforementioned mutant D839Y/N/E observed in a new strain from Europe. To this aim, we used PyMOL mutagenesis tool⁷ and evaluated the change in local conformation and energetics in the complex formed with TCR. The most probable rotamers were selected and energetically minimized in the presence of the bound TCR (conformation shown in **Fig. 1**) using OpenMM⁸. Binding affinities (ΔG) and dissociation constants (K_d) were obtained for (i) the full complex (with the intact spike and entire TCR as interactors) or (ii) a single spike subunit and TCRV β with the D839Y/N/E mutation on spike at 37 °C using PRODIGY server^{9,10}.

Supplemental Figures and Tables

Table S1: Binding affinities predicted for the interactions between the $\alpha\beta$ TCR and SARS-CoV-2 spike before/after the point mutation D839Y/N/E.

	Aspartic Acid (D)		Tyrosine (Y)		Glutamic Acid (E)		Asparagine (N)	
	ΔG (kcal mol ⁻¹)	K_d (nM)	ΔG (kcal mol ⁻¹)	K_d (nM)	ΔG (kcal mol ⁻¹)	K_d (nM)	ΔG (kcal mol ⁻¹)	K_d (nM)
Full complex	-11.0	18	-13.2	0.46	-13.1	0.56	-12.3	2.3
S subunit - TCR Vβ	-8.8	580	-10.3	53	-10.1	80	-9.5	190

* Binding affinities (ΔG) and dissociation constants (K_d) were obtained at 37 °C using PRODIGY server^{9,10}.

Supplemental Figures

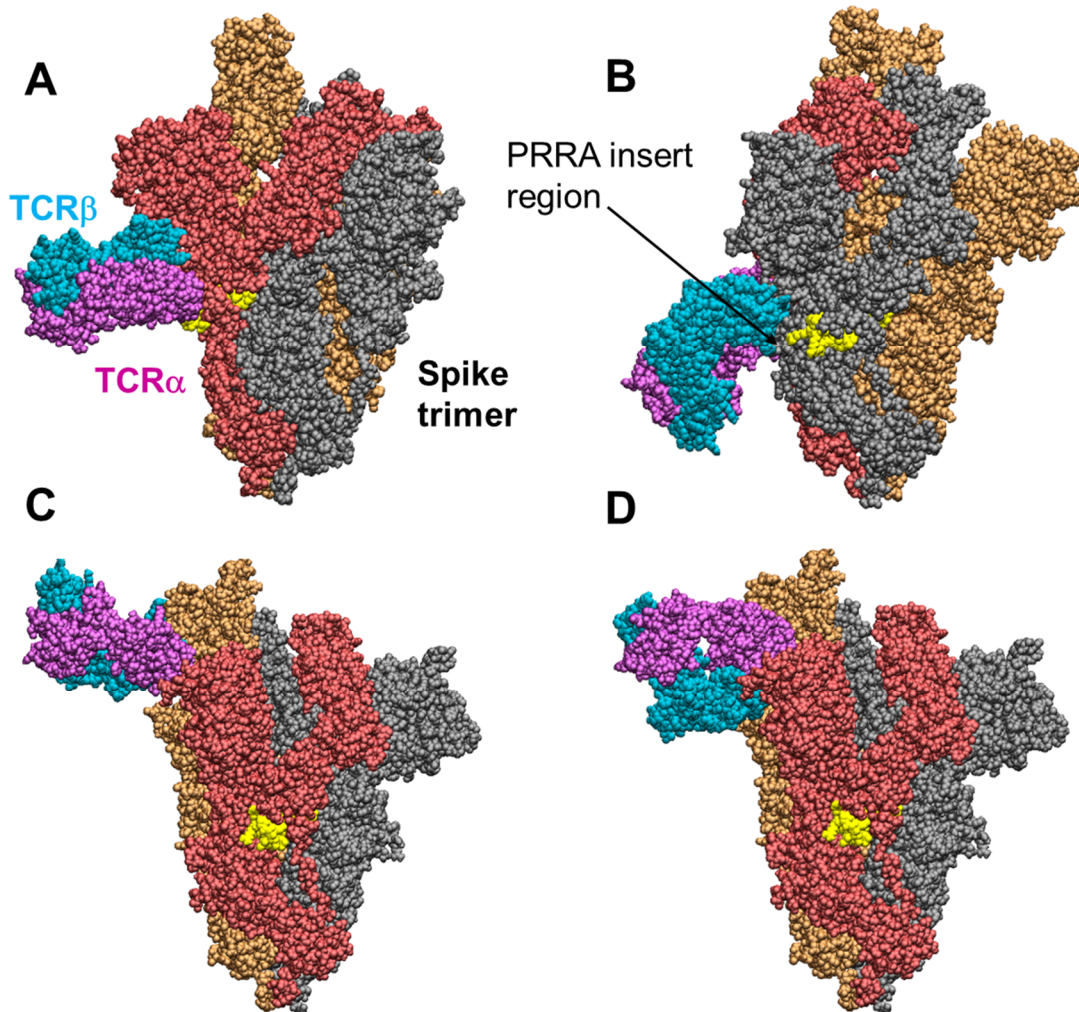


Figure S1: Top-ranking binary complexes of SARS-CoV-2 spike with the T cell receptor (TCR) predicted by ClusPro. (A-B) Binding of TCR near the “PRRA” insert region. (C-D) Binding of TCR near the RBD of a subunit. The spike trimer subunits are colored red, orange, and gray. The PRRA insert region (E661 to R685) is shown in yellow. TCR α - and β -chains are shown in cyan and magenta. See more details for the interaction between the PRRA insert region and TCR in Fig. 1.

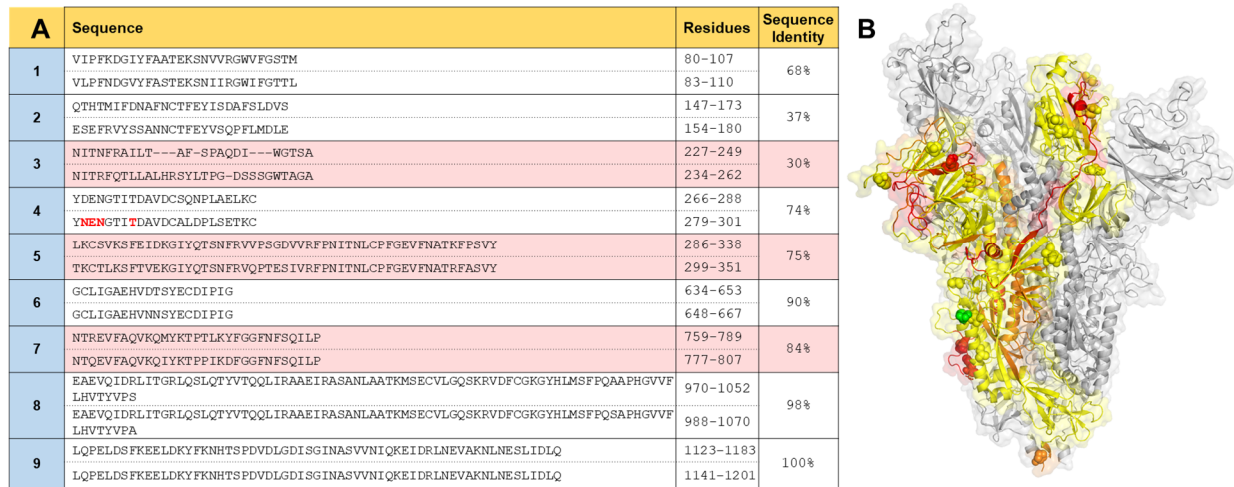


Figure S2: Motifs associated with superantigen, toxin, cytokine, and membrane surface proteins predicted for SARS-CoV spike and mapped onto SARS-CoV-2 spike sequence and structure. (A) Sequence alignment of these motifs on SARS-CoV (*upper rows*) and SARS-CoV-2 spikes (*lower rows*), corresponding residue numbers (3rd column) and sequence identity (4th/last column). Superantigenic and toxic-like motifs are highlighted in *pink*. Residues that interact with TCR V α are marked in *red*. **(B)** Predicted motifs mapped onto the trimeric structure of SARS-CoV-2 spike, with one of its subunit colored in *yellow*. The motifs are colored *red* (superantigenic and toxic-like) or *orange* (others). Mutation sites reported in recent work^{5,6} are shown in *spheres*. The mutation site D839Y/N/E is highlighted in *green*.

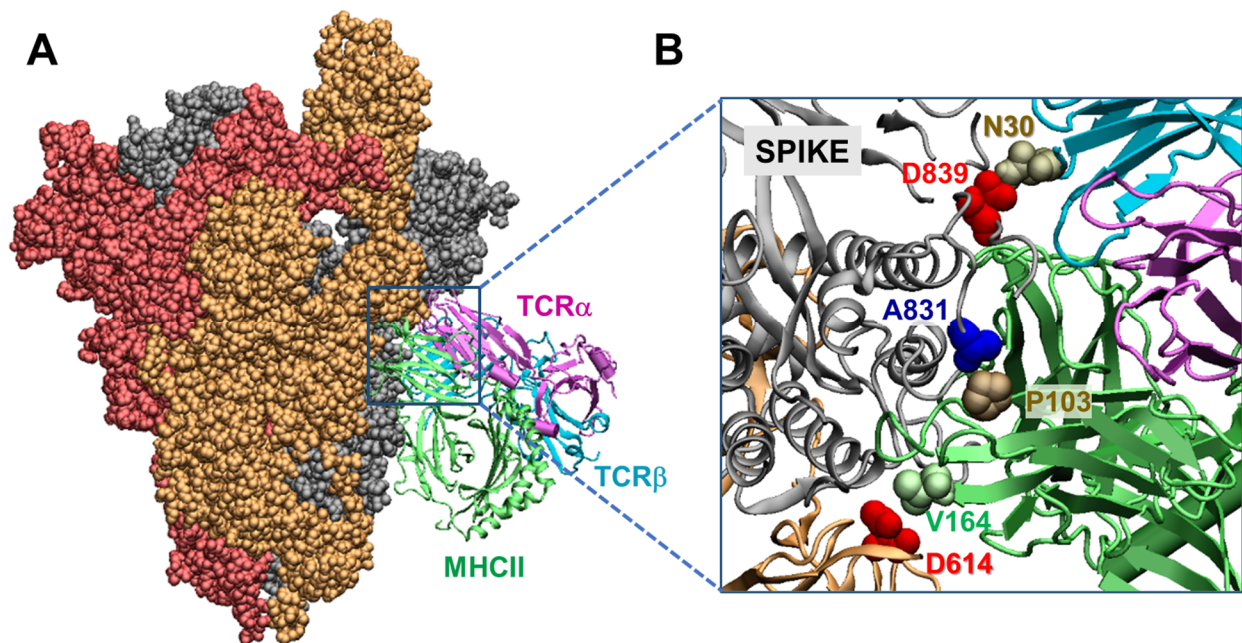


Figure S3: Modeled ternary complex formed by SARS CoV-2 spike, MHCII and $\alpha\beta$ TCR. (A) Side view of the complex. The trimeric spike subunits are shown in the same style and color as in Fig. S1, as well as TCR α and TCR β . The TCR retains a similar pose as in panel B in Fig. S1, now rotated by 180° along the z-axis. MHCII is displayed in *green* ribbon diagram. (B) Top view of the interfacial contacts. Note that three spike residues reported to mutate in recent strains observed in European and western countries^{5,6} (D614G, A831V and D839Y/N/E) are within 3-5 Å from either MHCII or TCR V β .

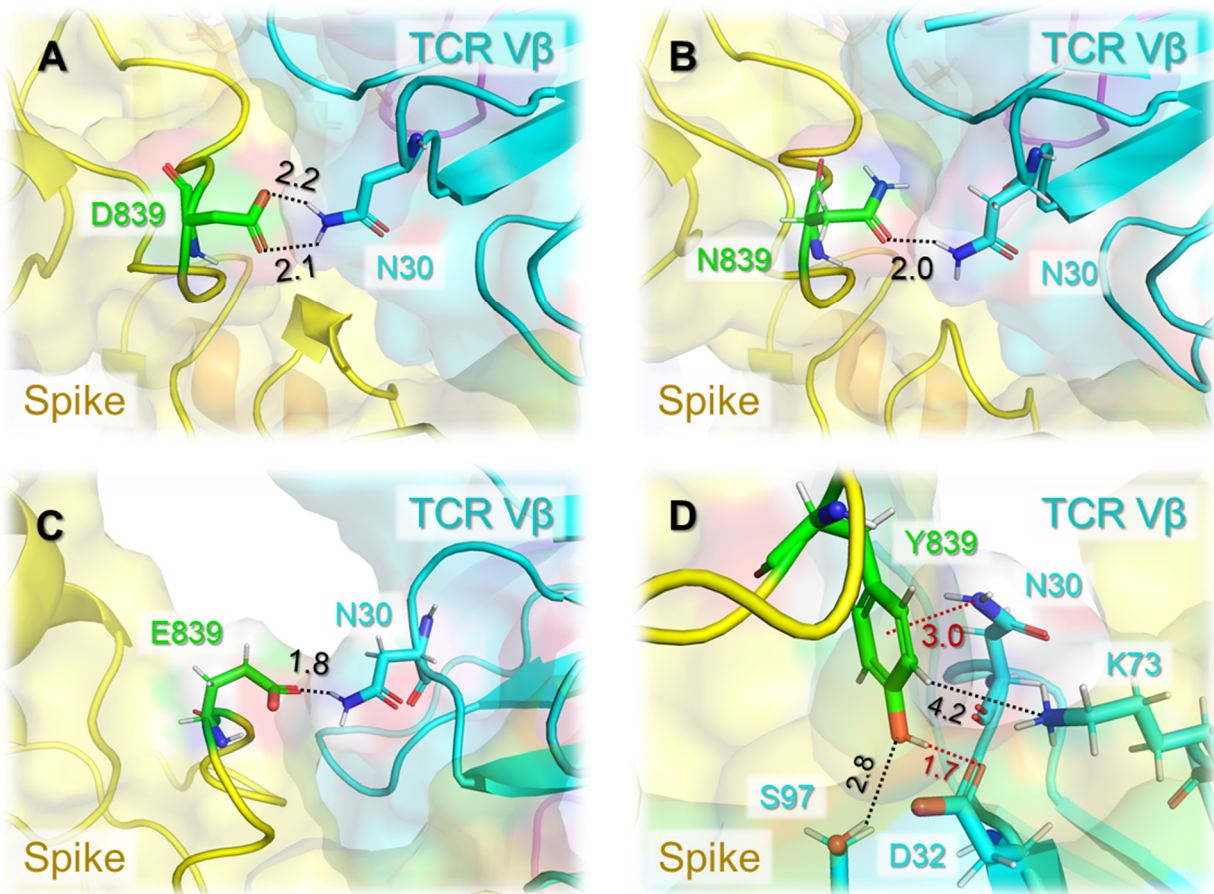


Figure S4: *In silico* mutagenesis analysis of SARS-CoV-2 spike protein residue D839Y/N/E. (A) A close-up view of the interaction between the wild type residue D839 in the spike and N30 of TCR V β . (B-D) Results obtained upon mutation to asparagine, glutamic acid, and tyrosine. The spike and TCR V β are shown in *yellow* and *cyan*, respectively. The mutation site is highlighted in *green*. Atomic interactions are indicated by *black dashed lines* along with their distances in Ångstroms.

Supplemental References

- 1 Bordoli, L. & Schwede, T. in *Homology Modeling* 107-136 (Springer, 2011).
- 2 Wrapp, D. *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science (New York, N.Y.)* **367**, 1260-1263, doi:10.1126/science.abb2507 (2020).
- 3 Salane, M. *et al.* The structure of superantigen complexed with TCR and MHC reveals novel insights into superantigenic T cell activation. *Nat. Commun* **1**, 119, doi:10.1038/ncomms1117 (2010).
- 4 Kozakov, D. *et al.* The ClusPro web server for protein–protein docking. *Nat. Protocols* **12**, 255 (2017).
- 5 Zhan, S. H., Deverman, B. E. & Chan, Y. A. SARS-CoV-2 is well adapted for humans. What does this mean for re-emergence? *bioRxiv*, 2020.2005.2001.073262, doi:10.1101/2020.05.01.073262 (2020).
- 6 Korber, B. *et al.* Spike mutation pipeline reveals the emergence of a more transmissible form of SARS-CoV-2. *bioRxiv*, 2020.2004.2029.069054, doi:10.1101/2020.04.29.069054 (2020).
- 7 DeLano, W. L. Pymol: An open-source molecular graphics tool. *CCP4 Newsletter on protein crystallography* **40**, 82-92 (2002).
- 8 Eastman, P. *et al.* OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS computational biology* **13**, e1005659 (2017).
- 9 Vangone, A. & Bonvin, A. M. Contacts-based prediction of binding affinity in protein–protein complexes. *elife* **4**, e07454 (2015).
- 10 Xue, L. C., Rodrigues, J. P., Kastritis, P. L., Bonvin, A. M. & Vangone, A. PRODIGY: a web server for predicting the binding affinity of protein–protein complexes. *Bioinformatics* **32**, 3676-3678 (2016).