CORONAVIRUS

Clonal analysis of immunodominance and cross-reactivity of the CD4 T cell response to SARS-CoV-2

Jun Siong Low1,2, Daniela Vaquerinheiro1,1, Federico Mele1,1, Mathilde Fogliieri1, Josipa Jerak1, Michela Perotti1, David Jarrossay1, Sandra Jovic1, Laurent Perez2, Rosalia Cacciatore2, Tatiana Terrot3, Alessandra Franzetti Pellanda4, Maira Biggiogero4, Christian Garzoni4, Paolo Ferrari5,6,7, Michela Perotti1, David Jarrossay1, Sandra Jovic1, Laurent Perez2, Rosalia Cacciatore2, Tatiana Terrot3, Alessandro Ceschi3,5,8,9, Antonio Lanzavecchia10, Federica Sallusto1,11

The identification of CD4+ T cell epitopes is instrumental for the design of subunit vaccines for broad protection against coronaviruses. Here, we demonstrate in COVID-19–recovered individuals a robust CD4+ T cell response to naturally processed severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spike (S) protein and nucleoprotein (N), including effecter, helper, and memory T cells. By characterizing 2943 S-reactive T cell clones from 34 individuals, we found that the receptor-binding domain (RBD) is highly immunogenic and that 33% of RBD-reactive clones and 94% of individuals recognized a conserved immunodominant S346–S365 region comprising nested human leukocyte antigen DR (HLA-DR)– and HLA-DP–restricted epitopes. Using pre– and post–COVID-19 samples and S proteins from endemic coronaviruses, we identified cross-reactive T cells targeting multiple S protein sites. The immunodominant and cross-reactive epitopes identified can inform vaccination strategies to counter emerging SARS-CoV-2 variants.

The identification of T cell epitopes in disease-causing organisms is challenging in view of the polymorphism of human leukocyte antigen (HLA) molecules and the variability of rapidly mutating pathogens. In the context of the COVID-19 pandemic, bioinformatic analysis (1) has been used to predict T cell epitopes in severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) proteins (2, 3) and to produce peptide pools to stimulate peripheral blood mononuclear cells (PBMCs) and enumerate antigen-specific T cells. These studies revealed a robust CD4+ and CD8+ T cell response against SARS-CoV-2 proteins in recovered patients (2–6) and a level of cross-reactivity with endemic coronaviruses in pre-pandemic samples (7–9).

A limitation of bioinformatics predictions is the difficulty in identifying immunodominant epitopes, because immunodominance is determined by multiple factors such as antigen processing, T cell repertoire, HLA alleles, and preexisting cross-reactive immunity (10–12). To identify naturally processed immunodominant CD4+ T cell epitopes, we took the unbiased approach of stimulating T memory (Tm) cells with protein-pulsed antigen-presenting cells (APCs), followed by the isolation of T cell clones to precisely map the epitope recognized (13).

PBMCs from a first cohort of 14 patients who had recovered from mild to severe COVID-19 (table S1) were used to isolate total CD4+ T cells or T central memory (Tcm) cells, T effector memory (Tem), and circulating T follicular helper (cTfh) cells (fig. S1A). The cells were labeled with carboxyfluorescein diacetate succinimidyl ester (CFSE) and stimulated with autologous monocytes in the presence of recombinant SARS-CoV-2 spike (S) protein or nucleoprotein (N). In all individuals, we observed a strong response to both antigens in terms of proliferation and interferon-γ (IFN-γ) production (Fig. 1, A and B, and fig. S1, B and C). Proliferating cells were detected at different levels in Tcm, Tem, and cTfh cells, consistent with a recent report (14), and over a 1-year period (fig. S1D). By contrast, the CD4+ T cell response to SARS-CoV-2 proteins in unexposed individuals was low or undetectable (Fig. 1B and fig. S1C), consistent with the presence of a few cross-reactive T cells primed by endemic coronaviruses (4, 5, 9).

The clonal composition of SARS-CoV-2–reactive T cell responses and the relationship between different memory subsets was studied in three individuals (P28, P31, and P33) by T cell receptor (TCR) Vβ sequencing. The Tcm, Tem, and cTfh cell lines comprised, on average, 908, 480, and 697 S-reactive clonotypes and 1452, 623, and 908 N-reactive clonotypes, respectively (Fig. 1C and fig. S2). Unexpectedly, several of the most expanded clonotypes were shared between two subsets, and even among all three subsets (Fig. 1, C and D), indicating a polyfunctional response consistent with previous studies on intraclonal diversification of antigen-primed CD4+ T cells (15, 16).

In view of the interest in the design of a subunit vaccine, we analyzed in depth the CD4+ T cell response to the S protein, in particular to the receptor-binding domain (RBD), which is the main target of neutralizing antibodies (17, 18). CD4+ T cells from a larger cohort of 34 COVID-19-recovered individuals (table S1) were stimulated with S protein–pulsed monocytes, and proliferating T cells were cloned by limiting dilution. We observed 2943 T cell clones and mapped their specificity using three pools of peptides spanning S1RBD, RBD, and S2 (Fig. 2, A and B). RBDb-specific T cell clones were found in 32 out of 34 donors, accounting for, on average, 20% of the response to the S protein (Fig. 2B). Using a matrix-based approach, we mapped the epitope specificity of 1254 RBDb-reactive CD4+ T cell clones (Fig. 2C) and found that, in each individual, the clones recognized multiple sites that collectively spanned almost all of the RBD sequence. However, certain regions emerged as immunodominant, such as those spanning residues S346–S385 and S446–S485. A 20–aminoc acid region (S346–S365) was recognized by 94% of the individuals (30 out of 32) and by 33% of the clones (408 out of 1254) (Fig. 2D). This region is highly conserved among human sarbecoviruses, including the recently emerged variants of concern and zoonotic sarbecoviruses (Fig. 2E) (19). RBDb- and S346–S365–specific T cell clones were found in different memory subsets of COVID-19-recovered individuals and were also isolated from individuals after SARS-CoV-2 mRNA vaccination (fig. S3). Thus, RBDb is highly immunogenic in vivo and contains a large number of naturally processed T cell epitopes, including a conserved immunodominant region.

To study the CD4+ T cell response to the immunodominant S346–S365 region, we sequenced TCR Vβ chains of 329 specific T cell clones. The 206 clonotypes identified used a broad spectrum of TCR Vβ genes and, even in the same individual, carried different CDR3 sequences (Fig. 3, A and B, and table S2). In P31 and P33, certain S346–S365 clonotypes were detected among the top 5% expanded Tc cells ex vivo (Fig. 3C). Using blocking antibodies, we determined that most of the T cell clones analyzed (n = 247 from 22 individuals) were HLA-DR restricted, whereas the remaining clones (n = 50 from five individuals) were HLA-DP restricted and one was HLA-DQ restricted (Fig. 3, D and E). Using truncated peptides and T cell clones from individuals

1Institute for Research in Biomedicine, Università della Svizzera italiana, 6500 Bellinzona, Switzerland. 2Laboratory of Immunogenetics, Department of Transfusion Medicine and Immunohematology, Fondazione IRCCS Policlinico S. Matteo, 27100 Pavia, Italy. 3Clinical Trial Unit, Ente Ospedaliero Cantonale, 6500 Bellinzona, Switzerland. 4Clinic of Internal Medicine and Infectious Diseases, Clinica Luganese Moncucco, 6900 Lugano, Switzerland. 5Faculty of Biomedical Sciences, Università della Svizzera italiana, 6900 Lugano, Switzerland. 6Department of Internal Medicine, Ente Ospedaliero Cantonale, 6500 Bellinzona, Switzerland. 7Prince of Wales Hospital Clinical School, University of New South Wales, Sydney, New South Wales 2052, Australia. 8Division of Clinical Pharmacology and Toxicology, Institute of Pharmacological Sciences of Southern Switzerland, Ente Ospedaliero Cantonale, 6000 Lugano, Switzerland. 9Department of Clinical Pharmacology and Toxicology, University Hospital Zurich, 8091 Zurich, Switzerland. 10National Institute of Molecular Genetics, 20122 Milano, Italy. 11Institute of Microbiology, ETH Zürich, 8093 Zurich, Switzerland.

*Corresponding author: federica.sallusto@irb.usi.ch (F.S.); antonino.cassotta@irb.usi.ch (A.C.)
†These authors contributed equally to this work.
‡These authors contributed equally to this work.
with different HLA types (table S3), we defined two HLA-DR-restricted epitopes (VYAWNKRK-RIS and RPASVYAWNKR) and one HLA-DP-restricted epitope (NRKRISNCVAD) (Fig. 3F). Thus, the S346–S365 region comprises at least three nested epitopes recognized in association with different allelic forms of HLA-DR or HLA-DP by T cell clones that use a large set of TCR Vβ genes and CDR3 of different sequence and length.

To address the extent of T cell cross-reactivity between different S proteins, SARS-CoV-2 S protein–specific T cell lines from P28 and P33 were relabeled with CFSE and stimulated with S proteins from endemic human coronaviruses. In these secondary cultures, a robust proliferation was observed in response to SARS-CoV and HKU1 (Fig. 4A). Unexpectedly, a sizeable fraction of clonotypes in SARS-CoV-2 primary cultures (ranging from 7 to 25%) were found in SARS-CoV-2 and/or HKU1 secondary cultures, consistent with a substantial degree of T cell cross-reactivity (fig. S4). To corroborate this finding, we isolated from secondary cultures several T cell clones that proliferated in response to two or even three different naturally processed S proteins (Fig. 4B and table S4).

Cross-reactive T cells may derive from preexisting memory T cells or from the priming of naïve T cells. We therefore analyzed a COVID-19–recovered patient from whom we had previously cryopreserved PBMCs. A robust CD4+ T cell proliferation in the pre–COVID-19 sample was detected against NL63 and 229E S proteins, whereas the response to HKU1 and OC43 was limited and the response to SARS-CoV and SARS-CoV-2 undetectable (Fig. 4C). Conversely, in the post–COVID-19 sample, strong T cell proliferation was observed not only in response to SARS-CoV-2, but also in response to all other alpha and beta coronavirus S proteins (Fig. 4, C and D), and shared clonotypes were detected between SARS-CoV-2 and endemic coronavirus S protein–stimulated cultures (fig. S5A). Furthermore, T cell clones isolated from cultures stimulated by SARS-CoV, OC43, or NL63 proliferated in response to the SARS-CoV-2 S-peptide pool, and their specificity was mapped primarily to the S2 region (Fig. 4, E and F), consistent with its high degree of sequence conservation (20–22). T cell clones that fully cross-reacted with all S proteins were mapped to the highly conserved fusion peptide (Fig. 4G).

To determine whether S-reactive T cells in the post–COVID-19 sample could be detected in pre-pandemic samples, we performed clonotypic analysis of total T cell subsets in the post–COVID-19 sample and on samples collected in 2014 and 2017. Most of the SARS-CoV-2–specific clonotypes identified above were found only in the post–COVID-19 sample, consistent with priming of naïve T cells (fig. S5B). By contrast, clonotypes specific to endemic coronaviruses were found at a comparable number at all time points. Some T cell clones against the highly conserved fusion peptide could be tracked back to the 2014 sample and were found to be expanded in the post–COVID-19 sample (fig. S5C). These findings demonstrate that preexisting cross-reactive T cell subsets are recalled and expanded upon SARS-CoV-2 infection.

The robust CD4+ T cell response to the RBD and the identification of the S346–S365 immunodominant region conserved in the emerging SARS-CoV-2 variants of concern provide the rationale for the development of a subunit vaccine based on the RBD because it is the target of most neutralizing antibodies (17, 18). These findings were not anticipated...
in previous studies based on bioinformatics predictions (2, 3) and short-term peptide stimulation of PBMCs, highlighting the value of combining T cell stimulation with protein antigens with cloning and TCR sequencing for the analysis of antigen-specific T cell repertoires.

The immunodominance of RBD S346–S365 at the individual level and at the population level may be due to the presence of three nested T cell epitopes presented by HLA-DR and HLA-DP and to the relative abundance of naturally processed peptides, as recently reported in an immunopeptidomics study (23). The S346–S365 region is also a contact site for the broadly reactive neutralizing antibody S309 (24), providing a good example of convergence of B and T cells around a conserved epitope.
Our study also provides evidence for the recall of preexisting cross-reactive T<sub>n</sub> cells upon SARS-CoV-2 infection. However, this phenomenon, reminiscent of the “original antigenic sin” (25), does not prevent a robust and persistent primary response to new epitopes of SARS-CoV-2 that is characterized by extensive intraclonal diversification into T<sub>cm</sub>, cT<sub>fh</sub>, and T<sub>em</sub> cells, which represent inflammatory, helper, and long-lived T<sub>em</sub> cells, respectively (26, 27). The availability of a large number of cross-reactive T cell clones is not only instrumental for defining target sites in relevant pathogens but also for understanding whether cross-reactivity is due to epitope structural similarities or to TCR-binding degeneracy (11, 28).

The possibility of leveraging a robust, cross-reactive T helper cell function against conserved sites will be instrumental in driving neutralizing antibody responses to adaptive vaccines that incorporate escape mutations found in emerging SARS-CoV-2 variants.

REFERENCES AND NOTES

1. R. Vita et al., Nucleic Acids Res. 47 (D1), D339–D343 (2019).
2. A. Grifoni et al., Cell Host Microbe 27, 671–680.e2 (2020).
3. A. Niedel et al., Nat. Immunol. 22, 74–85 (2021).
4. A. Grifoni et al., Cell 181, 1485–1501.e15 (2020).
5. J. Braun et al., Nature 587, 270–274 (2020).
6. A. Tarke et al., Cell Rep. Med. 2, 100204 (2021).
7. J. Mateus et al., Science 370, 89–94 (2020).
8. N. Le Bert et al., Nature 584, 457–462 (2020).
9. P. Bacher et al., Immunity 53, 1258–1271.e5 (2020).
10. A. Kim, S. Sadegh-Nasseri, Curr. Opin. Immunol. 34, 9–15 (2015).
11. A. K. Sewell, Nat. Rev. Immunol. 12, 669–677 (2012).
12. J. W. Yewdell, Immunity 25, 533–543 (2006).
13. A. Cassotta et al., J. Exp. Med. 217, e2020026 (2020).
14. L. B. Rodda et al., Cell 184, 169–183.e17 (2021).
15. N. J. Tubo et al., Cell 153, 796–796 (2013).
16. S. Becattini et al., Science 347, 400–406 (2015).
17. L. Piccoli et al., Cell 183, 1024–1042.e21 (2021).
18. A. J. Greaney et al., Cell Host Microbe 29, 44–57.e9 (2021).
19. R. Lu et al., Lancet 395, 565–574 (2020).
20. J. A. Jaimies, N. M. André, J. S. Chappie, J. K. Millet, G. R. Whitaker, J. Mol. Biol. 432, 3309–3325 (2020).
21. M. A. Tortorici, D. Veesler, Adv. Virus Res. 105, 93–116 (2019).
22. A. C. Walls et al., Nature 531, 114–117 (2016).
Fig. 4. Identification of coronavirus S protein cross-reactive T cell clones. (A and B) CFSE-labeled CD4+ Tm cells from P28 and P33 were stimulated with autologous monocytes pulsed with recombinant SARS-CoV-2 S protein. CFSE<sup>+</sup> cells were expanded with interleukin-2 for 10 days, relabeled with CFSE, and restimulated with S proteins from human beta (SARS-CoV, HKU1, and OC43) or alpha (NL63 and 229E) coronaviruses. T cell clones from proliferating secondary cultures were isolated and tested for cross-reactivity against different S proteins. (A) CFSE profiles from primary and secondary stimulation in the absence or presence of the indicated antigens. (B) Proliferative response of representative T cell clones, isolated from secondary cultures, to autologous APCs pulsed with titrated doses of recombinant S proteins from SARS-CoV-2, SARS-CoV, or HKU1. Proliferation was assessed on day 3 after a 16-hour pulse with 3H-thymidine and is expressed as counts per minute (cpm). (C to G) Multiple blood samples were obtained from donor P34 several years before and 1.5 months after COVID-19 (disease onset Oct 2020) and characterized regarding Tm cells and serum antibody levels. (C) T cell proliferation measured by CFSE dilution in pre–COVID-19 (Dec 2018) and post–COVID-19 (Dec 2020) samples in response to autologous monocytes pulsed with different S proteins. (D) Time course of serum immunoglobulin G antibodies against different coronavirus S proteins as determined by enzyme-linked immunosorbent assay [half-maximal serum dilution (EC<sub>50</sub>) values]. These data demonstrate that, together with a strong induction of serum antibodies to SARS-CoV-2, antibody titers against HKU1 and OC43 also increased in the post–COVID-19 sample. (E) Proliferative response (day 3 cpm) to a pool of peptides spanning SARS-CoV-2 S protein of T cell clones obtained from post–COVID-19 CFSE<sup>+</sup> cultures stimulated by SARS-CoV-2, SARS-CoV, OC43, or NL63. Pie charts show the total number of clones tested and the fraction of responsive clones is shown in gray. (F) Specificity of SARS-CoV-2 S peptide pool-reactive T cell clones isolated from each culture (E) was further mapped by stimulation.
with pools of peptides spanning the S1DRBD, RBD, and S2 regions of the SARS-CoV-2 S protein. Histograms show the percentage of clones specific for each region. The total number of clones tested is indicated at the top. (G) Characterization of representative cross-reactive T cell clones isolated from P34 post–COVID-19 sample. Left panels report the proliferative response (day 3 cpm) of T cell clones stimulated with titrated doses of recombinant S proteins in the presence of autologous monocytes. The peptides recognized are indicated on the right panels. Shown are sequence alignments of the recognized SARS-CoV-2 epitopes (S816–S830 and S981–S1000) with homologous sequences of endemic alpha and beta coronaviruses. Dots indicate amino acid residues identical to the SARS-CoV-2 reference strain.

Funding: This work was supported in part by the Henry Krenter Foundation, the Cariplo Foundation (grant 2020-1374 CoVIM), the Swiss National Science Foundation (grant 189331), and by EOC research funds. F.S. and the Institute for Research in Biomedicine are supported by the Helmut Horten Foundation.

Author contributions: A.Ca. and F.S. contributed to study concept and design. J.S.L., D.V., F.M., and J.J. contributed to experimental work. M.P. and L.P. produced recombinant proteins. D.J. performed cell sorting. S.J. provided technical support. M.F. performed bioinformatics analysis. R.C. performed HLA typing. T.T., A.F.P., M.B., C.G., P.F., and A.Ce. contributed clinical samples. A.Ca., A.L., and F.S. analyzed the data and wrote the manuscript. All authors contributed to interpretation of data and critical revision of the manuscript.

Competing interests: The authors declare no competing interests.

Data and materials availability: TCR Vβ sequences have been deposited in the ImmuneACCESS database (29). All other data are available in the main text or the supplementary materials. This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6548/1336/suppl/DC1

Materials and Methods

Tables S1 to S5

Figs. S1 to S5

References (30, 31)

MDAR Reproducibility Checklist

View/request a protocol for this paper from Bio-protocol.

3 February 2021; accepted 13 May 2021

Published online 18 May 2021

10.1126/science.abb9865