

## Article

# Potential Autoimmunity Resulting from Molecular Mimicry between SARS-CoV-2 Spike and Human Proteins

Janelle Nunez-Castilla <sup>1</sup>, Vitalii Stebliankin <sup>2</sup>, Prabin Baral <sup>3</sup>, Christian A. Balbin <sup>1</sup>, Masrur Sobhan <sup>4</sup>, Trevor Cickovski <sup>2</sup>, Ananda Mohan Mondal <sup>2,4,5</sup>, Giri Narasimhan <sup>2,5</sup>, Prem Chapagain <sup>3,5</sup>, Kalai Mathee <sup>5,6</sup> and Jessica Siltberg-Liberles <sup>1,5,\*</sup>

<sup>1</sup> Department of Biological Sciences, College of Arts, Sciences and Education, Florida International University, 11200 SW 8th St, Miami, FL 33199, USA; jnune033@fiu.edu (J.N.-C.); cbalbin@fiu.edu (C.A.B.)

<sup>2</sup> Bioinformatics Research Group (BioRG), Knight Foundation School of Computing and Information Sciences, Florida International University, 11200 SW 8th St, Miami, FL 33199, USA; vsteb002@fiu.edu (V.S.); tcickovs@fiu.edu (T.C.); amondal@fiu.edu (A.M.M.); giri@fiu.edu (G.N.)

<sup>3</sup> Department of Physics, College of Arts, Sciences and Education, Florida International University, 11200 SW 8th St, Miami, FL 33199, USA; pbara006@fiu.edu (P.B.); chapagap@fiu.edu (P.C.)

<sup>4</sup> Machine Learning and Data Analytics Group (MLDAG), Knight Foundation School of Computing and Information Sciences, Florida International University, 11200 SW 8th St, Miami, FL 33199, USA; msobh002@fiu.edu

<sup>5</sup> Biomolecular Sciences Institute, Florida International University, 11200 SW 8th St, Miami, FL 33199, USA; matheek@fiu.edu

<sup>6</sup> Department of Human and Molecular Genetics, Herbert Wertheim College of Medicine, Florida International University, 11200 SW 8th St, Miami, FL 33199, USA

\* Correspondence: jliberle@fiu.edu



**Citation:** Nunez-Castilla, J.; Stebliankin, V.; Baral, P.; Balbin, C.A.; Sobhan, M.; Cickovski, T.; Mondal, A.M.; Narasimhan, G.; Chapagain, P.; Mathee, K.; et al. Potential Autoimmunity Resulting from Molecular Mimicry between SARS-CoV-2 Spike and Human Proteins. *Viruses* **2022**, *14*, 1415. <https://doi.org/10.3390/v14071415>

Academic Editors: Franziska Hufsky, Alba Pérez-Cataluña, Walter Randazzo, Gloria Sanchez, Fernando González-Candelas and Manja Marz

Received: 30 April 2022

Accepted: 22 June 2022

Published: 28 June 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Molecular mimicry between viral antigens and host proteins can produce cross-reacting antibodies leading to autoimmunity. The coronavirus SARS-CoV-2 causes COVID-19, a disease curiously resulting in varied symptoms and outcomes, ranging from asymptomatic to fatal. Autoimmunity due to cross-reacting antibodies resulting from molecular mimicry between viral antigens and host proteins may provide an explanation. Thus, we computationally investigated molecular mimicry between SARS-CoV-2 Spike and known epitopes. We discovered molecular mimicry hotspots in Spike and highlight two examples with tentative high autoimmune potential and implications for understanding COVID-19 complications. We show that a TQLPP motif in Spike and thrombopoietin shares similar antibody binding properties. Antibodies cross-reacting with thrombopoietin may induce thrombocytopenia, a condition observed in COVID-19 patients. Another motif, ELDKY, is shared in multiple human proteins, such as PRKG1 involved in platelet activation and calcium regulation, and tropomyosin, which is linked to cardiac disease. Antibodies cross-reacting with PRKG1 and tropomyosin may cause known COVID-19 complications such as blood-clotting disorders and cardiac disease, respectively. Our findings illuminate COVID-19 pathogenesis and highlight the importance of considering autoimmune potential when developing therapeutic interventions to reduce adverse reactions.

**Keywords:** vaccine design; AlphaFold2; coronavirus; molecular dynamics; machine learning; protein structure comparison; COVID-19 complications; variant; omicron; tropomyosin

## 1. Introduction

The coronavirus SARS-CoV-2 is the causative agent of the COVID-19 pandemic. COVID-19 is an infectious disease whose typical symptoms include fever, cough, shortness of breath [1,2], and loss of taste or smell [3]. Curiously, despite over half a billion confirmed cases worldwide [4], roughly one-third are estimated to be asymptomatic [5]. Yet, other SARS-CoV-2 infected individuals may also experience a variety of disease-related complications including liver injury [6], kidney injury [7], and cardiovascular complications

including myocarditis, heart failure, thrombosis [8], and thrombocytopenia [9]. COVID-19 can trigger a range of antibody response levels [10] and an enrichment in autoantibodies that react with human proteins has been found for patients with severe disease [11]. While the reason for the variety of disease severity affecting people with COVID-19 is not well understood, molecular mimicry may provide an avenue for explanations.

Molecular mimicry occurs when unrelated proteins share regions of high molecular similarity, such that they can perform similar and unexpected interactions with other proteins. When molecular mimicry involves antigens to which antibodies are produced, cross-reactive antibodies can result. Molecular mimicry between pathogen antigens and human proteins can cause an autoimmune response, where antibodies against the pathogen erroneously interact with human proteins, sometimes leading to transient or chronic autoimmune disorders [12]. Alternatively, molecular mimicry could be viewed through the lens of heterologous immunity, where previous exposure to one pathogen antigen can result in short- or long-term complete or partial immunity to another pathogen with a similar antigen [13]. Moreover, antigen-driven molecular mimicry can also lead to cross-reactive antibody immunity which has been reported against SARS-CoV-2 for uninfected individuals [14].

The SARS-CoV-2 Spike protein is responsible for enabling SARS-CoV-2 entry into host cells [15]. Spike protrudes from the virus surface and is one of the main antigenic proteins for this virus [16]. Additionally, Spike is the primary component in the vaccines against SARS-CoV-2. Consequently, molecular mimicry between Spike and human proteins or Spike and other human pathogens can result in cross-reactive antibodies in response to SARS-CoV-2 infection or vaccination. Cross-reactive antibodies may yield complex outcomes such as diverse symptoms with varying severity across populations and developmental stages as observed for COVID-19. It must be noted that there are a variety of genetic and environmental factors that contribute to an individual's likelihood to develop an autoimmune response [17]. Still, identifying autoimmune potential and heterologous immunity through instances of molecular mimicry between Spike and proteins from humans or human pathogens can serve to better understand disease pathogenesis, improve therapeutic treatments, and inform vaccine design as they relate to SARS-CoV-2 infection. Previous studies have predicted molecular mimicry between SARS-CoV-2 Spike and human proteins using sequence similarity [18] to known epitopes in the Immune Epitope Database (IEDB) [19] and sequence and structural similarity in general [20,21]. We combine these approaches and investigate molecular mimicry between Spike and human proteins by considering both sequence and structural similarity and searching against known epitopes from IEDB [19]. We define molecular mimicry as a match of at least five identical consecutive amino acids that appear in both Spike and in a known epitope, where at least three amino acids are surface accessible on Spike and the match from the epitope has high structural similarity to the corresponding sequence from Spike. In light of our findings, we discuss the autoimmune potential and heterologous immunity with implications for vaccine design and the side effects of SARS-CoV-2 infection.

## 2. Methods

### 2.1. Identifying Epitopes with Molecular Mimicry

To identify known epitopes with positive assays from IEDB, we used Epitepedia [22] with a full-length Cryo-EM structure of Spike from SARS-CoV-2 (PDB id: 6XR8, chain A, RBD: 0up3down (solved residues: 14–69, 77–244, 254–618, 633–676, 689–1162) [23]) as input. Hits containing 5 or more consecutive residues with 100% sequence identity where at least 3 of the input residues are surface accessible are considered sequence-based molecular mimics (termed as “1D-mimics”). For all 1D-mimics with corresponding structural representation from either the Protein Data Bank (PDB) [24] or AlphaFold2 [25] 3D models of human proteins, TM-align [26] was used to generate a structural alignment and Root Mean Square Deviation (RMSD) for all input-hit (1D-mimic) alignment pairs using only the structural regions corresponding to the hit for the source antigenic protein

containing the epitope and the input. Epitopes with an RMSD  $\leq 1$  Å to Spike were considered structure-based molecular mimics (termed as “3D-mimics”).

### 2.2. Conformational Ensemble of TQLPP Structural Mimicry

To gather all structures of the TQLPP motif in Spike, an NCBI BLASTP search against PDB was performed with the SARS-CoV-2 Spike reference sequence as the query and a SARS-CoV-2 taxa filter. Of 75, close to full-length, hits (>88% query cover), 20 included a solved structure for the TQLPP motif. The TQLPP region of the PDB structure was extracted for all chains in the 20 structures (all were trimers, as in Spike’s biological state) resulting in a TQLPP Spike ensemble of 60 different chains from SARS-CoV-2. Each sequence in the TQLPP Spike ensemble was superimposed with chain X of the two PDB structures of human thrombopoietin (hTPO, PDB ids: 1V7M and 1V7N) to generate an RMSD value distribution for Spike’s conformational ensemble vs. hTPO for the structural mimicry region (Table S1).

### 2.3. Modeling Spike-Antibody Complexes

We constructed a composite model of the Spike-TN1 complex using the hTPO-TN1 complex (PDB id: 1V7M) as a template. For this, we first aligned the TQLPP segment of hTPO in the hTPO-TN1 complex with the TQLPP segment of the fully glycosylated model of Spike (PDB id: 6VSB [27]) retrieved from the CHARMM-GUI Archive [28]. We then removed hTPO, leaving TN1 complexed with Spike. For the Spike-TN1 simulations, only the TN1 interacting N-terminal domain of Spike (residues 1–272) was considered. Similarly, a composite model of the Spike-S2P6 complex was modeled by using the stem helix-S2P6 complex (PDB id: 7RNJ [29]) as a template. As with the TQLPP segment, we aligned ELDKY in the stem helix-S2P6 complex with ELDKY in the stem helix segment of Spike (PDB id: 6XR8) retrieved from the Protein Data Bank. We then removed the stem helix segment from the stem helix-S2P6 complex, leaving S2P6 complexed with Spike. For the Spike-S2P6 simulations, only the S2P6 interacting stem helix segment of Spike (residues 1146–1159) was considered. Geometrical alignments, as well as visualization, were performed with PyMOL version 2.5.0 [30] and Visual Molecular Dynamics (VMD 1.9.3 [31]).

To confirm that the modeled Spike TQLPP region is in agreement with the TQLPP region of solved Spike structures, these regions were extracted. TM-align was used to superimpose the TQLPP regions from the different structures, including the modeled TQLPP region from the Spike-TN1 complex, and to calculate the respective RMSD values. Three states of the model were included (before and after equilibration, and after molecular dynamics (described in the following paragraph)) together with the 60 experimentally determined Spike structures in Table S1 and compared in an all-against-all manner (Figure S1, Table S2). A Mann-Whitney U test was used to compare the TQLPP region from 60 experimentally determined Spike structures based on RBD state: (1) both down, (2) 1 down and 1 up, (3) both up. (Figure S2). Further, TM-align was used to calculate RMSD between wild-type TQLPP (PDB id: 6XR8, chain A) and the corresponding region in known variants of concern with available structures (Table S3).

### 2.4. Molecular Dynamics Simulation

A simulation system for the modeled Spike-antibody systems were prepared using CHARMM-GUI [32–34]. The complexes were solvated using a TIP3P water model and 0.15 M concentration of KCl and equilibrated for 1 ns at 303 K. All-atom simulations were performed with NAMD 2.14 [35] using CHARMM36m force-field. The production runs were performed under the constant pressure of 1 atm, controlled by a Nose–Hoover Langevin piston [36] with a piston period of 50 fs and a decay of 25 fs to control the pressure. The temperature was set to 303 K and controlled by Langevin temperature coupling with a damping coefficient of 1/ps. The Particle Mesh Ewald method (PME) [37] was used for long-range electrostatic interactions with periodic boundary conditions and all covalent bonds with hydrogen atoms were constrained by Shake [38]. The contact area

of the interface was calculated as  $(S_1 + S_2 - S_{12})/2$ , where  $S_1$  and  $S_2$  represent the solvent accessible surface areas of the antigen and antibody and  $S_{12}$  represents that for the complex (Figure S3). We performed MD simulations of the hTPO-TN1 complexes (PDB ids: 1V7M and 1V7N) as well as the Spike-TN1 complexes modeled from PDB ids: 1V7M and 1V7N to generate interaction matrices of protein-antibody hydrogen bonds during the last 50 ns of 200 ns MD simulation for each run.

### 2.5. Binding Affinity

The PRODIGY webserver [39] was used to calculate the binding affinity and intermolecular contacts for Spike-TN1 (described above) and hTPO-TN1 complexes (PDB ids: 1V7M and 1V7N) at the TQLPP region. We retrieved five intermediate structures from 200 ns MD simulations of each of these complexes at an interval of 40 ns. Similarly, PRODIGY was used to calculate the binding affinity and intermolecular contacts for the modeled Spike-S2P6 complex (from PDB id 7RNJ [29]) at the EDLKY region. We retrieved five intermediate structures from a 50 ns MD simulation at an interval of 10 ns. The average binding affinity for each complex is reported (Table S4).

### 2.6. Antibody Interface Complementarity

We used the MaSIF-search geometric deep learning tool designed to uncover and learn from complementary patterns on the surfaces of interacting proteins [40]. The surface properties of proteins are captured using radial patches. A radial patch is a fixed-sized geodesic around a potential contact point on a solvent-excluded protein surface [41]. In MaSIF-search, the properties include both geometric and physicochemical properties characterizing the protein surface [40]. This tool exploits that a pair of patches from the surfaces of interacting proteins exhibit interface complementarity in terms of their geometric shape (e.g., convex regions would match with concave surfaces) and their physicochemical properties. The data structure of the patch is a grid of 80 bins with 5 angular and 16 radial coordinates and ensures that its description is rotation invariant. Each bin is associated with 5 geometric and chemical features: shape index, distance-dependent curvature, electrostatics, hydrophathy, and propensity for hydrogen bonding. The model converts patches into 80-dimensional descriptor vectors, such that the Euclidian distance between interacting patches is minimized. Here, we define the binding confidence score as a measure of distance between the descriptor vectors of the two patches. Thus, lower “MaSIF binding confidence scores” represent better complementarity and therefore better matches. The pre-trained MaSIF-search model “sc05” with a patch radius of 12 Å was used.

Using the MaSIF protocol, we evaluated complexes of the TN1 antibody bound to Spike in the TQLPP region. The antibody-antigen patch pairs were extracted using scripts from the molecular mimicry search pipeline EMoMiS [42]. To accommodate multiple Spike configurations, we extracted patches from 40 SARS-CoV-2 Spike-antibody complexes from the SabDab structural antibody database [43]. Patches centered at Q23 from Spike and W33 from TN1 were selected as representative pairs for the Spike-TN1 interaction type because this potential contact point has the most hydrogen bonds in the modeled Spike-TN1 TQLPP region. Binding confidence scores of randomly formed complexes (Random), complexes between Spike and its native antibodies (Spike-Ab), and complexes between hTPO and TN1 (hTPO-TN1) were extracted and tabulated (Table S5). The distribution of binding confidence scores from randomly formed complexes was obtained by pairing patches from random locations on Spike with patches from its antibodies. For native antibody-antigen Spike-Ab and hTPO-TN1 complexes, we obtained patch pairs from known interface regions using the MaSIF-search strategy for the selection of interacting patches [40]. Columns “Contact AB” and “Contact AG” in Table S6 indicate the residue used as the center of the patch from the antibody and the corresponding antigen.



### 2.7. Evaluating Further Cross-Reactivity

All 3D-mimics and AlphaFold2-3D-mimics (termed as “AF-3D-mimics”) were split into pentapeptides (if mimicry motif exceeded 5 residues) which were used as queries for NCBI BLASTP searches against the RefSeq Select [44] set of proteins from the human proteome. Results for the BLAST searches can be found in Table S7.

For the TQLPP sequence motif, 10 representative isoforms in proteins containing the complete motif were found, including hTPO. The other 9 proteins lacked a solved structure containing TQLPP. However, AlphaFold2 3D models were available for all 10 of these RefSeq Select sequences [25,45], allowing us to extract the region corresponding to TQLPP in these hits and structurally superimpose this region with Spike TQLPP (from PDB id 6XR8) with TM-align as described above.

TN1-protein complexes were generated for three of the remaining 9 proteins (Fc receptor-like protein 4 (residues 190–282), serine/threonine-protein kinase NEK10 (residues 1029–1146), ALG12 (Mannosyltransferase ALG12 homolog (residues 1–488)). The TQLPP segment in hTPO was structurally aligned with each of the TQLPP segments of the modeled proteins, after which, hTPO was removed resulting in the complex of TN1 with the modeled proteins following the methods mentioned for Spike above. The equilibrated structures of these complexes show that TN1 stays firmly with these proteins without any structural clash. Further, to evaluate the shape complementarity of these three proteins and TN1, MaSIF was used to calculate binding confidence scores as described above (Table S8).

It should also be noted that two additional human genes (GeneIDs 8028 and 57110) also have one TQLPP motif, but not in the RefSeq Select isoforms. Since no structure or structural prediction was available for these proteins, they were excluded from further analysis.

For the ELDKY sequence motif, 6 additional representative isoforms containing the complete motif were found, in addition to the human proteins identified by Epitepedia to contain 3D-mimics of the motif. Solved structures of the ELDKY motif were available for 3 of the proteins, while the others had AlphaFold2 3D models available. In all instances, the region corresponding to the ELDKY motif was extracted and structurally superimposed with Spike ELDKY (from PDB id 6XR8) with TM-align as previously described.

### 2.8. Statistical Analysis

Distributions were visualized as violin plots with ggpubr (Version 0.40) and ggplot2 (Version 3.3.6) from R (Version 4.2.1). Following Shapiro-Wilk normality testing, a statistical analysis comparing the different distributions was performed using Mann-Whitney U with SciPy [46] (Version 1.7.1) from Python (Version 3), followed by a simplified Bonferroni correction ( $\alpha/n$  comparisons) when appropriate.

## 3. Results and Discussion

We used Epitepedia [22] to predict molecular mimicry for the structure of the SARS-CoV-2 Spike protein (PDB id: 6XR8, chain A, RBD: 0up3down [23]) against all linear epitopes in IEDB, excluding those from Coronaviruses. Epitepedia returned 789 sequence-based molecular mimics (termed as “1D-mimics”). One-dimensional-mimics are protein regions from epitopes that share at least five consecutive amino acids with 100% sequence identity to a pentapeptide in SARS-CoV-2 Spike, where at least three of the amino acids are surface accessible on Spike. Most 1D-mimics (627 epitopes) were found in humans. Additionally, 1D-mimics were found in non-human vertebrates (65 epitopes, 7 species), viruses (58 epitopes, 17 species), bacteria (18 epitopes, 7 species), parasitic protists (12 epitopes, 2 species), plants (5 epitopes, 2 species), and invertebrates (4 epitopes, 2 species). Seemingly redundant 1D-mimics from the same protein may represent different epitopes and, thus, all 789 1D-mimics were kept at this step.

Structural representatives from the Protein Data Bank (PDB) were identified for 284 1D-mimics based on their source sequence using the minimum cutoffs of 90% for sequence identity and 20% for query cover. The 284 1D-mimics are represented by 7992 redundant structures from 1514 unique PDB chains. From these, structure-based molecular mimics

(termed as “3D-mimics”) were identified. Three-dimensional-mimics are 1D-mimics that share structural similarity with SARS-CoV-2 Spike as determined by an RMSD of at most 1 Å. We found 20 3D-mimics for Spike. Unsurprisingly, as with the 1D-mimics, most 3D-mimics were found for human proteins. Additionally, one 3D-mimic was found for *Mus musculus* (mouse), *Mycobacterium tuberculosis*, *Phleum pratense* (Timothy grass), and respiratory syncytial virus, respectively (Table 1). For each 3D-mimic, Epitepedia computes a Z-score based on the distribution of RMSD values for all resulting hits for the input structure. This allows for a comparative assessment of the quality of a hit, with respect to RMSD, to other hits for a given run. Epitepedia also computes an EpiScore for each hit. EpiScore, calculated as (*motif length/RMSD*), favors longer motifs over shorter ones with the same RMSD values.

**Table 1.** 3D-mimics found for SARS-CoV-2 Spike.

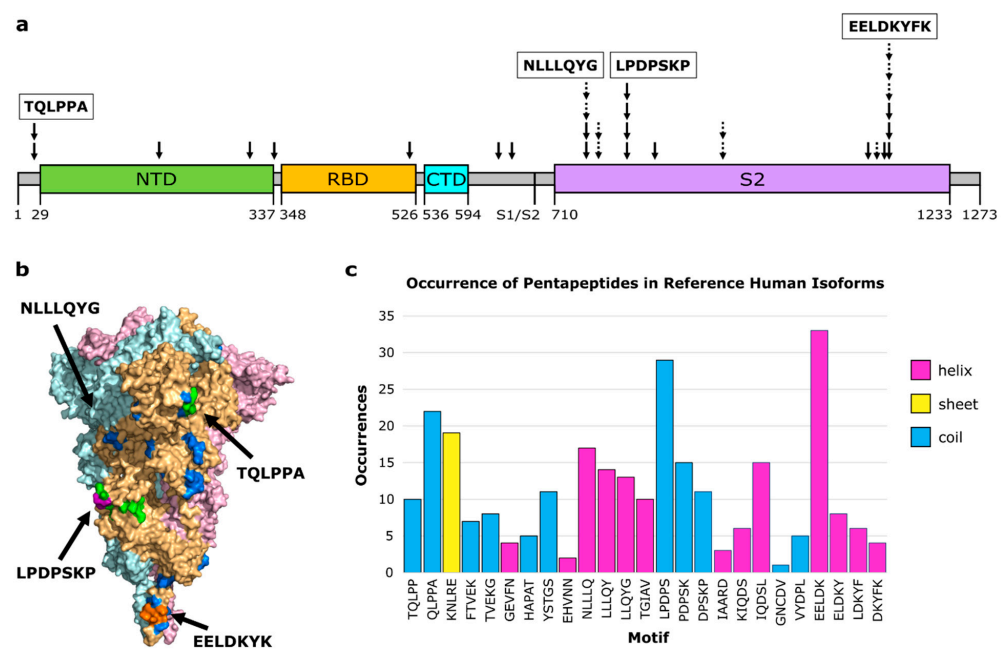
Motif	Protein	Species	RMSD (Å)	Z-Score	EpiScore	PDB_Chain
TQLPP	Thrombopoietin	Human	0.46	−1.34	10.87	1V7N_X
QLPPA	SMYD3 protein	Human	0.38	−1.42	13.16	5CCL_A
KNLRE	Toll-like receptor 8	Human	0.87	−0.92	5.75	6WML_D
FTVEKG	Pollen allergen Phl p2	<i>Phleum pratense</i>	0.76	−1.03	7.89	1WHP_A
GEVFN	Integrin beta 1	Human	0.63	−1.16	7.94	7NWL_B
HAPAT	Activator of 90 kDa heat shock protein ATPase homolog 1	Human	0.74	−1.05	6.76	7DME_A
YSTGS	Argininosuccinate lyase	Human	0.48	−1.31	10.42	1K62_B
EHVNN	Casein kinase 2 alpha isoform	Human	0.29	−1.51	17.24	2ZJW_A
NLLLQ	DNA polymerase subunit gamma 1	Human	0.57	−1.22	8.77	5C51_A
LLQYG	Ankyrin 1	Human	0.20	−1.60	25.00	1N11_A
LPDPS	BRCA1-A complex subunit BRE	Human	0.32	−1.48	15.62	6GVW_C
LPDPS	Semaphorin 7a	Human	0.84	−0.91	5.95	3NVQ_A
DPSKP	60S ribosomal protein L3	Human	0.10	−1.70	50.00	6LU8_B
DPSKP	Alanine and proline-rich secreted protein apa precursor	<i>Mycobacterium tuberculosis</i>	0.21	−1.59	23.81	5ZXA_A
IAARD	Talin	<i>Mus musculus</i>	0.74	−1.05	6.76	6R9T_A
GNCDV	Tryptophan-tRNA ligase	Human	0.91	−0.88	5.49	1O5T_A
SFKEE	Small subunit processome component 20 homolog	Human	0.32	−1.48	15.62	7MQA_SP
EELDK	Kynureninase	Human	0.22	−1.58	22.73	2HZP_A
ELDKY	Fusion glycoprotein F0	Respiratory syncytial virus	0.12	−1.68	41.67	6EAE_F
DKYFK	Cytoplasmic FMR1-interacting protein 1	Human	0.14	−1.66	35.71	4N78_A

For the 402 human 1D-mimics where no PDB structural representative could be identified for their source sequence, AlphaFold2 3D models were used. Three-dimensional model representatives were found for 102 human 1D-mimics. Of these, 10 are predicted to be AlphaFold2-3D-mimics (termed as “AF-3D-mimics”) based on the RMSD (Table 2).

**Table 2.** Human AF-3D-mimics for SARS-CoV-2 Spike.

Motif	Protein	RMSD (Å)	Z-Score	EpiScore	AlphaFold2 ID
NLLLQ	Ankyrin 3	0.61	−1.18	8.20	AF-Q12955-F1-model_v1_A
LLQYG	Olfactory receptor 10Q1	0.66	−1.13	7.58	AF-Q8NGQ4-F1-model_v1_A
TGIAV	Phosphofructokinase	0.17	−1.63	29.41	AF-P17858-F1-model_v1_A
TGIAV	Low affinity immunoglobulin gamma Fc region receptor II-b	0.17	−1.63	29.41	AF-P31995-F1-model_v1_A
KIQDSL	Phosphorylase b kinase regulatory subunit beta	0.19	−1.61	31.58	AF-Q93100-F1-model_v1_A
KIQDSL	Long-chain-fatty-acid-CoA ligase 5	0.37	−1.43	16.22	AF-Q9ULC5-F1-model_v1_A
VYDPL	Actin-binding protein IPP	0.17	−1.63	29.41	AF-Q9Y573-F1-model_v1_A
EELDK	Tight junction-associated protein 1	0.20	−1.60	25.00	AF-Q5JTD0-F1-model_v1_A
EELDKY	Keratin, type I cytoskeletal 18	0.22	−1.58	27.27	AF-P05783-F1-model_v1_A
ELDKY	Tropomyosin alpha-3 chain	0.18	−1.62	27.78	AF-P06753-F1-model_v1_A

The 3D- and AF-3D-mimics (hereinafter referred to as “molecular mimics”) mapped to a few clusters on Spike. Ten molecular mimics were singletons, six overlapping molecular mimics were found as pairs in three small clusters, and the remaining 14 were found in three larger clusters with at least four overlapping molecular mimics (Figure 1a). The largest cluster, with six molecular mimics, was also adjacent to three additional molecular mimics. All mimics are displayed on the surface of Spike’s functional trimer, but the large cluster centered around LLLQY is in a deep pocket and is an unlikely antibody binding epitope in this conformation (Figure 1b). Only one molecular mimic is predicted for the RBD, despite RBD being an immunodominant region in Spike to which many antibodies naturally bind [47]. This molecular mimic (HAPAT) corresponds to the activator of 90-kDa heat shock protein ATPase homolog 1 (AHA1). Two molecular mimics are predicted near the S1/S2 boundary that is a site for proteolytic cleavage [48]. The first is YSTGS from argininosuccinate lyase. The second is EHVNN from casein kinase 2 alpha (CK2). CK2 has been found to play an important role in SARS-CoV-2 infection [49]. Activation of CK2 is promoted by SARS-CoV-2 infection [50] and inhibiting CK2 has been suggested as a therapeutic strategy against both SARS-CoV and SARS-CoV-2 [49]. If a cross-reactive antibody intended for SARS-CoV-2 can interact with CK2, it may impact its activity and perhaps the antibody can stabilize conformations that makes CK2 more active, but these are speculations and more work along these lines is needed.

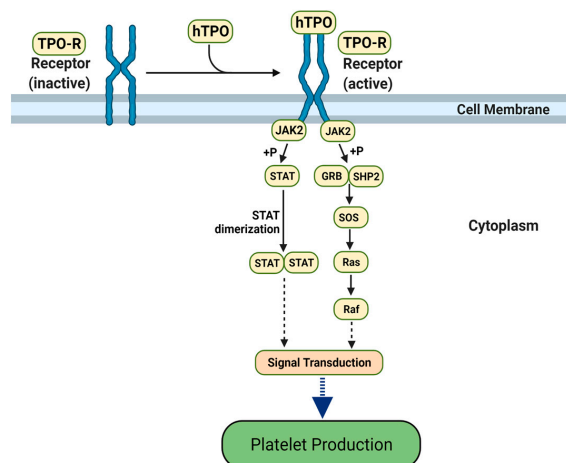


**Figure 1.** Molecular mimicry with autoimmune potential across SARS-CoV-2 Spike. (a) Overview of molecular mimics (solid arrow: 3D-mimic, dashed arrow: AF-3D-mimic) for Spike in the linear sequence showing Spike domains (NTD: N-terminus domain of S1 subunit (green), RBD: receptor binding domain of S1 subunit (orange), CTD: C-terminus domain of S1 subunit (cyan), S2: S2 domain (purple)) as predicted by Pfam [51] based on the NCBI reference sequence (YP:009724390.1). The boundary between the S1 and S2 subunits is indicated at S1/S2. (b) Surface representation of Spike (PDB id: 6XR8 [23]) colored by subunit (pink, beige, light blue) with residues colored by number of occurrences in a molecular mimic (blue: 1, green: 2, purple: 3, orange: 4 or more). Structural visualization generated with PyMOL 2.5.0 [30]. (c) The number of occurrences of the sequence motif in human RefSeq Select isoforms arranged in order from the N-terminus to the C-terminus and colored by predominant secondary structure element (magenta:  $\alpha$ -helix, yellow:  $\beta$ -sheet, blue: coil) based on Spike PDB id 6XR8 chain A.

To further evaluate the autoimmune potential of the human mimics, we identified all occurrences of the motifs in the human RefSeq Select proteome [44]. The pentapeptides

from the molecular mimicry regions are found from four to 33 times in human proteins (Figure 1c, Table S7). The human protein thrombopoietin that includes the 3D-mimic TQLPP (Table 1) also has an occurrence of the sequence mimic LPDPS (Table S7). Further, another protein family that occurs twice for the same pentapeptide is tropomyosin. Tropomyosin alpha-3 is an AF-3D-mimic (Table 2), and tropomyosin alpha-1 has one occurrence of the same pentapeptide (ELDKY). The same motif, ELDKY, is a 3D-mimic in the fusion F0 glycoprotein of respiratory syncytial virus (Table 1). Altogether, based on the known epitopes in IEDB, heterologous immunity is rare with Spike while regions of autoimmune potential form hotspots.

To further evaluate molecular mimicry and, indirectly, autoimmune potential, we performed a deeper investigation of two motifs, TQLPP and ELDKY, that mapped to positions 22–26 (small cluster) and 1151–1155 (largest cluster) in Spike, respectively. For TQLPP, a 3D-mimic with human thrombopoietin was identified. The only structure in our dataset where a 3D-mimic was located at an antibody interface was for human thrombopoietin (hTPO). Thrombopoietin is a cytokine that regulates platelet production [52] (Figure 2). Interestingly, COVID-19 patients often suffer from thrombocytopenia, characterized by low platelet levels [53], which correlates with an almost 5-fold increase in mortality [54]. Thrombocytopenia in COVID-19 patients resembles immune thrombocytopenia (ITP), where hTPO and/or its receptor are mistakenly targeted by autoantibodies leading to reduced platelet count [55]. Treatments with hTPO Receptor Agonists improve thrombocytopenia in both general [56] and COVID-19 [57] patients, suggesting the mistaken targeting occurs before hTPO activates the hTPO receptor. ITP is a heterogeneous disease caused by numerous mechanisms. In ITP patients, about half have antibodies against the major platelet glycoproteins while 28.1% have autoantibodies against hTPO, 21.8% against the hTPO receptor, and 6.8% against the hTPO-hTPO receptor complex. While autoantibodies often seem to play a role in ITP, other mechanisms are possible [55]. It has been suggested that autoimmunity is a likely mechanism for ITP in COVID-19 patients [58]. For ELDKY, we identified one 3D-mimic in the fusion F0 glycoprotein of respiratory syncytial virus (Table 1) and two AF-3D-mimics from keratin type I cytoskeletal 18 and tropomyosin alpha-3 (Table 2). Additional 3D-mimics partially overlapping with ELDKY were identified. The ELDKY motif in Spike is part of a highly reactive epitope [59] found in an  $\alpha$ -helix located towards the C-terminus. This motif is conserved across beta-coronaviruses and can bind an S2P6 antibody effective against all human-infecting beta-coronaviruses [29]. Altogether, the numerous molecular mimics of the ELDKY motif suggest a potential for both protective and autoimmune cross-reactivity.

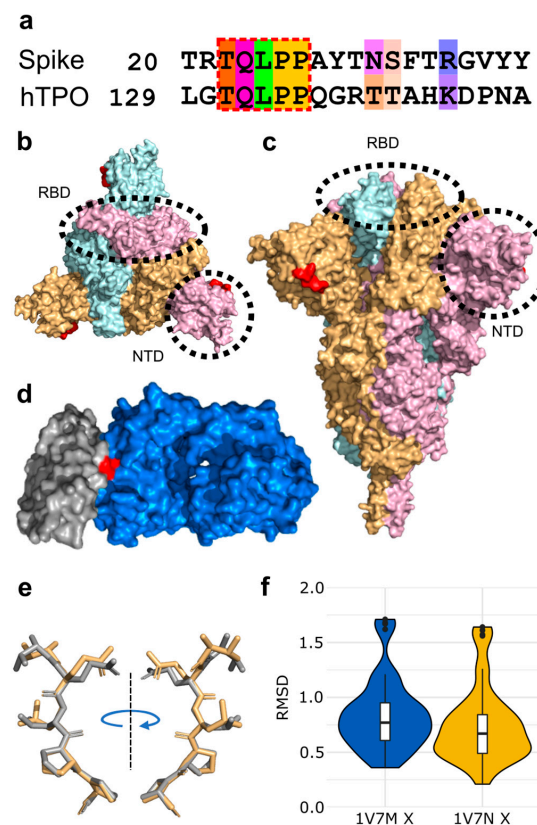


**Figure 2.** The hTPO pathway to induce platelet production. Simplified JAK-STAT signaling pathway in megakaryocytes where hTPO activates the TPO receptor and triggers signaling cascades that stimulate platelet production [60,61]. Created with BioRender.com (accessed on 12 August 2021).



### 3.1. Molecular Mimicry between Spike and Thrombopoietin Mediated through TQLPP

The shared five-amino acid motif, TQLPP (Figure 3a), is located on the surface of Spike's N-terminal Domain (NTD) (Figure 3b,c), whereas it is found at the interface with a neutralizing antibody in hTPO [62] (Figure 3d). The TQLPP motifs from the two proteins are found in coil conformations but exhibit high structural similarity (Figure 3e,f). On Spike, the motif is adjacent to the NTD supersite that is known to be targeted by neutralizing antibodies [63]. We hypothesized that COVID-19 may trigger the production of TQLPP-specific antibodies against this epitope that can cross-react with hTPO. In the absence of Spike TQLPP antibodies, we used molecular modeling and machine learning to investigate the binding of the neutralizing mouse Fab antibody (TN1) from the hTPO structure [64] to the Spike TQLPP epitope.



**Figure 3.** Structural mimicry between a TQLPP motif in SARS-CoV-2 Spike and an antibody binding epitope in thrombopoietin. (a) Pairwise sequence alignment for the TQLPP motif in the epitope for human thrombopoietin (hTPO, IEDB Epitope ID: 920946) and Spike, amino acids colored by Taylor [65] for sites with  $\geq 50\%$  conservation in the amino acid property [66]. The region of molecular mimicry is highlighted in the red dashed box. Surface representation of Spike from (b) the top and (c) the side, with Spike trimer (PDB id: 6XR8 [23]) colored by subunit (pink, beige, light blue) and red indicating the location of the TQLPP epitope fragment, illustrating the surface accessibility of TQLPP and highlighting the location of RBD (dashed oval) and NTD (dashed circle). (d) Surface representation shown for hTPO (gray, PDB id: 1V7M [62]) and its TN1 antibody (blue) with the TQLPP motif (red) at the interface. (e) TM-align generated structural alignment for TQLPP in Spike (beige) and hTPO (gray), with RMSD = 0.61 Å. (f) Violin plots of RMSD values resulting from the comparison of the TQLPP region in 20 Spike trimer structures (60 chains) vs. TQLPP in two hTPO structures (PDB ids: 1V7M and 1V7N, chain X for both [62]). Statistical analysis with Mann-Whitney U reveals no statistical significance between the sets. Box plots, bounded by the 1st and 3rd quartiles, show median value (horizontal solid bold line), vertical lines (whiskers) represent  $1.5 \times \text{IQR}$ , while outliers are marked as black points. For further details, see methods. Alignment representations were generated with Jalview 2.11.2.2 [66] and structural visualizations were generated with PyMOL 2.5.0 [30].

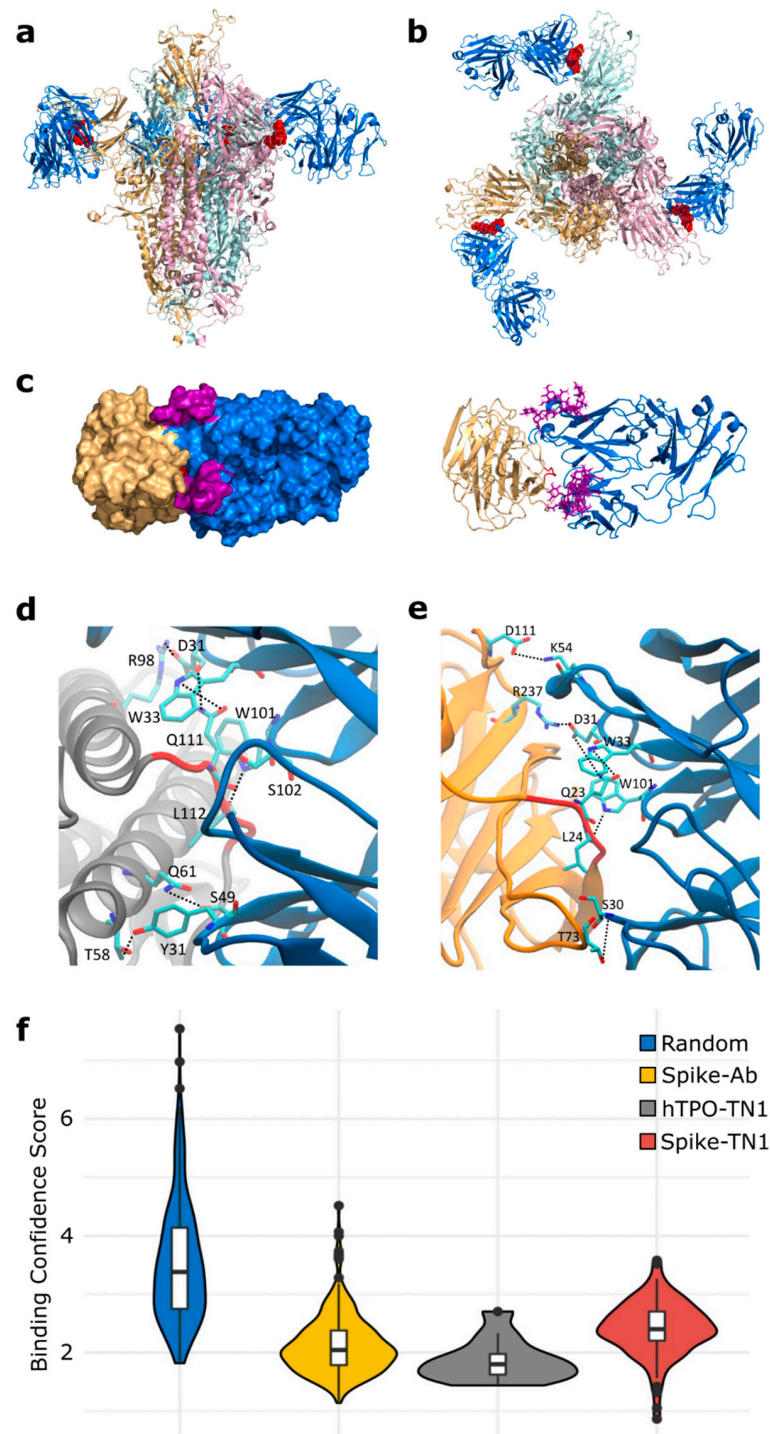
To construct a composite model of Spike and TN1 Fab, a full-length glycosylated model of the Spike trimer, based on PDB id 6VSB [27] with the first 26 residues (including the TQLPP motif) reconstructed [67], was coupled to three copies of TN1 Fab from the structure of hTPO complexed with TN1 Fab [62]. The Spike-TN1 complex was energy minimized and equilibrated with molecular dynamics (MD) simulation. The final model of the Spike trimer complexed with three TN1 Fab antibodies (Figure 4a,b) shows that the TQLPP epitope is accessible to the antibody and the adjacent glycan chains do not shield the antibody-binding site (Figures 4c and S4). To confirm the conformation of TQLPP, we calculated the RMSD for TQLPP regions from 20 Spike trimer structures (60 chains) from PDB, plus the modeled states (before and after equilibration, and upon 200 ns MD simulation) in an all-vs-all manner (Figure S1, Table S2), paying particular attention to the orientation (up or down) of the RBD. For 1953 pairwise comparisons, 1306 had an RMSD  $\leq 1$  Å and 32 had an RMSD  $\geq 2$  Å. Three groups were compared using a Mann-Whitney U test based on RBD state: (1) both down (N = 666, mean = 0.78 Å, median = 0.66 Å), (2) 1 down 1 up (N = 962, mean = 0.81 Å, median = 0.73 Å), and (3) both up (N = 325, mean = 0.85 Å, median = 0.78 Å). Here, comparisons between groups 1 and 2 ( $p$ -value = 0.030) and 1 and 3 ( $p$ -value = 0.003) were significantly different, while that between groups 2 and 3 ( $p$ -value = 0.055) was not (Figure S2). The reconstructed TQLPP region falls within the conformational ensemble from PDB, suggesting that the modeled representation of TQLPP is valid. Furthermore, the Spike-TN1 complexes (with TN1 from PDB ids 1V7M and 1V7N) and hTPO-TN1 complexes (PDB ids 1V7M and 1V7N) are all stable and have comparable binding affinities, with averages ranging from  $-9.2$  to  $-9.56$  kcal/mol (Table S4). The predominant intermolecular contacts for these four complexes are between polar-apolar and apolar-apolar residues (Table S4).

To evaluate the molecular mimicry between the antibody interface areas, we performed MD simulations of hTPO and Spike NTD with TQLPP complexed with the TN1 antibody. The hydrogen bonds were calculated between the TN1 antibody with hTPO and Spike, respectively, from the last 50 ns of both trajectories (Figure S3). Both the Spike-TN1 and the hTPO-TN1 complexes showed similar contact areas (Figure S3). Notably, critical hydrogen bonds were observed for residues Q and L in the TQLPP motif with TN1 for both Spike and hTPO (Figures 4d,e and S3).

To further support our findings, we evaluated the antibody-antigen interface complementarity with MaSIF-search, a recent tool that uses deep learning techniques [40], on a pair of circular surface regions (patches) from an antibody-antigen complex. MaSIF-search produces a score associated with the confidence of binding when forming a stable complex. We refer to this score here as the binding confidence score, where lower scores indicate a higher probability of protein-protein binding. The results show that Spike-TN1 complexes have a better (lower) binding confidence score than random complexes and that complexes including Spike from PDB id 7LQV [63] have three of the four best binding confidence scores (0.86, 1.05, 1.42) and may bind to TN1 as well as, or better than, hTPO (Figure 4f, Tables S5 and S6). Notably, in 7LQV, COVID-19 antibodies bind to Spike at the NTD supersite [63]. These results strongly argue for the possibility of cross-reactivity between Spike and hTPO driven by the molecular mimicry of TQLPP (Figure 4).

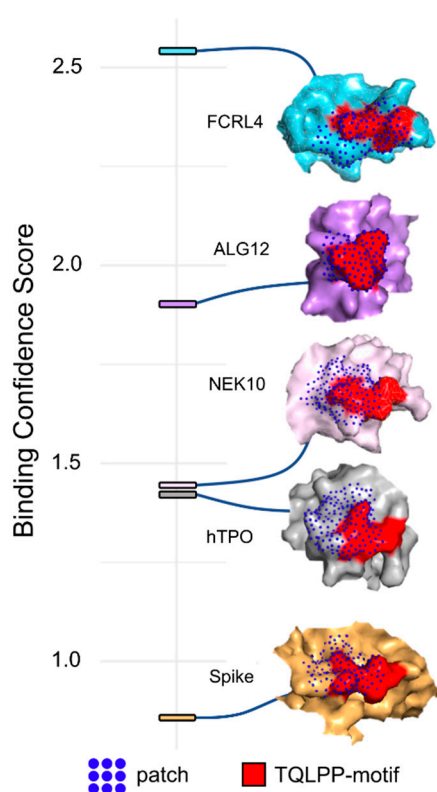
The human proteome contains nine additional occurrences of the TQLPP motif. Two of these motifs, found in Hermansky-Pudlak syndrome 4 protein and ALG12 (Mannosyltransferase ALG12 homolog), have been associated with thrombosis and hemostasis disorder [68]. To evaluate structural mimicry between Spike-TQLPP and all human-TQLPP motifs, we utilized AlphaFold2 3D models [25,45] (Figure S5). The closest structural mimicry region is in hTPO (RMSD = 0.39 Å), followed by coiled-coil domain-containing protein 185, Fc receptor-like protein 4 (FCRL4), and far upstream element-binding protein 1 (Figure S5). These results indicate that TQLPP motifs have similar conformations (Figure S1), strengthening the notion of structural mimicry. We investigated the potential cross-reactivity of an antibody targeting TQLPP in these proteins, after discarding six that display the TQLPP motif in low confidence or unstructured regions. The remaining three

proteins, NEK10 (ciliated cell-specific kinase), FCRL4, and ALG12 were complexed with TN1 (Figure 5). The binding confidence score for NEK10-TN1 (1.44) is comparable to the hTPO-TN1 complex (Figure 5). NEK10 regulates motile ciliary function responsible for expelling pathogens from the respiratory tract [69]. Dysfunction of NEK10 can impact mucociliary clearance and lead to respiratory disorders such as bronchiectasis [69]. Based on our results, it is plausible that the function of NEK10 and thus mucociliary clearance can be affected by cross-reactive Spike antibodies targeting TQLPP.



**Figure 4.** Binding of SARS-CoV-2 Spike to TN1 Fab antibody. Equilibrated structure (1 ns) of the modeled TN1 Fab antibody (blue, PDB id: 1V7M) complexed with Spike trimer model (pink beige,

light blue) shown from (a) the side and (b) the top, with TQLPP shown as red spheres. (c) The Spike NTD (beige) and TN1 Fab complex used for MD simulations (200 ns), with adjacent glycans at N17 and N74 highlighted in purple. The representative amino acids contributing to hydrogen bonds (dashed lines) during the last 50 ns of simulations for the (d) hTPO-TN1 and (e) Spike-TN1 complexes are highlighted as cyan sticks. (f) Violin plot showing the distribution of the MaSIF binding confidence scores for randomly selected patch pairs (blue), the interacting region of Spike-antibody (yellow) and hTPO-TN1 (gray) complexes, and for modeled Spike-TN1 complexes across 40 Spike configurations (red). Statistical analysis with Mann-Whitney U shows that all pairwise comparisons except for Spike-Ab and hTPO-TN1 are significantly different after Bonferroni correction (Table S6). Box plots, bounded by the 1st and 3rd quartiles, show median value (horizontal solid bold line), vertical lines (whiskers) represent  $1.5 \times \text{IQR}$ , while outliers are marked as black points. For further details, see methods. Structural visualizations were generated with PyMOL 2.5.0 [30] and VMD 1.9.3 [31].



**Figure 5.** Predicted interaction patches between TN1 Fab antibody (PDB id: 1V7N) and the TQLPP motif. The best (lowest) binding confidence score is shown for Spike (PDB id: 7LQV, chain A, beige), hTPO (PDB id: 1V7N, chain X, gray), NEK10 (Uniprot: Q6ZWH5, pink), ALG12 (Uniprot: Q9BV10, purple), and FCRL4 (Uniprot: Q96PJ5, light blue). For all, red indicates the TQLPP motif and dark blue dots represent the surface points included in the predicted MaSIF patches.

### 3.2. Molecular Mimicry between Spike, RSV, and Many Human Proteins Mediated through ELDKY

Another motif, ELDKY, is in a region with several partially overlapping pentamer motifs including three 3D-mimics and three AF-3D-mimics (Figure 6a). For the 3D-mimics, two are from the human proteins kynureninase (hKYNU; motif: EELDK) and cytoplasmic FMR1-interacting protein 1 (hCYFIP1; motif: DKYFK), while the last is found in the fusion F0 glycoprotein of respiratory syncytial virus (RSV; motif: ELDKY). For the AF-3D-mimics, the motif is found in human tight junction-associated protein 1 (hTJAP1; motif: EELDK), keratin type I cytoskeletal 18 (hkRT18; motif: EELDKY), and tropomyosin alpha-3 (hTPM3; motif: ELDKY). In Spike, the ELDKY motif is in a stem helix region near the C-terminus. This motif is well-conserved across beta-coronaviruses and is found

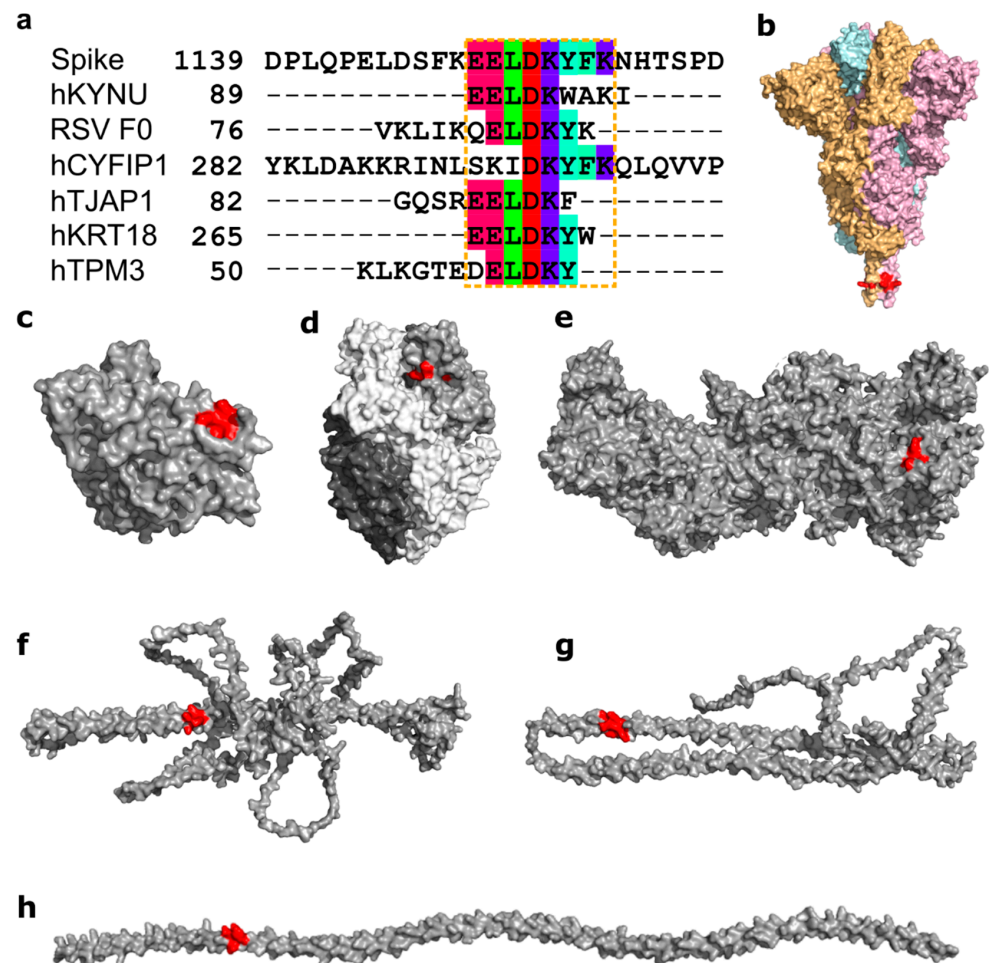


in a highly reactive epitope [59] that has been shown to bind to a broadly neutralizing antibody (S2P6) effective against all human-infecting beta-coronaviruses [29]. The S2P6 antibody (from PDB id 7RNJ [29]) forms a stable complex with the Spike helix, with an average binding affinity of  $-9.52 \pm 0.26$  kcal/mol (Table S4). Here, the predominant intermolecular contacts are formed between charged-apolar, polar-apolar, and apolar-apolar residues (Table S4). In COVID-19, stronger antibody responses to the epitope containing the ELDKY motif have been recorded for severe (requiring hospitalization) vs. moderate cases, while fatal cases had a weaker response than surviving cases [16]. A synthetic epitope containing the ELDKY motif has also been shown to elicit antibody production following COVID-19 immunization [70]. Together with the 3D mimics identified here, these results suggest interesting possibilities for the ELDKY motif from the perspective of both protective immunity and an autoimmune response. First, while not an example of molecular mimicry but evolutionary conservation across beta-coronaviruses, prior exposure to an endemic cold-causing coronavirus (ex. HCoV-OC43) could result in the production of a broadly neutralizing antibody against an epitope containing the ELDKY motif that would be effective against SARS-CoV-2 infection, which could result in milder or asymptomatic infection. Further, a protective effect due to molecular mimicry is suggested by the 3D-mimic identified for the fusion F0 glycoprotein of RSV, a common virus that infects most children in the United States by the time they are 2 years old [71], where antibodies against the ELDKY-containing epitope in RSV may be effective in combatting SARS-CoV-2 infection. In contrast, the potential for an autoimmune response against this motif is suggested by its presence in both two human 3D- and AF-3D-mimics (Figure 6).

There are six additional occurrences of the ELDKY motif in the human proteome (Figure S6). Structural similarity between Spike-ELDKY and human-ELDKY was assessed based on experimentally determined structures (if available) or AlphaFold2 3D models. RMSDs for the ELDKY motif ranged from 0.12–0.20 Å for 5 of the structures, with one hit being an outlier at an RMSD = 0.46 Å. In all instances, the ELDKY motif is found in an  $\alpha$ -helix, resulting in the high degree of structural similarity found for this motif across proteins and bolstering the possibility for molecular mimicry. The ELDKY occurrence with the largest RMSD (0.46 Å) is found in the leucine-zipper dimerization domain of cGMP-dependent protein kinase 1 (PRKG1) (Figure S6) whose phosphorylation targets have roles in the regulation of platelet activation and adhesion [72], smooth muscle contraction [73], and cardiac function [74]. Additionally, PRKG1 regulates intracellular calcium levels via a multitude of signaling pathways [75]. The ELDKY motif is also found in tropomyosin alpha-1 (TPM1), a homolog of the AF-3D-mimic tropomyosin alpha-3 (TPM3). Tropomyosins (TPMs) are involved in the regulation of the calcium-dependent contraction of striated muscle [76]. TPM1 is a 1D-mimic but due to a discrepancy in IEDB it was not identified as a 3D-mimic, although there is a high structural similarity between ELDKY in Spike and ELDKY in TPM1 (Figure S6). A previous study identified a longer match with 53% sequence identity between Spike and TPM1 that included the ELDKY motif [77]. However, in a separate search for structural similarity, Marrama and colleagues were unable to identify structural mimicry at the ELDKY motif due to using a structure for Spike lacking the motif, leading to a conclusion against molecular mimicry contributing to myocarditis in COVID-19 [77], in contrast to our work. These results illustrate the importance of structural representative selection when performing structural comparisons and in taking both sequence and structural similarity together into account when performing molecular mimicry searches, as we have done. For PRKG1, cross-reactive Spike antibodies targeting ELDKY may react with the motif, affecting PRKG1's role in the regulation of platelet activation and adhesion and thus providing another avenue for thrombocytopenia or other blood clotting disorders. Antibodies that cross-react with PRKG1 may also alter calcium levels, thus affecting TPM function. For TPM1, cross-reactive Spike antibodies targeting the ELDKY motif may result in coronary heart disease, as low-level autoantibodies against this protein have been associated with increased risk for this condition [78] and TPM1 and TPM3 are cardiac disease-linked antigens [77]. Cardiac disease, including myocardial



injury and arrhythmia, can be induced by SARS-CoV-2 infection [79] and myocarditis has been found to develop in some individuals following vaccination against SARS-CoV-2 [80]. Furthermore, COVID-19 has been found to increase the risk and long-term burden of several cardiovascular diseases, with COVID-19 severity being proportionate to increased risk and incidence [81].



**Figure 6.** Structural mimicry between an ELDKY motif in SARS-CoV-2 Spike and epitopes in six other proteins. (a) Sequence alignment between SARS-CoV-2 Spike and the epitopes containing the 3D-mimicry motif for human kynureninase (hKYNU, IEDB Epitope ID: 1007556), respiratory syncytial virus fusion F0 glycoprotein (RSV F0, IEDB Epitope ID: 1087776), human cytoplasmic FMR1-interacting protein 1 (hCYFIP1, IEDB Epitope ID: 1346528), human tight junction-associated protein 1 (hTJAP1, IEDB Epitope ID: 1016424), human keratin type I cytoskeletal 18 (hKRT18, IEDB Epitope ID: 1331545), and human tropomyosin alpha-3 (hTPM3, IEDB Epitope ID: 938472). Residues in the molecular mimicry motifs are colored by Taylor [65]. The extended molecular mimicry region is highlighted by the orange dashed box. (b) Surface representation of Spike (PDB id: 6XR8) colored by subunit (beige, pink, light blue) with ELDKY motif indicated in red. Surface representation of proteins (gray) with full or partial 3D-mimics of the ELDKY motif (red): (c) hKYNU (PDB id: 2HZP), (d) RSV F0 (PDB id: 6EAE), (e) hCYFIP1 (PDB id: 4N78), (f) hTJAP1 (Uniprot: Q5JTD0), (g) hKRT18 (Uniprot: P05783), (h) hTPM3 (Uniprot: P06753). Alignment representations were generated with Jalview 2.11.2.2 [66] and structural visualizations were generated with PyMOL 2.5.0 [30].

#### 4. Conclusions

We find that molecular mimics with high autoimmune potential are often found in clusters within Spike. Some clusters have several molecular mimics whose motifs are also found multiple times in the human proteome. Molecular mimics located in  $\alpha$ -helices tend

to have high structural similarity as can be expected based on the regular conformation of the helix, but also some molecular mimics in coil regions are remarkably similar. Our results on the TQLPP motif, located in a coil region, suggest a worrisome potential for cross-reactivity due to molecular mimicry between Spike and hTPO involving the TQLPP epitope that may affect platelet production and lead to thrombocytopenia. Further, cross-reactivity with other TQLPP-containing proteins such as NEK10 cannot be dismissed based on our in-silico results, but in-vivo validation is required. The presence of neutralizing antibodies against peptides with TQLPP in COVID-19 patients' convalescent plasma [82], particularly in severe and fatal cases [16] adds credence to our result. It is also interesting to note that antibodies against a TQLPP-containing peptide were found in the serum of pre-pandemic, unexposed individuals [83]. Prior infection with a different human coronavirus cannot explain the cross-reactivity observed in the unexposed group because TQLPP is situated in a region with low amino acid conservation [83]. Rather, this suggests the presence of an antibody for an unknown epitope with an affinity for the TQLPP region in Spike. The COVID-19 vaccines designed to deliver the Spike protein from SARS-CoV-2, like COVID-19 infection itself, can cause thrombocytopenia [53,84–86] and it is plausible that cross-reactivity can titrate the serum concentration of free hTPO. The TQLPP motif is changing in the SARS-CoV-2 variants and evolutionary trends of the motif suggest it may not remain in Spike. RMSD values between wild-type TQLPP and TQLPP in five variants of concern range from 0.21–1.78 Å (Table S3). In the Gamma variant, a P26S mutation has changed TQLPP to TQLPS and two additional mutations are located just before the motif at L18F and T20N in the NTD supersite, while the Delta variant is mutated at T19R [87]. The first Omicron variant (21K or BA.1), however, has no amino acid substitutions near the TQLPP motif, while a closely related Omicron variant (21L or BA.2) contains a 9 nucleotide deletion that results in the loss of 60% of the TQLPP motif (L24-, P25-, P26-) [87]. Neutralizing antibodies targeting the NTD supersite may rapidly lose efficacy against the evolving SARS-CoV-2. While the current COVID-19 vaccines remain safe and efficacious, we postulate that protein engineering of the TQLPP motif and possibly the NTD supersite for future COVID-19 vaccines may reduce the risk for thrombocytopenia and improve long-term vaccine protection against evolving variants.

We illuminated the cross-reactivity mediated through the ELDKY motif between Spike and PRKG1, TPM1, and TPM3. While PRKG1 provides a connection between blood clotting disorders and cardiac complications, it has a larger RMSD than other ELDKY motifs. ELDKY motifs in  $\alpha$ -helices have high similarity and make good candidates for molecular mimicry. We find ELDKY in the homologous proteins TPM1 and TPM3 suggesting a conserved importance for structure and function. In contrast to TQLPP, the ELDKY motif is highly conserved among beta-coronaviruses [29] and there are presently no SARS-CoV-2 variants with mutations in this region [87]. Further, while the existence of a broadly neutralizing antibody against an epitope containing ELDKY [29] illustrates the potential of this motif as a pan-coronavirus vaccine target, the viability may be diminished by the possibility of autoimmune cross-reactivity due to this motif.

We present an extended application of Epitepedia [22] to identify molecular mimicry between Spike and known epitopes. We do not attempt to discover all possible molecular mimicry epitopes for Spike. Epitepedia is only capable of predicting molecular mimicry for linear epitopes with positive assays that have been deposited in IEDB [19] and cannot predict molecular mimicry de novo. By design, Epitepedia does not predict molecular mimicry for conformational epitopes. Epitepedia relies primarily on structures available in PDB [24] when assessing structural similarity between 1D-mimics and the corresponding region on SARS-CoV-2 Spike. This can result in the nonidentification of potentially genuine molecular mimics if they are only present as 1D-mimics but have yet to have their structure experimentally determined. Moreover, the composition of the PDB is biased towards proteins that crystallize well, thus a molecular mimic can additionally go nonidentified if the 1D-mimic is found in an intrinsically disordered protein region. Proteins are dynamic molecules and the structures present in PDB may only represent a fraction of a protein's full

conformational ensemble [88]. Further, IEDB and PDB both have a biased data composition in that more well-studied proteins are likely to be the ones whose functions and structures are published while other proteins are underrepresented. Lastly, it is important to be mindful that Epitepedia output is strictly a prediction and can have false positives. It is therefore of utmost importance to follow up on the results with both literature searches and experimental validation.

We highlight two epitopes of particular interest in our investigation of molecular mimicry in SARS-CoV-2. For one epitope, we find the TQLPP motif and an interacting antibody with which we perform a computational investigation into antibody binding properties of the tentative molecular mimic. The results show that the same antibody may be able to bind TQLPP-containing epitopes in different proteins and that the TQLPP motif tends to be found in similar conformations despite being in a loop or coil. For the other epitope, we find the ELDKY motif with potential for protective immunity and with high structural similarity. High structural similarity can be expected for  $\alpha$ -helical structures, and, if the sequence is identical, molecular mimicry results. Altogether, these are examples of molecular mimicry that may play a role in the autoimmune or cross-reactive potential of antibodies generated by the immune system against SARS-CoV-2 Spike, but it must be noted that these results have not been experimentally verified. Still, computational investigations into the autoimmune potential of pathogens like SARS-CoV-2 are important for therapeutic intervention and when designing vaccines to avoid potential predictable autoimmune interference.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/v14071415/s1>, Figure S1: RMSD value distribution for solved and modeled Spike TQLPP regions; Figure S2: Comparison of RMSD values for TQLPP region from 20 Spike trimer structures based on RBD state; Figure S3: Molecular dynamics simulations overview; Figure S4: SARS-CoV-2 Spike bound to TN1 Fab antibody; Figure S5: TQLPP motif for 10 human proteins modeled by AlphaFold2; Figure S6: Structure of ELDKY motif for 5 human proteins; Table S1: RMSD values resulting from the alignment of the TQLPP region of 1V7M chain X and 1V7N chain X against the TQLPP region of 60 Spike structures; Table S2: RMSD values resulting from the alignment of the TQLPP region from 60 Spike structures and three modeled states, representing a conformational ensemble of TQLPP in Spike, sorted by RMSD; Table S3: RMSD values for SARS-CoV-2 Spike wild-type TQLPP compared to corresponding region in known variants of concern; Table S4: PRODIGY binding affinities for antigen-antibody complexes; Table S5: Distribution of MaSIF binding confidence scores; Table S6: Statistical comparison of MaSIF binding confidence scores for antibody complexes; Table S7: RefSeq Select human isoforms that contain pentapeptides found in the 3D-mimics and AF-3D-mimics for SARS-CoV-2 Spike; Table S8: MASIF binding confidence scores of other human proteins in complex with TN1.

**Author Contributions:** J.S.-L., J.N.-C., C.A.B., G.N., P.C., K.M., T.C., A.M.M., P.B., and V.S. designed the overall method and approach. J.S.-L., G.N., P.C., A.M.M., K.M. supervised the research. J.N.-C., C.A.B., and J.S.-L. identified molecular mimicry. P.B. and P.C. performed modeling. G.N. and V.S. performed MaSIF-search. J.N.-C., C.A.B., J.S.-L., P.C., G.N., P.B. and V.S. analyzed the data. M.S., J.N.-C., C.A.B., J.S.-L., V.S. and P.B. performed visualization. G.N. and J.S.-L. performed project administration. J.S.-L. and J.N.-C. acted as lead authors. C.A.B., G.N., P.C., K.M., T.C., A.M.M., P.B. and V.S. contributed to writing the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Science Foundation under Grant No. 2037374 (J.N.-C., P.B., V.S., K.M., G.N., P.C., A.M.M., J.S.-L.).

**Data Availability Statement:** No experimental data was generated for this study.

**Acknowledgments:** We thank Sathibalan Ponniah (Immune Analytics LLC, Columbia, MD, USA), Charles Dimitroff (Florida International University), Sixto Leal Jr (University of Alabama—Birmingham), and Kevin Bennett (Florida International University) for discussions. The authors would also like to acknowledge the Instructional & Research Computing Center (IRCC) at Florida International

University for providing HPC computing resources that have contributed to the research results reported within this article.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Guan, W.; Ni, Z.; Hu, Y.; Liang, W.; Ou, C.; He, J.; Liu, L.; Shan, H.; Lei, C.; Hui, D.S.C.; et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N. Engl. J. Med.* **2020**, *382*, 1708–1720. [[CrossRef](#)] [[PubMed](#)]
2. Wang, D.; Hu, B.; Hu, C.; Zhu, F.; Liu, X.; Zhang, J.; Wang, B.; Xiang, H.; Cheng, Z.; Xiong, Y.; et al. Clinical Characteristics of 138 Hospitalized Patients with 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *JAMA* **2020**, *323*, 1061–1069. [[CrossRef](#)] [[PubMed](#)]
3. Dawson, P.; Rabold, E.M.; Laws, R.L.; Connors, E.E.; Gharpure, R.; Yin, S.; Buono, S.A.; Dasu, T.; Bhattacharyya, S.; Westergaard, R.P.; et al. Loss of Taste and Smell as Distinguishing Symptoms of Coronavirus Disease. *Clin. Infect. Dis.* **2021**, *72*, 682–685. [[CrossRef](#)]
4. Dong, E.; Du, H.; Gardner, L. An Interactive Web-Based Dashboard to Track COVID-19 in Real Time. *Lancet Infect. Dis.* **2020**, *20*, 533–534. [[CrossRef](#)]
5. Sah, P.; Fitzpatrick, M.C.; Zimmer, C.F.; Abdollahi, E.; Juden-Kelly, L.; Moghadas, S.M.; Singer, B.H.; Galvani, A.P. Asymptomatic SARS-CoV-2 Infection: A Systematic Review and Meta-Analysis. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2109229118. [[CrossRef](#)] [[PubMed](#)]
6. Saviano, A.; Wrensch, F.; Ghany, M.G.; Baumert, T.F. Liver Disease and Coronavirus Disease 2019: From Pathogenesis to Clinical Care. *Hepatology* **2021**, *74*, 1088–1100. [[CrossRef](#)]
7. Han, X.; Ye, Q. Kidney Involvement in COVID-19 and Its Treatments. *J. Med. Virol.* **2021**, *93*, 1387–1395. [[CrossRef](#)]
8. Long, B.; Brady, W.J.; Koyfman, A.; Gottlieb, M. Cardiovascular Complications in COVID-19. *Am. J. Emerg. Med.* **2020**, *38*, 1504–1507. [[CrossRef](#)]
9. Mei, H.; Luo, L.; Hu, Y. Thrombocytopenia and Thrombosis in Hospitalized Patients with COVID-19. *J. Hematol. Oncol.* **2020**, *13*, 161. [[CrossRef](#)]
10. Wei, J.; Matthews, P.C.; Stoesser, N.; Maddox, T.; Lorenzi, L.; Studley, R.; Bell, J.I.; Newton, J.N.; Farrar, J.; Diamond, I.; et al. Anti-Spike Antibody Response to Natural SARS-CoV-2 Infection in the General Population. *Nat. Commun.* **2021**, *12*, 6250. [[CrossRef](#)]
11. Wang, E.Y.; Mao, T.; Klein, J.; Dai, Y.; Huck, J.D.; Jaycox, J.R.; Liu, F.; Zhou, T.; Israelow, B.; Wong, P.; et al. Diverse Functional Autoantibodies in Patients with COVID-19. *Nature* **2021**, *595*, 283–288. [[CrossRef](#)] [[PubMed](#)]
12. Getts, D.R.; Chastain, E.M.; Terry, R.L.; Miller, S.D. Virus Infection, Antiviral Immunity, and Autoimmunity. *Immunol. Rev.* **2013**, *255*, 197–209. [[CrossRef](#)] [[PubMed](#)]
13. Agrawal, B. Heterologous Immunity: Role in Natural and Vaccine-Induced Resistance to Infections. *Front. Immunol.* **2019**, *10*, 2631. [[CrossRef](#)] [[PubMed](#)]
14. Fraley, E.; LeMaster, C.; Banerjee, D.; Khanal, S.; Selvarangan, R.; Bradley, T. Cross-Reactive Antibody Immunity against SARS-CoV-2 in Children and Adults. *Cell. Mol. Immunol.* **2021**, *18*, 1826–1828. [[CrossRef](#)]
15. Shang, J.; Wan, Y.; Luo, C.; Ye, G.; Geng, Q.; Auerbach, A.; Li, F. Cell Entry Mechanisms of SARS-CoV-2. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 11727–11734. [[CrossRef](#)]
16. Voss, C.; Esmail, S.; Liu, X.; Knauer, M.J.; Ackloo, S.; Kaneko, T.; Lowes, L.; Stogios, P.; Seitova, A.; Hutchinson, A.; et al. Epitope-Specific Antibody Responses Differentiate COVID-19 Outcomes and Variants of Concern. *JCI Insight* **2021**, *6*, e148855. [[CrossRef](#)]
17. Segal, Y.; Shoenfeld, Y. Vaccine-Induced Autoimmunity: The Role of Molecular Mimicry and Immune Crossreaction. *Cell. Mol. Immunol.* **2018**, *15*, 586–594. [[CrossRef](#)]
18. Kanduc, D. From Anti-SARS-CoV-2 Immune Responses to COVID-19 via Molecular Mimicry. *Antibodies* **2020**, *9*, 33. [[CrossRef](#)]
19. Vita, R.; Mahajan, S.; Overton, J.A.; Dhanda, S.K.; Martini, S.; Cantrell, J.R.; Wheeler, D.K.; Sette, A.; Peters, B. The Immune Epitope Database (IEDB): 2018 Update. *Nucleic Acids Res.* **2019**, *47*, D339–D343. [[CrossRef](#)]
20. O'donoghue, S.I.; Schafferhans, A.; Sikta, N.; Stolte, C.; Kaur, S.; Ho, B.K.; Anderson, S.; Procter, J.B.; Dallago, C.; Bordin, N.; et al. SARS-CoV-2 Structural Coverage Map Reveals Viral Protein Assembly, Mimicry, and Hijacking Mechanisms. *Mol. Syst. Biol.* **2021**, *17*, e10079. [[CrossRef](#)]
21. Khavinson, V.; Terekhov, A.; Kormilets, D.; Maryanovich, A. Homology between SARS-CoV-2 and Human Proteins. *Sci. Rep.* **2021**, *11*, 17199. [[CrossRef](#)] [[PubMed](#)]
22. Balbin, C.A.; Nunez-Castilla, J.; Stebliankin, V.; Baral, P.; Sobhan, M.; Cickovski, T.; Mondal, A.M.; Narasimhan, G.; Chapagain, P.; Mathee, K.; et al. Epitopedia: Identifying Molecular Mimicry of Known Immune Epitopes. *BioRxiv* **2021**. [[CrossRef](#)]
23. Cai, Y.; Zhang, J.; Xiao, T.; Peng, H.; Sterling, S.M.; Walsh, R.M.; Rawson, S.; Rits-Volloch, S.; Chen, B. Distinct Conformational States of SARS-CoV-2 Spike Protein. *Science* **2020**, *369*, 1586–1592. [[CrossRef](#)]
24. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. [[CrossRef](#)] [[PubMed](#)]



25. Tunyasuvunakool, K.; Adler, J.; Wu, Z.; Green, T.; Zielinski, M.; Židek, A.; Bridgland, A.; Cowie, A.; Meyer, C.; Laydon, A.; et al. Highly Accurate Protein Structure Prediction for the Human Proteome. *Nature* **2021**, *596*, 590–596. [[CrossRef](#)]
26. Zhang, Y.; Skolnick, J. TM-Align: A Protein Structure Alignment Algorithm Based on the TM-Score. *Nucleic Acids Res.* **2005**, *33*, 2302–2309. [[CrossRef](#)]
27. Wrapp, D.; Wang, N.; Corbett, K.S.; Goldsmith, J.A.; Hsieh, C.L.; Abiona, O.; Graham, B.S.; McLellan, J.S. Cryo-EM Structure of the 2019-NCoV Spike in the Prefusion Conformation. *Science* **2020**, *367*, 1255–1260. [[CrossRef](#)]
28. Choi, Y.K.; Cao, Y.; Frank, M.; Woo, H.; Park, S.J.; Yeom, M.S.; Croll, T.I.; Seok, C.; Im, W. Structure, Dynamics, Receptor Binding, and Antibody Binding of the Fully Glycosylated Full-Length SARS-CoV-2 Spike Protein in a Viral Membrane. *J. Chem. Theory Comput.* **2021**, *17*, 2479–2487. [[CrossRef](#)]
29. Pinto, D.; Sauer, M.M.; Czudnochowski, N.; Low, J.S.; Tortorici, M.A.; Housley, M.P.; Noack, J.; Walls, A.C.; Bowen, J.E.; Guarino, B.; et al. Broad Betacoronavirus Neutralization by a Stem Helix-Specific Human Antibody. *Science* **2021**, *373*, 1109–1116. [[CrossRef](#)]
30. Schrödinger. *The PyMOL Molecular Graphics System*; Version 2.5.0; Schrödinger: New York, NY, USA, 2015.
31. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38. [[CrossRef](#)]
32. Jo, S.; Kim, T.; Iyer, V.G.; Im, W. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859–1865. [[CrossRef](#)] [[PubMed](#)]
33. Brooks, B.R.; Brooks, C.L.; Mackerell, A.D.; Nilsson, L.; Petrella, R.J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545–1614. [[CrossRef](#)] [[PubMed](#)]
34. Lee, J.; Cheng, X.; Swails, J.M.; Yeom, M.S.; Eastman, P.K.; Lemkul, J.A.; Wei, S.; Buckner, J.; Jeong, J.C.; Qi, Y.; et al. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **2016**, *12*, 405–413. [[CrossRef](#)]
35. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802. [[CrossRef](#)] [[PubMed](#)]
36. Nosé, S.; Klein, M.L. Constant Pressure Molecular Dynamics for Molecular Systems. *Mol. Phys.* **1983**, *50*, 1055–1076. [[CrossRef](#)]
37. Essmann, U.; Perera, L.; Berkowitz, M.L.; Darden, T.; Lee, H.; Pedersen, L.G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577–8593. [[CrossRef](#)]
38. Ryckaert, J.P.; Ciccotti, G.; Berendsen, H.J.C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341. [[CrossRef](#)]
39. Xue, L.C.; Rodrigues, J.P.; Kastriitis, P.L.; Bonvin, A.M.; Vangone, A. PRODIGY: A Web Server for Predicting the Binding Affinity of Protein-Protein Complexes. *Bioinformatics* **2016**, *32*, 3676–3678. [[CrossRef](#)] [[PubMed](#)]
40. Gainza, P.; Sverrisson, F.; Monti, F.; Rodolà, E.; Boscaini, D.; Bronstein, M.M.; Correia, B.E. Deciphering Interaction Fingerprints from Protein Molecular Surfaces Using Geometric Deep Learning. *Nat. Methods* **2019**, *17*, 184–192. [[CrossRef](#)]
41. Sanner, M.F.; Olson, A.J.; Spehner, J.C. Reduced Surface: An Efficient Way to Compute Molecular Surfaces. *Biopolymers* **1996**, *38*, 305–320. [[CrossRef](#)]
42. Stebliankin, V.; Baral, P.; Balbin, C.; Nunez-Castilla, J.; Sobhan, M.; Cickovski, T.; Mohan Mondal, A.; Siltberg-Liberles, J.; Chapagain, P.; Mathee, K.; et al. EMOmIS: A Pipeline for Epitope-Based Molecular Mimicry Search in Protein Structures with Applications to SARS-CoV-2. *BioRxiv* **2022**. [[CrossRef](#)]
43. Dunbar, J.; Krawczyk, K.; Leem, J.; Baker, T.; Fuchs, A.; Georges, G.; Shi, J.; Deane, C.M. SAbDab: The Structural Antibody Database. *Nucleic Acids Res.* **2014**, *42*, D1140–D1146. [[CrossRef](#)] [[PubMed](#)]
44. NCBI RefSeq Select. Available online: [https://www.ncbi.nlm.nih.gov/refseq/refseq\\_select/](https://www.ncbi.nlm.nih.gov/refseq/refseq_select/) (accessed on 4 August 2021).
45. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Židek, A.; Potapenko, A.; et al. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [[CrossRef](#)] [[PubMed](#)]
46. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods* **2020**, *173*, 261–272. [[CrossRef](#)]
47. Premkumar, L.; Segovia-Chumbez, B.; Jadi, R.; Martinez, D.R.; Raut, R.; Markmann, A.J.; Cornaby, C.; Bartelt, L.; Weiss, S.; Park, Y.; et al. The Receptor-Binding Domain of the Viral Spike Protein Is an Immunodominant and Highly Specific Target of Antibodies in SARS-CoV-2 Patients. *Sci. Immunol.* **2020**, *5*, 8413. [[CrossRef](#)]
48. Takeda, M. Proteolytic Activation of SARS-CoV-2 Spike Protein. *Microbiol. Immunol.* **2022**, *66*, 15–23. [[CrossRef](#)]
49. Borgo, C.; D’Amore, C.; Sarno, S.; Salvi, M.; Ruzzene, M. Protein Kinase CK2: A Potential Therapeutic Target for Diverse Human Diseases. *Signal Transduct. Target. Ther.* **2021**, *6*, 183. [[CrossRef](#)]
50. Bouhaddou, M.; Memon, D.; Meyer, B.; White, K.M.; Rezelj, V.V.; Correa Marrero, M.; Polacco, B.J.; Melnyk, J.E.; Ulferts, S.; Kaake, R.M.; et al. The Global Phosphorylation Landscape of SARS-CoV-2 Infection. *Cell* **2020**, *182*, 685–712.e19. [[CrossRef](#)]
51. Finn, R.D.; Bateman, A.; Clements, J.; Coggill, P.; Eberhardt, R.Y.; Eddy, S.R.; Heger, A.; Hetherington, K.; Holm, L.; Mistry, J.; et al. Pfam: The Protein Families Database. *Nucleic Acids Res.* **2014**, *42*, D222–D230. [[CrossRef](#)]
52. Varghese, L.N.; Defour, J.-P.; Pecquet, C.; Constantinescu, S.N. The Thrombopoietin Receptor: Structural Basis of Traffic and Activation by Ligand, Mutations, Agonists, and Mutated Calreticulin. *Front. Endocrinol.* **2017**, *8*, 59. [[CrossRef](#)]
53. Yang, X.; Yang, Q.; Wang, Y.; Wu, Y.; Xu, J.; Yu, Y.; Shang, Y. Thrombocytopenia and Its Association with Mortality in Patients with COVID-19. *J. Thromb. Haemost.* **2020**, *18*, 1469–1472. [[CrossRef](#)] [[PubMed](#)]



54. Shi, C.; Wang, L.; Ye, J.; Gu, Z.; Wang, S.; Xia, J.; Xie, Y.; Li, Q.; Xu, R.; Lin, N. Predictors of Mortality in Patients with Coronavirus Disease 2019: A Systematic Review and Meta-Analysis. *BMC Infect. Dis.* **2021**, *21*, 663. [CrossRef] [PubMed]
55. Nazy, I.; Kelton, J.G.; Moore, J.C.; Clare, R.; Horsewood, P.; Smith, J.W.; Ivetic, N.; D'Souza, V.; Li, N.; Arnold, D.M. Autoantibodies to Thrombopoietin and the Thrombopoietin Receptor in Patients with Immune Thrombocytopenia. *Br. J. Haematol.* **2018**, *181*, 234–241. [CrossRef] [PubMed]
56. Audia, S.; Bonnotte, B. Emerging Therapies in Immune Thrombocytopenia. *J. Clin. Med.* **2021**, *10*, 1004. [CrossRef] [PubMed]
57. Watts, A.; Raj, K.; Gogia, P.; Gahona, C.C.T.; Porcelli, M. Secondary Immune Thrombocytopenic Purpura Triggered by COVID-19. *Cureus* **2021**, *13*, e14501. [CrossRef]
58. Frankel, A.E.; Wylie, D.; Peters, B.; Marrama, D.; Ahn, C. Bioinformatic Analysis Underpinning the Frequent Occurrence of Immune Thrombocytopenic Purpura in COVID-19 Patients. *Isr. Med. Assoc. J.* **2022**, *24*, 320–326.
59. Mishra, N.; Huang, X.; Joshi, S.; Guo, C.; Ng, J.; Thakkar, R.; Wu, Y.; Dong, X.; Li, Q.; Pinapati, R.S.; et al. Immunoreactive Peptide Maps of SARS-CoV-2. *Commun. Biol.* **2021**, *4*, 225. [CrossRef]
60. Kanehisa, M.; Furumichi, M.; Sato, Y.; Ishiguro-Watanabe, M.; Tanabe, M. KEGG: Integrating Viruses and Cellular Organisms. *Nucleic Acids Res.* **2021**, *49*, D545–D551. [CrossRef]
61. Kuter, D.J. The Biology of Thrombopoietin and Thrombopoietin Receptor Agonists. *Int. J. Hematol.* **2013**, *98*, 10–23. [CrossRef]
62. Feese, M.D.; Tamada, T.; Kato, Y.; Maeda, Y.; Hirose, M.; Matsukura, Y.; Shigematsu, H.; Muto, T.; Matsumoto, A.; Watarai, H.; et al. Structure of the Receptor-Binding Domain of Human Thrombopoietin Determined by Complexation with a Neutralizing Antibody Fragment. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 1816–1821. [CrossRef]
63. Cerutti, G.; Guo, Y.; Zhou, T.; Gorman, J.; Lee, M.; Rapp, M.; Reddem, E.R.; Yu, J.; Bahna, F.; Bimela, J.; et al. Potent SARS-CoV-2 Neutralizing Antibodies Directed against Spike N-Terminal Domain Target a Single Supersite. *Cell Host Microbe* **2021**, *29*, 819–833.e7. [CrossRef] [PubMed]
64. Tahara, T.; Kuwaki, T.; Matsumoto, A.; Morita, H.; Watarai, H.; Inagaki, Y.; Ohashi, H.; Ogami, K.; Miyazaki, H.; Kato, T. Neutralization of Biological Activity and Inhibition of Receptor Binding by Antibodies against Human Thrombopoietin. *Stem Cells* **1998**, *16*, 54–60. [CrossRef]
65. Taylor, W.R. Residual Colours: A Proposal for Aminochromography. *Protein Eng.* **1997**, *10*, 743–746. [CrossRef]
66. Waterhouse, A.M.; Procter, J.B.; Martin, D.M.A.; Clamp, M.; Barton, G.J. Jalview Version 2—A Multiple Sequence Alignment Editor and Analysis Workbench. *Bioinformatics* **2009**, *25*, 1189–1191. [CrossRef]
67. Woo, H.; Park, S.J.; Choi, Y.K.; Park, T.; Tanveer, M.; Cao, Y.; Kern, N.R.; Lee, J.; Yeom, M.S.; Croll, T.I.; et al. Developing a Fully Glycosylated Full-Length SARS-CoV-2 Spike Protein Model in a Viral Membrane. *J. Phys. Chem. B* **2020**, *124*, 7128–7137. [CrossRef] [PubMed]
68. Kanduc, D. Thromboses and Hemostasis Disorders Associated with Coronavirus Disease 2019: The Possible Causal Role of Cross-Reactivity and Immunological Imprinting. *Glob. Med. Genet.* **2021**, *8*, 162–170. [CrossRef] [PubMed]
69. Chivukula, R.R.; Montoro, D.T.; Leung, H.M.; Yang, J.; Shamseldin, H.E.; Taylor, M.S.; Dougherty, G.W.; Zariwala, M.A.; Carson, J.; Daniels, L.A.; et al. A Human Ciliopathy Reveals Essential Functions for NEK10 in Airway Mucociliary Clearance. *Nat. Med.* **2020**, *26*, 244. [CrossRef]
70. Andries, J.; Viranaicken, W.; Cordonin, C.; Herrscher, C.; Planesse, C.; Roquebert, B.; Lagrange-Xelot, M.; El-Kalamouni, C.; Meilhac, O.; Mavingui, P.; et al. The SARS-CoV-2 Spike Residues 616/644 and 1138/1169 Delineate Two Antibody Epitopes in COVID-19 MRNA COMINARTY Vaccine (Pfizer/BioNTech). *Sci. Rep.* **2022**, *12*, 5999. [CrossRef]
71. Respiratory Syncytial Virus (RSV) | NIH: National Institute of Allergy and Infectious Diseases. Available online: <https://www.niaid.nih.gov/diseases-conditions/respiratory-syncytial-virus-rsv> (accessed on 8 January 2022).
72. Li, Z.; Xi, X.; Gu, M.; Feil, R.; Ye, R.D.; Eigenthaler, M.; Hofmann, F.; Du, X. A Stimulatory Role for CGMP-Dependent Protein Kinase in Platelet Activation. *Cell* **2003**, *112*, 77–86. [CrossRef]
73. Sauzeau, V.; Le Jeune, H.; Cario-Toumaniantz, C.; Smolenski, A.; Lohmann, S.M.; Bertoglio, J.; Chardin, P.; Pacaud, P.; Loirand, G. Cyclic GMP-Dependent Protein Kinase Signaling Pathway Inhibits RhoA-Induced Ca<sup>2+</sup> Sensitization of Contraction in Vascular Smooth Muscle. *J. Biol. Chem.* **2000**, *275*, 21722–21729. [CrossRef]
74. Francis, S.H. The Role of CGMP-Dependent Protein Kinase in Controlling Cardiomyocyte CGMP. *Circ. Res.* **2010**, *107*, 1164. [CrossRef] [PubMed]
75. Francis, S.H.; Busch, J.L.; Corbin, J.D. CGMP-Dependent Protein Kinases and CGMP Phosphodiesterases in Nitric Oxide and CGMP Action. *Pharmacol. Rev.* **2010**, *62*, 525. [CrossRef] [PubMed]
76. Szent-Györgyi, A.G. Calcium Regulation of Muscle Contraction. *Biophys. J.* **1975**, *15*, 707–723. [CrossRef]
77. Marrama, D.; Mahita, J.; Sette, A.; Peters, B. Lack of Evidence of Significant Homology of SARS-CoV-2 Spike Sequences to Myocarditis-Associated Antigens. *EBioMedicine* **2022**, *75*, 103807. [CrossRef]
78. Zhang, Y.; Zhao, H.; Liu, B.; Li, L.; Zhang, L.; Bao, M.; Ji, X.; He, X.; Yi, J.; Chen, P.; et al. Low Level Antibodies Against Alpha-Tropomyosin Are Associated With Increased Risk of Coronary Heart Disease. *Front. Pharmacol.* **2020**, *11*, 195. [CrossRef]
79. Nishiga, M.; Wang, D.W.; Han, Y.; Lewis, D.B.; Wu, J.C. COVID-19 and Cardiovascular Disease: From Basic Mechanisms to Clinical Perspectives. *Nat. Rev. Cardiol.* **2020**, *17*, 543–558. [CrossRef]
80. Patone, M.; Mei, X.W.; Handunnetthi, L.; Dixon, S.; Zaccardi, F.; Shankar-Hari, M.; Watkinson, P.; Khunti, K.; Harnden, A.; Coupland, C.A.C.; et al. Risks of Myocarditis, Pericarditis, and Cardiac Arrhythmias Associated with COVID-19 Vaccination or SARS-CoV-2 Infection. *Nat. Med.* **2021**, *28*, 410–422. [CrossRef]

81. Xie, Y.; Xu, E.; Bowe, B.; Al-Aly, Z. Long-Term Cardiovascular Outcomes of COVID-19. *Nat. Med.* **2022**, *28*, 583–590. [[CrossRef](#)]
82. Li, Y.; Lai, D.; Zhang, H.; Jiang, H.; Tian, X.; Ma, M.; Qi, H.; Meng, Q.; Guo, S.; Wu, Y.; et al. Linear Epitopes of SARS-CoV-2 Spike Protein Elicit Neutralizing Antibodies in COVID-19 Patients. *Cell. Mol. Immunol.* **2020**, *17*, 1095–1097. [[CrossRef](#)]
83. Stoddard, C.I.; Galloway, J.; Chu, H.Y.; Shipley, M.M.; Sung, K.; Itell, H.L.; Wolf, C.R.; Logue, J.K.; Magedson, A.; Garrett, M.E.; et al. Epitope Profiling Reveals Binding Signatures of SARS-CoV-2 Immune Response in Natural Infection and Cross-Reactivity with Endemic Human CoVs. *Cell Rep.* **2021**, *35*, 109164. [[CrossRef](#)]
84. Helms, J.M.; Ansteatt, K.T.; Roberts, J.C.; Kamatam, S.; Foong, K.S.; Labayog, J.M.S.; Tarantino, M.D. Severe, Refractory Immune Thrombocytopenia Occurring after SARS-CoV-2 Vaccine. *J. Blood Med.* **2021**, *12*, 221–224. [[CrossRef](#)] [[PubMed](#)]
85. Schultz, N.H.; Sørvoll, I.H.; Michelsen, A.E.; Munthe, L.A.; Lund-Johansen, F.; Ahlen, M.T.; Wiedmann, M.; Aamodt, A.-H.; Skattør, T.H.; Tjønnfjord, G.E.; et al. Thrombosis and Thrombocytopenia after ChAdOx1 NCoV-19 Vaccination. *N. Engl. J. Med.* **2021**, *384*, 2124–2130. [[CrossRef](#)] [[PubMed](#)]
86. Greinacher, A.; Thiele, T.; Warkentin, T.E.; Weisser, K.; Kyrle, P.A.; Eichinger, S. Thrombotic Thrombocytopenia after ChAdOx1 NCoV-19 Vaccination. *N. Engl. J. Med.* **2021**, *384*, 2092–2101. [[CrossRef](#)] [[PubMed](#)]
87. Hodcroft, E.B. CoVariants: SARS-CoV-2 Mutations and Variants of Interest. 2022. Available online: <https://covariants.org/shared-mutations> (accessed on 8 January 2022).
88. Slabinski, L.; Jaroszewski, L.; Rodrigues, A.P.C.; Rychlewski, L.; Wilson, I.A.; Lesley, S.A.; Godzik, A. The Challenge of Protein Structure Determination—Lessons from Structural Genomics. *Protein Sci.* **2007**, *16*, 2472–2482. [[CrossRef](#)] [[PubMed](#)]