Superantigenic character of an insert unique to SARS-CoV-2 spike supported by skewed TCR repertoire in patients with hyperinflammation

Mary Hongying Cheng*, She Zhang*, Rebecca A. Porritt*, Magali Noval Rivas*, Lisa Paschold*, Edith Willscher*, Mascha Binder*, Moshe Arditib,c,1,2, and Ivet Bahar*,1,2

*Department of Computational and Systems Biology, School of Medicine, University of Pittsburgh, Pittsburgh, PA 15213; bDepartment of Pediatrics, Division of Infectious Diseases and Immunology, Cedars-Sinai Medical Center, Los Angeles, CA 90048; cDepartment of Medical Sciences, Infectious and Immunologic Diseases Research Center, Cedars-Sinai Medical Center, Los Angeles, CA 90048; and dDepartment of Internal Medicine IV, Oncology/Hematology, Martin Luther University Halle-Wittenberg, 06120 Halle (Saale), Germany

Contribution by Ivet Bahar, June 29, 2020 (sent for review May 26, 2020; reviewed by Talat A. Chatila, Ruth Nussinov, and Celia A. Schiffer)

Multisystem Inflammatory Syndrome in Children (MIS-C) associated with COVID-19 is a newly recognized condition in children with recent severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection. These children and adult patients with severe hyperinflammation present with a constellation of symptoms that strongly resemble toxic shock syndrome, an escalation of the cytotoxic adaptive immune response triggered upon the binding of pathogenic superantigens to T cell receptors (TCRs) and/or major histocompatibility complex class II (MHCI) molecules. Here, using structure-based computational models, we demonstrate that the SARS-CoV-2 spike (S) glycoprotein exhibits a high-affinity motif for binding TCRs, and may form a ternary complex with MHCI. The binding epitope on S harbors a sequence motif unique to SARS-CoV-2 (not present in other SARS-related coronaviruses), which is highly similar in both sequence and structure to the bacterial superantigen staphylococcal enterotoxin B. This interaction between the virus and human T cells could be strengthened by a rare mutation (D839Y/N/E) from a European strain of SARS-CoV-2. Furthermore, the interfacing region includes selected residues from an intercellular adhesion molecule (ICAM)-like motif shared between the coronaviruses, which is highly similar in both sequence and structure to the bacterial superantigen staphylococcal enterotoxin B. This interaction between the virus and human T cells could be strengthened by a rare mutation (D839Y/N/E) from a European strain of SARS-CoV-2. Furthermore, the interfacing region includes selected residues from an intercellular adhesion molecule (ICAM)-like motif shared between the SARS viruses from the 2003 and 2019 pandemics. A neurotoxin-like sequence motif on the receptor-binding domain also exhibits a high tendency to bind TCRs. Analysis of the TCR repertoire in adult COVID-19 patients demonstrates that those with severe hyperinflammation disease exhibit TCR skewing consistent with superantigen activation. These data suggest that SARS-CoV-2 S may act as a superantigen to trigger the development of MIS-C as well as cytokine storm in adult COVID-19 patients, with important implications for the development of therapeutic approaches.

COVID-19 | superantigen | SARS-CoV-2 spike | toxic shock syndrome | TCR binding

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the virus that causes COVID-19, is a betacoronavirus (β-CoV) closely related to SARS-CoV and Middle East Respiratory Syndrome (MERS)-CoV (1). COVID-19 can manifest in adults as a severe interstitial pneumonia with hyperinflammation, while severe respiratory manifestations are rare in children (2–4). Recently, however, multisystem inflammatory syndrome in children (MIS-C) has been observed in patients that either tested positive for COVID-19 (by PCR or serology) or had epidemiological links to COVID-19 (5–7). After initial reports in the United Kingdom (5), many cases have been reported in Europe (6, 7) and New York (Centers for Disease Control and Prevention). However, no such cases have been reported in China, Japan, or South Korea, which have also been severely impacted by the COVID-19 pandemic (European Centre for Disease Prevention and Control).

MIS-C manifests as persistent fever and hyperinflammation with multiorgan system involvement including cardiac, gastrointestinal (GI), renal, hematologic, dermatologic, and neurologic symptoms (5–7). These symptoms are highly reminiscent of toxic shock syndrome (TSS) (8, 9) (Table 1), rather than Kawasaki disease (KD), due to marked demographic, clinical, and laboratory differences (6). Indeed, a recent uncontrolled retrospective case study concluded that MIS-C is distinct from KD and KD shock syndrome (10). The similarities to TSS and the association of MIS-C with COVID-19 led us to hypothesize that SARS-CoV-2 may possess superantigenic fragments that induce an inflammatory cascade and may also contribute to the hyperinflammation and cytokine storm observed in severe adult COVID-19 patients (3, 4). The question we raised is, does SARS-CoV-2 spike (S) possess superantigenic fragments that could elicit such reactions upon binding proteins involved in the host cell cytotoxic adaptive immune response? Such a reaction was not observed in the SARS-CoV pandemic of 2003 (SARS1 hereafter). What is unique to SARS-CoV-2, and how might recent mutations in SARS-CoV-2 S have promoted such an increased virulence?

TSS can be caused by two types of superantigens (SAgs): bacterial or viral. Bacterial SAgs have been broadly studied. They include proteins secreted by Staphylococcus aureus and Streptococcus pyogenes that stimulate massive production of toxic shock syndrome (TSS) (8, 9) (Table 1), rather than Kawasaki disease (KD), due to marked demographic, clinical, and laboratory differences (6). Indeed, a recent uncontrolled retrospective case study concluded that MIS-C is distinct from KD and KD shock syndrome (10). The similarities to TSS and the association of MIS-C with COVID-19 led us to hypothesize that SARS-CoV-2 may possess superantigenic fragments that induce an inflammatory cascade and may also contribute to the hyperinflammation and cytokine storm observed in severe adult COVID-19 patients (3, 4). The question we raised is, does SARS-CoV-2 spike (S) possess superantigenic fragments that could elicit such reactions upon binding proteins involved in the host cell cytotoxic adaptive immune response? Such a reaction was not observed in the SARS-CoV pandemic of 2003 (SARS1 hereafter). What is unique to SARS-CoV-2, and how might recent mutations in SARS-CoV-2 S have promoted such an increased virulence?

**Significance**

A hyperinflammatory syndrome reminiscent of toxic shock syndrome is observed in severe COVID-19 patients, including children with Multisystem Inflammatory Syndrome in Children (MIS-C). TSS is typically caused by pathogenic superantigens stimulating excessive activation of the adaptive immune system. We show that SARS-CoV-2 spike contains sequence and structure motifs highly similar to those of a bacterial superantigen and may directly bind T cell receptors. We further report a skewed T cell receptor repertoire in COVID-19 patients with severe hyperinflammation, in support of such a superantigenic effect. Notably, the superantigen-like motif is not present in other SARS family coronaviruses, which may explain the unique potential for SARS-CoV-2 to cause both MIS-C and the cytokine storm observed in adult COVID-19.
inflammatory cytokines and toxic shock. Typical examples are TSS toxin 1, and staphylococcal enterotoxins B (SEB) and H (SEH). They are highly potent T cell activators that can bind to major histocompatibility complex (MHC) class II (MHCII) molecules and/or to T cell receptors (TCRs) of both CD4+ and CD8+ T cells. The ability of SAgs to bypass the antigen specificity of the TCRs results in broad activation of T cells and a cytokine storm, leading to toxic shock (11, 12). Notably, SAgs do not bind the major (antigenic) peptide-binding groove of MHCII, but instead bind other regions as well as the αβ TCRs, directly. While early studies showed that bacterial SAgs activate T cells by binding to the β-chain of dimeric TCRs at their variable domain (V) (13-15), more-recent studies revealed that they can bind to either α- or β-chains, or both (16). The question is then, does SARS-CoV-2 S possess any superantigenic fragments/domains that could bind to TCRs?

Here, we used computational modeling to determine whether SARS-CoV-2 S possesses SAg-like fragments and activity. We demonstrate that a polybasic insert present in SARS-CoV-2 S, which is absent in the S glycoprotein of other SARS-related β-CoVs, mediates high-affinity, nonspecific binding to the TCR. Notably, a motif of ~20 amino acids enclosing this insert unique to SARS-CoV-2 S has sequence and structure features highly similar to those of the toxin SEB. Our analysis further indicates that a rare SARS-CoV-2 S mutation detected in a European strain may potentially enhance TCR binding. Moreover, analysis of a cohort of adult COVID-19 patients reveals that those with severe hyperinflammatory disease exhibit skewing of the TCR repertoire consistent with SAg activity. Therefore, these findings have important implications for the management and treatment of both MIS-C and COVID-19 patients with hyperinflammatory syndrome.

Results and Discussion

SARS-CoV-2 Spike Harbors a High-Affinity Site for TCR β-Chain Binding, which Contains An Insertion, P_681RRA_684 Unique to SARS2. We first examined whether SARS-CoV-2 S could bind to a TCR α- and/or β-chain. To this aim, we constructed a structural model for SARS-CoV-2 S based on the cryo-electron microscopy (cryo-EM) structure resolved for the S glycoprotein (17). We used then the X-ray structure of αβ TCR resolved in a ternary complex with SEH and MHCII (16) to generate a series of structural models for possible SARS-CoV-2 S–TCR complex formation using ClusPro (18). Our simulations revealed two most probable TCR-binding sites on each monomer of S: one on the receptor-binding domain (RBD; residues R319 to K529), and the other near the S1/S2 cleavage site between the subunits S1 and S2. The former was also shared by SARS1 and MERSCoV S, while the latter was unique and strongly preferred in SARS-CoV-2 S, as described in detail in SI Appendix, Figs. S1 and S2. Therefore, we focused on the latter, shown in Fig. 1A, as the most probable mechanism of complex formation that distinguishes SARS-CoV-2 S from other β-CoVs. A close-up view of the interface between S and TCRβ (Fig. 1B) reveals strong interatomic interactions involving spike residues S680-R683, and Vβ residues Q52, D56, R70 to E74 (CDR2), and S96 to Q103 (CDR3).

We note that the TCRβ-binding epitope on SARS-CoV-2 S is centered around a polybasic sequence motif, P_681RRA_684 (PRRA hereafter), and includes the sequential and spatial neighbors of this motif. Comparison to other β-CoV S sequences shows that SARS-CoV-2 S is distinguished by the existence of this four-residue insertion, PRRA, preceding the S1/S2 cleavage site (R685–S686 peptide bond) (Fig. 2A). Structural comparison of the S glycoproteins between SARS1 and SARS-CoV-2 further shows their close structural similarity in general [except for their RBDs engaged in specific interfacial interactions (17)], but the two S glycoproteins significantly differ near the PRRA motif unique to SARS-CoV-2, which is exposed to the exterior (Fig. 2B).

Notably, the exposure of this motif and its close sequential neighbors is further accentuated in the S1 trimer (Fig. 2C) after cleavage by the human proteases (TMPRSS2 or furin) to enable the activation of the fusion trimer of S2 subunits.

### Table 1. Similarities between clinical and laboratory features of MIS-C and pediatric TSS

<table>
<thead>
<tr>
<th>Clinical Features</th>
<th>MIS-C*</th>
<th>Pediatric TSS†</th>
</tr>
</thead>
<tbody>
<tr>
<td>High fever</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Skin rash</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Conjunctivitis</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Oral mucosal involvement</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Myalgia</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Hypotension</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Myocardial involvement (dysfunction)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Gastro-intestinal symptoms (vomiting, diarrhea, abdominal pain)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Renal involvement</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>CNS symptoms, altered mental state</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Headache</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>High C-reactive protein (CRP)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>High ferritin</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>High IL-6</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>High D-dimers</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>High procalcitonin</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Lymphopenia</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Reduced platelet count</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Increased neutrophil count</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Increased aspartate amino transferase (AST) and alanine transaminase (ALT)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>High pro b-type natriuretic peptide (pro-BNP)</td>
<td>+</td>
<td>NA</td>
</tr>
<tr>
<td>High troponin</td>
<td>+</td>
<td>NA</td>
</tr>
<tr>
<td>Isolation of TSS inducing bacteria (Staphylococcus or Streptococcus)</td>
<td>–</td>
<td>+</td>
</tr>
</tbody>
</table>

*NA, not available; + represents association with reported cases; –, no association.
*Taken from refs. 5 to 7.
†Taken from refs. 8, 9, 61, and 62.
Further Examination of the Motif near PRRA Reveals Close Structural Similarity to the SEB Superantigen as well as Sequence Similarities to Neurotoxins and a Viral SAg. The insertion PRRA together with seven sequentially preceding residues and succeeding R685 (conserved among β-CoVs) form a motif, Y_{674}^{678}N_{679}^{685}PRRRAR, homologous to those of neurotoxins from Ophiophagus (cobra) and Bungarus genera, as well as the neurotoxin-like regions from three RABV strains (20) (Fig. 2D). We further noticed that the same segment bears close similarity to the HIV-1 glycoprotein gp120 SAg motif F164 to V174. This close sequence similarity to both bacterial and viral SAgs, in support of the potential superantigenic character of the stretch Y674 to R685 of SARS-CoV-2 S, directed us to further analyze its local sequence and structure.

Our analysis led to an interesting sequence similarity between the fragment T678 to Q690 of SARS-CoV-2 S and the SEB superantigenic peptide T_{150}^{161}KKKATQELD (Fig. 3A). This dodecapeptide sequence shows strong conservation among a broad range of staphylococcal and streptococcal SAgs (21, 22). We note that the sequentially aligned segment of SARS1 (S664 to K672) bears minimal similarity to the SEB SAg (Fig. 3A, left). What is even more interesting is that SARS-CoV-2 motif showed a palindromic behavior with respect to this superantigenic SEB sequence, in the sense that a broader stretch, from E661 to R685, could be aligned to the SAg peptide in the reverse direction as well (Fig. 3A, right). This brings to our attention the versatility and high propensity of the SARS-CoV-2 S TCRβ-β-binding site residues to potentially elicit an SAg-like response.

Significantly, the structure of the SARS-CoV-2 S SAg-like segment and that of SEB peptide also exhibit a remarkable similarity (Fig. 3B and C): A salt bridge (E159–K152 in SEB and E661–R685 in SARS-CoV-2 S) stabilizes both structural motifs; the relative orientations of three positively charged residues (K152, K153, and K154 in SEB and R682, R683, and R685 in SARS-CoV-2 S) are maintained; and an asparagine (N151 in SEB, N679 in SARS-CoV-2) completes this motif. All three features are absent in SARS1 S (Fig. 3D). A β-hairpin that apparently serves as a scaffold is conserved in all three spikes, and we observe a pair of cysteines that may potentially form a disulfide bond in SARS-CoV-2 and SARS1 spikes (C662–C671 and C648–C657, respectively).

This analysis overall indicates that the segment T_{150}^{161}KKKATQELD may potentially form a putatively superantigenic core, consistently aligned against various bacterial or viral SAgs (Figs. 2C and 3A–C) with or without the participation of the adjoining amino acids. However, combined broader sequence and structure analysis in Fig. 3A (Right) and B and C, reveals an even more compelling feature: This putative SAg core is structurally consolidated by spatial proximity to a conserved acidic segment, E_{661}^{663}CD_{665}, which forms a highly stable salt bridge with the polybasic segment PRRAR of SARS-CoV-2 S, much in the same way as the salt bridge observed in SEB (but not in SARS1 S), complemented by an asparagine shared between SARS-CoV-2 S and SEB (but not SARS1 S), and the SAg character may be conferred by this type of structural scaffolding.

We note that the SEB superantigenic peptide Y_{150}^{161}KKKATQELD has been reported to bind CD28 (21), a TCR that provides costimulatory signals required for T cell activation and survival. CD28 and TCRVβ domains share the same (immunoglobulin, Ig) fold (Fig. 3E), and the binding mechanism shown in Fig. 1B could adapt, with minor rearrangements, to interactions with other Ig-fold molecules including neutralizing antibodies. Because of the homologous superantigenic segment of SEB binding CD28, we also tested the potential binding of SARS-CoV-2 spike residues E661 to R685 onto CD28. Our simulations indicated that the same segment can equally bind to CD28, further supporting the strong propensity of the fragment to stimulate T cell activation.

An ICAM-1–like Motif Shared between SARS1 and SARS-CoV-2 Spikes Interacts with TCRβ and Further Stabilizes the S-TCR Complex. The existence of potential superantigenic, toxic, or intercellular adhesion molecule (ICAM)-like fragments in SARS1 was thoroughly examined by Li et al. (23) following the 2003 pandemic. This led to the identification of the nine sequence stretches including three Botulinum neurotoxin type D or G precursors and two motifs that are highly similar to ICAM-1. Comparative analysis with SARS-CoV-2 S sequence revealed that seven of these motifs are conserved (with >65% sequence identity) between SARS1 and SARS-CoV-2 spikes (SI Appendix, Fig. S3). Among them, the ICAM-1 (CD54)-like motif Y_{375}^{388}NENGTTIDAVDCALDLPLSETKC also participates in the association between the SARS-CoV-2 S and αβTCR as shown in Fig. 4.

ICAM-1 involvement is critical to mediating immune and inflammatory responses. The observed interaction of the ICAM-1–like motif of SARS-CoV-2 S with TCRβ, in tandem with the interaction of the above discussed putative SAg motif (around
the insert PRRA) with TCRβ, is likely to further strengthen the association of the virus with the T cell and the ensuing activation. Precisely, T286 (S) makes close contacts with S94 (CDR3); E281 (S) forms a hydrogen bond with T51 (CDR2); and N280 and N282 (S) closely associate with R69 (Fig. 4D).

**A Neurotoxin-like Fragment at the RBD May also Bind αβTCR, Thus Further Enhancing the Immune Response.** Further examination of the SARS-CoV-2 S segments sequentially homologous to the neurotoxin-like sequences identified (23) for SARS1 S (rows colored green, Fig. 5A) pointed to two motifs conserved between the two CoVs: SARS-CoV-2 S1 residues 299 to 351 partially overlapping with the RBD and the S2 residues 777 to 807. Our simulations in search of possible binding of TCRs to these neurotoxin-like motifs showed the high propensity of the motif T299 to Y351 (second highest affinity site after the PRRA region) and its counterpart in SARS1 (first highest-affinity site) to bind TCRs. Fig. 5C illustrates these complexes. A recent study (24) detected significant T cell reactivity against 66 epitopes on the SARS-CoV-2 S glycoprotein in people who have not been exposed to the virus, inviting attention to possible memory response acquired upon exposure to human CoVs (HCoVs) such as common cold HCoV-OC43, HCoV-HKU1, HCoV-NL63, and HCoV-229E, which share sequence homology with SARS-CoV-2 genome. A total of 142 such cross-reactive epitopes were identified upon screening 474 peptides in the SARS-CoV-2 proteome (24).

We examined whether the neurotoxin-like regions identified here were among these cross-reactive epitopes. Notably, of the top-ranking four epitopes (ranked by T cell reactivity measured by spot-forming cells [SFC]/10^6 PBMCs), two (peptides 321 to 335 and 316 to 330) belong to the neurotoxin-like fragment T299 to Y351, and one (peptide 236 to 250) to the fragment 234 to 262. In fact, the former was completely spanned by eight partially overlapping cross-reactive epitopes as illustrated in SI Appendix, Fig. S4, pointing to the distinctive ability of this region to trigger CD4+ T cell response. These observations provide strong support to the predicted high affinity of this motif to bind TCR (Fig. 5C).

Overall, this neurotoxin-like sequence T299 to Y351 deserves attention as a possible source of central nervous system (CNS) disorders in COVID-19 patients.

Among the 66 epitopes, we note 661 to 675, which lies within the SAg-like region E661-R685 (Fig. 5), albeit at low reactivity and frequency. The absence of the insert PRRA among the cross-reactive epitopes is not surprising given that this insert is unique to SARS-CoV-2 S among all SARS-related βCoVs, and cross-reactivity increases with sequence similarity to antigens/peptides to which the donors have been already exposed to. The sequence identity between SARS-CoV-2 and SARS1 is 40% in the portion 671 to 685 of the SAg-like region, and the percentage of cross-reactive peptides having 33 to 40% sequence identity is reported to be 1% (24). On the other hand, it is interesting to note that, in a recent study (25) on epitopes that show strong T cell reactivity in convalescent patients who experienced severe COVID-19, an epitope (680 to 688) overlapping with the PRRA-containing part of SAg region is predicted to be one of the highest-affinity epitopes binding to HLA.

**A Rare Mutation, D839Y/E, Recently Observed in a SARS2 Strain from Europe May Contribute to Stabilizing the Interaction with TCR.** Interestingly, the SARS-CoV-2 S binding region harbors three residues that have been recently reported to have mutated in new strains from Europe and the United States (26, 27): D641G, A831V, and D839Y/N/E. The former two may potentially interact with MHCIIB, based on a ternary model we generated for SARS-CoV-2 S, MHCIIB, and TCR (**SI Appendix, Fig. S5**), while the latter (D839) is close to TCRβ and strongly interacts with N30 (Fig. 4A and C and **SI Appendix, Fig. S6**). The substitution of D839 by tyrosine strengthens the interactions between the spike and TCRβ. The interfacial interactions in the D839Y
mutant are stabilized by a hydrogen bond between Y839 and D839, an aromatic (polar-π) interaction between Y839 and N30, and possible electrostatic interactions with K73 and S97. The change in binding affinity between the spike and TCR upon mutating D839 to tyrosine is ΔΔG_D→Y = −0.9 ± 0.7 kcal/mol, indicating an approximate fourfold increase in binding affinity upon substituting the aspartic acid by a tyrosine at this position. The same qualitative effect is valid, but to a weaker extent, in the mutations to asparagine or glutamic acid. See SI Appendix, Table S1 for details on the method and results.

**TCR Repertoire Analysis Shows TCRβ Skewing and Junctional Diversity Indicating SAg Effect in Patients with Severe and Hyperinflammatory COVID-19.** SAg binding to specific TCR Vβ chains results in Vβ skewing, such that T cells with specific Vβ chains and diverse antigen specificity dominate the TCR repertoire (11, 13). If the motif we identified in SARS-CoV-2 S acts as an SAg, we reasoned that patients with mild/moderate COVID-19 disease courses and recovery without hyperinflammation would show adaptive immune responses mediated by T cells recognizing SARS-CoV-2 epitopes in a CDR3-mediated fashion, whereas patients with severe/hyperinflammatory COVID-19 would show immune responses consistent with at least partial SAg recognition. We analyzed, to this aim, next generation sequencing (NGS) immunosequencing data from 38 patients (42 samples) with mild/moderate COVID-19 and 8 patients (24 samples) with severe, hyperinflammatory COVID-19, which were part of a previously studied cohort (28). Principal component analysis (PCA) of the TCR β-chain variable gene (TRBV) repertoires corresponding to the two groups revealed that patients with mild/moderate COVID-19 cluster apart from those with severe/hyperinflammatory COVID-19 (Fig. 6A).

Differential gene usage analysis showed that several TRBV genes were overrepresented in the severe/hyperinflammatory COVID-19 patient group (Fig. 6B). In contrast, PCA of J gene distribution showed much less skewing, suggesting a selective pressure was preferentially exerted on the V gene distribution (Fig. 6C). To further investigate J gene diversity specifically for the V genes overrepresented in the severe/hyperinflammatory...
SARS-CoV (SARS1) S - TCR complex

266-288 74%
227-249 30%
759-789 84%

Downloaded from https://www.pnas.org by 71.234.245.10 on March 19, 2022 from IP address 71.234.245.10.

SI Appendix

identical to the protein product of these three genes. As shown in the SARS-CoV-2 S at its SAg-like region and form ternary spike. Overall, these simulations showed that these TCRs interactions also indicated the possibility of energetically favorable ternary complex formation between these TCRs, MHCII, and T cells, as well as IL-1 and TNFα from antigen-presenting cells (12). This cytokine storm leads to multiorgan tissue damage like what is now observed in MIS-C. Our results suggest that the hyperinflammatory syndrome originates from superantigenic activity by SARS-CoV-2 S glycoprotein. Furthermore, these findings also raise the possibility that the hyperinflammation observed in severe cases of COVID-19 in adults may also be driven by the SAg-like activity of the S protein. Indeed, SAgS induce an inflammatory cytokine signature similar to that which predicts severity and death in COVID-19, including IL-6, TNFα, IL-8, and IL-10 (12, 29). Moreover, our analysis of the T cell immune response in COVID-19 patients shows that those with more severe and hyperinflammatory clinical courses exhibit TCRβ skewing consistent with SAg activity.

Interestingly, a thorough study of SARS1 immunogenicity, conducted with a cohort of 128 individuals who have recovered from SARS1 (30), showed that the SARS1 spike 18-mer D649-L666 (DIPIGAGICASYHTVSLL) is one of the peptides most frequently recognized by T cells, among the screened 1,843 peptides that span the whole SARS1 CoV proteome [table 3 in Li et al. (30)]. This segment coincides with the SARS1 S region E647 to R667 that is sequentially (and structurally) homologous to our SARS-CoV-2 spike SAg-like motif E661 to R685 (Fig. 34, bottom alignment). This provides a very strong support for the T cell stimulatory ability of our SAg motif, given that it shares 12/18 amino acids with that SARS1 18-mer. And the remaining amino acids (including the insert PRRA, not present in SARS1 S) would endow even stronger superantigenic properties by virtue of their close similarity to the aligned SEB fragment.

Our findings raise the exciting possibility that immunomodulatory therapeutic options used for TSS may also be effective for MIS-C, including intravenous immune globulin (IVIG) and steroids. Indeed, initial published and unpublished reports suggest that MIS-C patients respond well to IVIG with or without steroids (5–7). IVIG contains antibodies that neutralize SEB (31). Given structural similarities between SEB and the SARS-CoV-2 S protein SAg motif, there is potential for cross-reactivity of these immunoglobulins, which may, in part, explain the response of MIS-C cases to IVIG. Other Food and Drug Administration-approved antiinflammatory drugs tested in models of SEB TSS may also be effective, including CTLA4-Ig which can inhibit CD28 costimulation (32), and the mammalian target of rapamycin (mTOR) inhibitor rapamycin (33), which is already in use for COVID-19. In addition, humanized monoclonal anti-SEB Abs (34, 35) could be of potential therapeutic benefit in MIS-C patients. Notably, it has been shown, in the mouse model of TSS, that lethal SEB superantigen challenge can be prevented by short peptide mimetics of its superantigen motif (21). It would be interesting to examine whether short peptide mimetics of the SARS-CoV-2 spike superantigen region might be employed to prevent/attenuate inflammatory cytokine gene induction and toxic shock in MIS-C patients.

At present, the majority of COVID-19 antibody therapies under investigation are designed to target the RBD, and some the N-terminal domain (NTD), of SARS-CoV-2 spike (36–44).
Our simulations also indicate that the RBD might potentially interact with TCRs. However, our study suggests that the “stem” of the spike, or its SAg-like motif, might also serve as a target. Compared with RBDs, relatively fewer mutations are found in the SAg region of SARS-CoV-2; notably, the PRRA insert is unique to SARS-CoV-2 and retained among all of its isolates sequenced to date (26, 27). It is also important to note that most of the cryo-EM structures reported for SARS-CoV-2 S protein have GSAS or GSGS substitution at R682RAR685, following the original work by Wrapp et al. (17). This same segment is either mutated (42–45) or removed (38–41) to enable protein expression, in cryo-EM studies of antibody-, nanobody-, or Fab-bound spike. The “mutant spike” is thus devoid of the unique character that would be otherwise endowed by the polybasic insert R681RAR685 and the high reactivity of R682RAR685 may have eluded these studies. It might be constructive to design antibodies or drugs targeting this SAg region, to not only modulate the SAg-induced inflammatory cytokine gene induction (12) but also block the cleavage essential to enabling viral entry (1, 19). Alternatively, combination therapies that target both the SAg-like region and the RBD can prove useful.

Fortunately, severe respiratory manifestations of COVID-19 in children as well as development of MIS-C are rare. This may be due to trained immunity (2). T and B cells play an important role in the antiviral response. CD4+ and CD8+ T cells from convalescent COVID-19 patients can recognize a range of SARS-Cov-2 epitopes, and the S protein represents a major target. Interestingly, T cells from unexposed individuals also respond to S protein epitopes from SARS-CoV-2, which supports the hypothesis of cross-viral immunity from other coronavirus strains (24, 46). However, why only a fraction of infected children develop MIS-C is unclear. We show that the mutation D839Y found in a European strain of SARS-CoV-2 enhances the binding affinity of the SAg motif to the TCR. This could (at least partly) explain the geographical skewing of MIS-C to areas where the European strain is endemic. It is also possible that a poor initial antibody response to the virus fails to neutralize the SAg, as recently shown in MIS-C patients (47), leading to immune enhancement following reexposure. Certain HLA types are more permissive of binding SAg, and, indeed, HLA has been shown to play a role in COVID-19 susceptibility (48). Of the nine cases initially reported in the United Kingdom, six were of Afro-Caribbean descent, which also suggests a potential genetic component to susceptibility (5).

Fig. 6. Skewing of TRBV usage in severe/hyperinflammatory COVID-19 patients; 24 repertoires of severe/hyperinflammatory COVID-19 cases versus 42 repertoires of mild/moderate COVID-19 cases were analyzed with and without 23 repertoires of age-matched healthy donors (age-matched to severe/hyperinflammatory COVID-19 group). (A) PCA of TRBV usage. Principal components 1 and 2 are shown; percentage of axis contributions are given in parentheses. Statistical analysis was performed using multivariate analysis of variance (MANOVA) Pillai-Bartlett test. (B) TRBV usage. The fraction of individual TRBV genes per repertoire is shown as mean ± SEM. TRBV genes are sorted to enriched fractions in severe/hyperinflammatory versus mild/moderate COVID-19 disease in ascending order from bottom to top. The top TRBVs enriched in severe/hyperinflammatory COVID-19 patients (TRBV5-6, TRBV14, TRBV13, and TRBV24-1) are enlarged in Inset. (C) PCA of TRBJ usage as described in A. See also SI Appendix, Fig. S7.
addition, ~80% of individuals over age 12 y harbor anti-SEB antibodies (49, 50), which may provide protection against the SAgs effects of SARS-CoV-2 S protein. The prevalence of preexisting anti-SEB antibodies could also contribute to the age distribution of severe COVID-19 cases in adults, as protective SEB titers fall in older adults after age 70 y. It is interesting to note that approximately a third or fewer of MIS-C patients tested positive for SARS-CoV-2, but the majority (but not all) had serologic evidence of infection or a history of exposure to COVID-19 (5–7). This may suggest that the SARS-CoV-2 SAg causes a delayed hyperinflammation response in certain children. SAgs have been implicated in autoimmunity by triggering self-reactive T cells (11). Antibody-mediated enhancement upon reexposure to the virus may also contribute to uncontrolled infection and inflammation (51). It is also possible that, despite a negative nasopharyngeal PCR test, the virus may still be present in the GI tract (52). MIS-C patients demonstrate unusually severe GI symptoms, abdominal pain, vomiting, and diarrhea, in addition to severe myocardial dysfunction and cardiac shock (5–7) and such severe GI symptoms are also frequently associated with the response to SAgs (9). In the case of SEB, cleavage and release of a specific fragment is responsible for induction of GI symptoms. Whether the SARS-CoV-2 SAg-like neurotoxin-like fragment could be similarly cleaved and would underlie the GI symptoms observed in MIS-C patients remains to be determined. We also observed that a neurotoxin-like segment (T299 to Y351) partially overlapping with the RBD exhibited a high affinity to bind TCRs. Notably, this region was recently observed to elicit strong and frequent T cell reactivity mediated by CD4+ T cells in donors who have not been exposed to SARS-CoV-2 (24). This invites attention to its ability to potentially trigger neurotoxic immune response in individuals who have not been exposed to CoVs that contain sequentially homologous peptides. Moving forward, it will be important to establish the significance of this region in the neurological disorders that are also frequently reported in children with MIS-C as well as adults.

In summary, we made five major observations: 1) PRRAR and sequential neighbors interact with TCRβ CDRs, and this association closely resembles that of SEB SAg with TCRβ; 2) nearby D839 participates in this interaction, and its mutation to tyrosine could strengthen the association with TCRβ; 3) a sequence motif (N280 to T286) typical of ICAM-1 interacts with the TCRα, further stabilizing the association between the spike and host cell TCR; 4) a neurotoxin-like motif (T299 to Y351) shows a high tendency to bind TCRs and potentially trigger neurotoxic responses, and this latter effect may be attenuated if the SARS-CoV-2 SAg–infected individual has been exposed HCoVs that contain homologous segments, as suggested (24) in a recent study; and 5) adult patients with severe/hyperinflammatory COVID-19 exhibit a skewed TCRβ repertoire distinguishing them from patients with mild/moderate COVID-19. Overall, our results from both computational modeling and NGS immunosequencing of TCRβ analysis of human samples indicate that strategies used for the treatment of SEB-mediated TSS or approaches to block the interaction of the S protein with TCRs may help reduce hyperinflammatory manifestations or (neuro)toxic effects of COVID-19 in both adults and children.

Materials and Methods

**SARS-CoV-2 (P0DTC2) and SARS-CoV (CVHSA_P59594)** spike models were generated using SWISS-MODEL (53), based on the resolved spike glycoprotein structures of SARS-CoV-2 (17) (PDB id code 6VSB) and SARS-CoV (54) (PDB id code 6M00). The crystal structures were built using libraries of backbone fragments (55) or by constraint space de novo reconstruction of these backbone segments (56). Two mutants associated with European COVID-19 patients (26) were constructed using CHARMM-GUI (57): one is the main strain mutant D614G, and the other contains four mutations including Q239K, A313V, D614G, and D839Y. These two SARS-CoV-2 spike mutants together with the SARS-CoV-2 (P0DTC2) originally taken from Wuhan were used to investigate the binding to αTCR, and MHCII (PDB id code 2XN9) (16) using ClusPro (18) and PRODIGY (58). See details in **SI Appendix**. The NGS immunosequencing data were analyzed using the methods described earlier (59, 60), summarized in **SI Appendix**.

**Data Availability.** All study data are included in the article and **SI Appendix**.

**ACKNOWLEDGMENTS.** We gratefully acknowledge support from NIH Awards P41 GM103712 (to I.B.) and R01 AI072726 (to M.A.), and useful comments from Drs. A. M. Brufsky, M. T. Lotze, S.-J. Gao, and S. D. Shapiro at the University of Pittsburgh Medical Center. We thank the clinical COVID-19 team at Halle for contributing samples to the Biobank.

---

10. M. Hoffmann et al., SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell 181, 271–280.e8 (2020).
44. Y. Cao et al., Potent neutralizing antibodies against SARS-CoV-2 identified by high-throughput single-cell sequencing of convalescent patients’ B cells. Cell 182, 73–84.e16 (2020).