Lack of evidence of significant homology of SARS-CoV-2 spike sequences to myocarditis-associated antigens

Daniel Marrama,[a] Jarjapu Mahita,[a] Alessandro Sette,[a,b,1] and Bjoern Peters,[a,b,1]*

[a]La Jolla Institute for Allergy and Immunology, La Jolla, CA 92037, USA
[b]Department of Medicine, University of California, San Diego, CA 92093, USA

EBioMedicine 2022;75: 103807
Published online xxx
https://doi.org/10.1016/j.ebiom.2021.103807

Summary

Background COVID-19 mRNA vaccines have proven to be highly safe and effective. Myocarditis is an adverse event associated with mRNA vaccination, especially in young male subjects. These events are rare and, in the majority of cases, resolve quickly. As myocarditis can be driven by autoimmune responses, we wanted to determine if the SARS-CoV-2 spike protein antigen encoded in the mRNA COVID vaccines had potential cross-reactivity with auto-antigens previously associated with myocarditis.

Methods We performed a sequence identity comparison between SARS-CoV-2 spike protein-derived peptides and myocarditis-associated antigens. We also performed a structural analysis of these antigens and the SARS-CoV-2 spike protein to identify potential discontinuous 3-D epitope similarities.

Findings We found no significant enrichment in the frequency of spike-derived peptides similar to myocarditis-associated antigens as compared to several controls.

Interpretation Our results do not support the notion that increased occurrence of myocarditis after SARS-CoV-2-spike vaccination is mediated by a cross-reactive adaptive immune response.

Keywords: SARS-CoV-2; Myocarditis; mRNA vaccines; Antigen homology; Autoimmune disease

Introduction

In late 2019, severe acute respiratory coronavirus 2 (SARS-CoV-2) emerged causing a global pandemic of COVID-19 disease resulting in widespread morbidity and mortality. COVID-19 typically presents as a dry cough, sore throat, fever, and loss of taste and smell,[1] but more rare complications have arisen as well including heart injury.[2] Following the rapid development and approval for emergency use of several different SARS-CoV-2 vaccines, as of December 2021, over eight billion COVID-19 vaccine doses have been administered worldwide.[3] Rare occurrences of myocarditis and pericarditis have been reported as associated with COVID-19 vaccination in the context of mRNA,[4,5] but only extremely rarely with viral vector-based vaccines which are in turn associated with a different class of adverse event such as increased frequency of blood clots.[6] The etiology of these rare side effects is poorly understood, but the possibility of autoimmune adaptive reactions needs to be investigated. As the two currently authorized mRNA vaccines BNT162b2 (Pfizer/BioNTech) and mRNA-1273 (Moderna) are both encoding the SARS-CoV-2 spike protein as the vaccine immunogen, we set out to determine if specific sequences contained in the spike protein could lead to a cross-reactive immune response to autoantigens associated with autoimmune myocarditis in particular.[7–9]

Methods

Myocarditis associated auto-antigens (cardiac proteins)

To compile a list of myocarditis-associated antigens, we first queried the Immune Epitope Database (IEDB),[10] which includes myocarditis-associated epitopes and their respective source antigens. A search for positive assays that included disease entries of "myocarditis" (DOID: 820), "rheumatic myocarditis" (DOID: 8481), and "experimental autoimmune myocarditis" (ONTIE ID: 0003439) revealed 66 human epitopes, which were contained in eight protein antigens. In addition, we reviewed the autoimmune myocarditis literature,[7–9]

*Corresponding author.
E-mail address: bpeters@lji.org (B. Peters).
Funding: This work was funded by 75N93019C00001 from the National Institutes of Health.
[1] Contributed equally to the work.

Research in context

Evidence before this study

An increase of Myocarditis events has been observed following COVID-19 mRNA vaccination. These adverse events are extremely rare and often not biopsy proven. There is no clear evidence for a mechanism that could lead to these events. One possible mechanism is the triggering of autoimmune reactions, based on cross-reactivity between the SARS-CoV-2 spike protein encoded by the mRNA vaccines and auto-antigens associated with myocarditis.

Added value of this study

We performed a similarity analysis between the SARS-CoV-2 spike protein encoded in the vaccine, and myocarditis-associated antigens reported in the literature. We focused on regions in these proteins that could be the targets of T cell- or B-cell responses. We did not find significant similarity between them, making it doubtful that there is cross-reactive recognition occurring in individuals who developed myocarditis post-COVID-19 mRNA vaccination.

Implications of all the available evidence

Without much evidence of an adaptive cross-reactive response occurring in these individuals, the incidents of post-vaccination myocarditis are unlikely to be T-cell or B-cell mediated and more consistent with an innate response.

which provided 23 additional antigens that had known associations with myocarditis and four antigens that were mentioned for potential associations with myocarditis, but either weak or no evidence was noted. In total, we compiled this list of 35 antigens (Table 1) to use for this conservation analysis.

Randomized human protein control sets

We compiled 1000 sets of 35 proteins each that were randomly selected from the human proteome (UniProt proteome ID: UP000005640) using custom Python scripts. These sets provide a control on how human proteins not specifically selected to be associated with myocarditis compare to the set described above.

Spike protein-derived peptides and shuffled controls

The SARS-CoV-2 spike protein (UniProtID: P0DTC2) is 1273 amino acids in length. Since cross-reactivity at the level of either $CD8^+$or $CD4^+$ T cells is of potential concern, we considered 9-mers and 15-mers, as these epitope sizes are associated with $CD8^+$or $CD4^+$ T cell epitopes, respectively. To identify possible peptides of relevance, we split the spike protein sequence into all possible 9-mers, overlapping by eight amino acids, and all possible 15-mers, overlapping by 14 amino acids using custom Python scripts. In total, we compiled 1265 9-mers and 1259 15-mers. As a control, we also generated shuffled sequences of all peptides using the Python shuffle function.

Conservation analysis

We considered different levels of sequence identity to identify potentially relevant hits for CD4- and CD8 T cell immune responses. Previous studies[11] support the notion that 50% is a conservative identity threshold for cross-reactivity for CD4 T cells, which are typically 15 residues in length. For CD8 epitopes, which are typically 9 residues in length, more than two substitutions are in general non-cross-reactive.[12] Both the spike peptide and shuffled peptide sets were searched for matches in the cardiac proteins, as well as the 1000 control sets, using PEPMatch, a tool developed by the IEDB (manuscript in progress; https://github.com/IEDB/PEPMatch). PEPMatch is optimized for short peptide searches, and guarantees finding complete sets of results in contrast to, for example, BLAST[13] with default settings.

3-D Structural analysis

To consider the potential for discontinuous 3-D epitope cross-reactivity from B cells, we analysed structural similarities between the SARS-CoV-2 spike protein and the myocarditis-associated antigens using TM-align.[14] PDB files for each antigen were extracted from the Protein Data Bank website (https://www.rcsb.org). Where PDB structures were not available for a protein, predicted structures created by AlphaFold[15] were used. This analysis was also repeated with 1000 control sets each containing 35 randomly selected proteins from the human proteome (UniProt proteome ID: UP000005640). Since TM-align normalizes its scores based on protein length, the proteins selected for these controls were made to fall within 30% of the average length of the myocarditis-associated proteins.

The solvent-accessible surface area of the residues making up the region of spike that have a TM-align score of 0.5 or above compared with the myocarditis-associated antigens were calculated with the program NACCESS[16] using a default probe size of 1.4 Angstroms.

Statistics

Statistics were performed using custom Python scripts with implementation of the SciPy library. For the conservation analyses, we used a Fisher's exact test to determine the association between homology and spike or shuffled peptides.

Protein Name	Gene	UniProt ID	Source
Myosin-6	MYO6	Q9UM54	IEDB
Myosin-7	MYH7	P12883	IEDB
Muscarinic acetylcholine receptor M2	CHRM2	P08172	IEDB
Myosin-binding protein C - cardiac-type	MYBPC3	Q14896	IEDB
Myosin-binding protein C - fast-type	MYBPC2	Q14324	IEDB
Beta-2-glycoprotein 1	APOH	P02749	IEDB
Laminin subunit alpha-1	LAMA1	P25391	IEDB
Transmembrane protease serine 4	TMPRSS4	Q9NRS4	IEDB
Troponin I	TNNI3	P19429	Review Literature
Troponin T	TNNT2	P45379	Review Literature
Beta-1 adrenergic receptor	ADRB1	P08588	Review Literature
Actin, alpha cardiac muscle 1	ACTC1	P68032	Review Literature
Tropomyosin alpha-1 chain	TPM1	P09493	Review Literature
Tropomyosin beta chain	TPM2	P07951	Review Literature
Tropomyosin alpha-3 chain	TPM3	P06753	Review Literature
Cytoplasmic aconitate hydratase	ACO1	P21399	Review Literature
ADP/ATP translocase 1	SLC25A4	P12235	Review Literature
Creatine kinase B-type	CKB	P12277	Review Literature
Creatine kinase S-type, mitochondrial	CKMT2	P17540	Review Literature
Creatine kinase U-type, mitochondrial	CKMT1A	P12532	Review Literature
Creatine kinase M-type	CKM	P06732	Review Literature
Desmin	DES	P17661	Review Literature
Dihydrolipoyl dehydrogenase, mitochondrial	DLD	P09622	Review Literature
60 kDa heat shock protein, mitochondrial	HSPD1	P10809	Review Literature
Heat shock 70 kDa protein 1A	HSPA1A	P0DMV8	Review Literature
Vimentin	VIM	P08670	Review Literature
E3 ubiquitin-protein ligase TRIM21	TRIM21	P19474	Review Literature
Lupus La protein	SSB	P05455	Review Literature
Pyruvate kinase	PKLR	P30613	Review Literature
Ubiquinol-cytochrome-c reductase complex assembly factor 1	UQCC1	Q9NVA1	Review Literature
Sodium/potassium-transporting ATPase subunit alpha-1	ATP1A1	P05023	Review Literature
Natriuretic peptides B	NPPB	P16860	Review Literature
Natriuretic peptides A	NPPA	P01160	Review Literature
Troponin C, slow skeletal and cardiac muscles	TNNC1	P63316	Review Literature
Transmembrane protein 65	TMEM65	Q6PI78	Review Literature

***Table 1*: Myocarditis-Associated Cardiac Antigens.**

Role of funders

This work was supported by the National Institutes of Health. The funders had no role in the design, collection of data, statistical analysis, or writing of the paper.

Results

To evaluate the occurrence of peptides in SARS-CoV-2 spike that have high similarity to peptides in proteins associated with cardiac autoimmunity (cardiac proteins for short hereafter), we generated a set of 1259 15-mers overlapping by 14 residues spanning the entire spike protein. 15-mer peptides were considered first, as the typical length of MHC-II restricted CD4 T cell epitopes. We compared these peptides to a set of 35 cardiac proteins associated with cardiac autoimmunity. We found zero peptides in the spike that matched any of these cardiac antigens at a sequence identity of 60% or more. Relaxing the identity threshold further, at 53% homology, we found 13 matches for peptides from the spike protein. However, we also found 14 matches from shuffled peptides, which means there is no statistically significant increased sequence identity of actual spike peptides as compared to shuffled controls at the 53% threshold ($p = 1.0$, OR=0.928 (Table 2)).

Next, we examined the homology of 9-mer peptide fragments, which is the length of typical MHC-I restricted CD8 T cell epitopes. At 78% homology or more (two substitutions), three spike peptides and one shuffled peptide were found in cardiac proteins, which is not a significant enrichment ($p = 0.63$). At the 67% homology level, we found 77 homologous peptides

	Match in Cardiac Proteins	No Match in Cardiac Proteins	Total Peptides
Spike Peptides	13	1246	1259
Shuffled Peptides	14	1245	1259
Total Peptides	27	2491	2518

***Table 2*: SARS-CoV-2 Spike 15-mers vs. Shuffled 15-mers (Homology >= 53%).**

	Match in Cardiac Proteins	No Match in Cardiac Proteins	Total Peptides
Spike Peptides	77	1188	1265
Shuffled Peptides	55	1210	1265
Total Peptides	132	2398	2530

***Table 3*: SARS-CoV-2 Spike 9-mers vs. Shuffled 9-mers (Homology >= 67%).**

from spike and 55 homologous from shuffled peptides (Table 3), which is also not a statistically significant increase ($p = 0.06$).

While these analyses do show a trend for a higher number of 9-mer peptides in spike that match the cardiac proteins, that enrichment is not statistically significant, and thus does not support the notion that spike protein sequences are significantly enriched in peptides that are potential epitopes with significant sequence identity to human self-proteins associated with autoimmune myocarditis. Conversely, the analysis also identifies 13 15-mer and 77 9-mer peptides that could be further evaluated experimentally for their potential to mediate cross-reactive responses in individuals experiencing post-vaccination myocarditis (Supplemental Table 1).

As an alternative control, we randomly selected sets of human proteins to match the cardiac protein set. We then repeated the peptide match analysis. For 9-mers at the 56% homology level, 89.5% of sets were below the cardiac protein set and 10.5% were at or above it in terms of peptide match frequency. At the 67% homology level, 84.8% were below and 15.2% were at or above the cardiac protein set (Figure 1). This shows a trend for increased hits in cardiac proteins rather than in randomly selected proteins, but as before this trend is not statistically significant at the conventional $p = 0.05$ cut-off. Only spike 15-mers at the 53% homology level had matches within the cardiac protein set and 48.7% of the randomly selected protein sets were below it and 51.3% were at or above it in terms of peptide match frequency (Figure 2). This is also not considered significant.

All PDB files for the myocarditis-associated antigens were compared to the SARS-Cov-2 Spike protein using the TM-align program (https://zhanggroup.org/TM-align/) with the structure of the spike protein (PBD ID: 7DDD). TM-align scores are considered significant when greater than or equal to 0.5. Four substructures of these antigens had significant scores (Table 4). Since these are only fragments of the antigen, we mapped their residues onto the 3-D structure of the spike, which

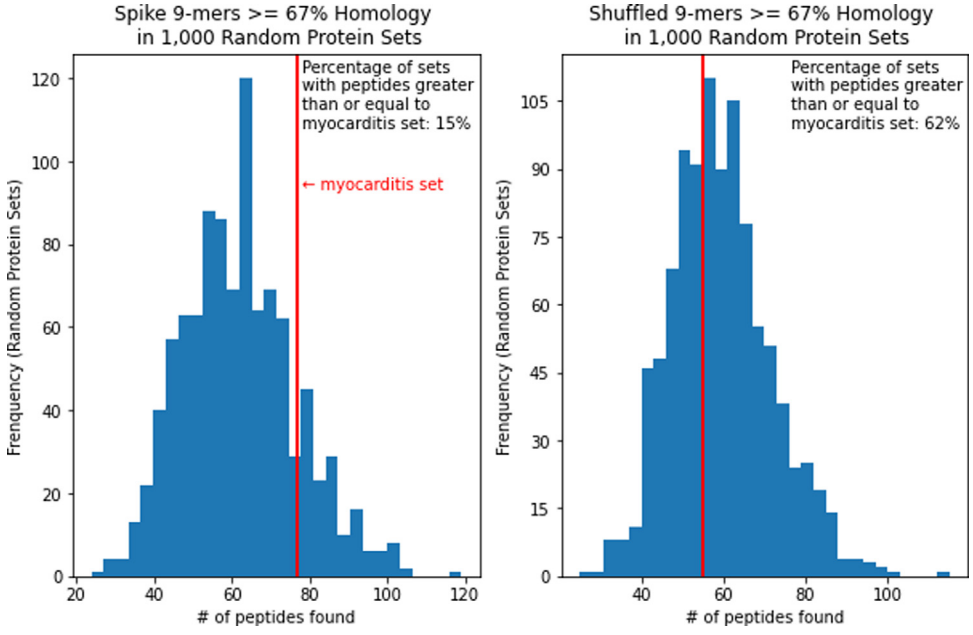

Figure 1. Spike vs shuffled 9-mers >= 67% homology match distribution of 1000 random protein sets.

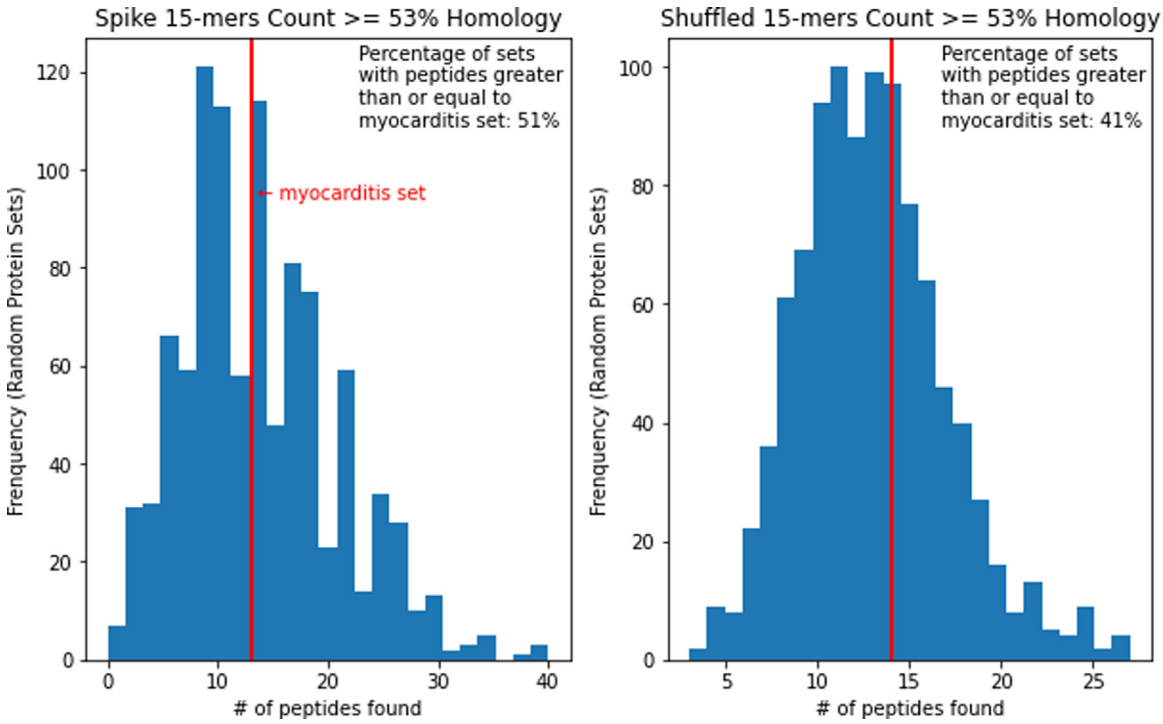

Figure 2. Spike vs shuffled 15-mers >= 53% homology match distribution of 1000 random protein sets.

PDB ID	Source Protein	UniProt ID	TM-align Score	Residue count
3SSU	Vimentin	P08670	0.742	91
5KHT	Tropomyosin alpha-1 chain	P09493	0.857	29
5WLQ	Myosin-7	P12883	0.516	79
6OTN	Tropomyosin alpha-3 chain	P06753	0.621	74

***Table 4*: Significant TM-align scores for antigen fragments compared with spike.**

shows the location of these regions and their proximity to the surface (Figure 3). Using NACCESS to calculate solvent accessibility, we found that all of these residues had values under 100 square Angstroms. Since only residues with values between 100 and 120 square Angstroms are considered fully exposed, these residues are considered to have a low solvent-accessible surface area.

We then repeated this analysis with 1000 random protein sets that have an average length within 30% of the average protein length in the cardiac set. This was to create a distribution of average TM-align scores normalized for the non-spike proteins. We found that 84.5% of sets were below the cardiac protein set and 14.5% were at or above it in terms of mean TM-align score. This would not be considered significant (Figure 4).

Discussion

Myocarditis is an inflammatory disease that affects the muscles of the heart which can be caused by an autoimmune mechanism.[8] There have been a number of cases of myocarditis occurring in humans after SARS-CoV-2 infection and with COVID-19 vaccination.[4,17] Although these occurrences are extremely rare and often not biopsy proven,[18] investigation into a possible adaptive immune response is warranted. Here, we examined the potential for a cross-reactivity link based on sequence similarity of the SARS-CoV-2 spike protein encoded in mRNA COVID-19 vaccines and myocarditis-associated proteins. We did not find statistically significant overlap in terms of linear peptide sequences between cardiac proteins and the spike protein when considering various controls, which would be potential targets of T cell responses. When considering potential 3-D epitope cross-reactivity, which would be targeted by antibodies, we did not find these antigens were significantly higher in structural similarity compared with controls. The antigens that had some structural similarities were similar in spike regions that appear inaccessible, making them unlikely epitope targets of antibody cross-reactivity. This does not support the hypothesis that myocarditis adverse events post-mRNA COVID-19 vaccination are due to cross-reactive reactions of the adaptive immune system. This is further supported by the fact that the median onset of myocarditis incidents occurring post-vaccination was three and a half days and for

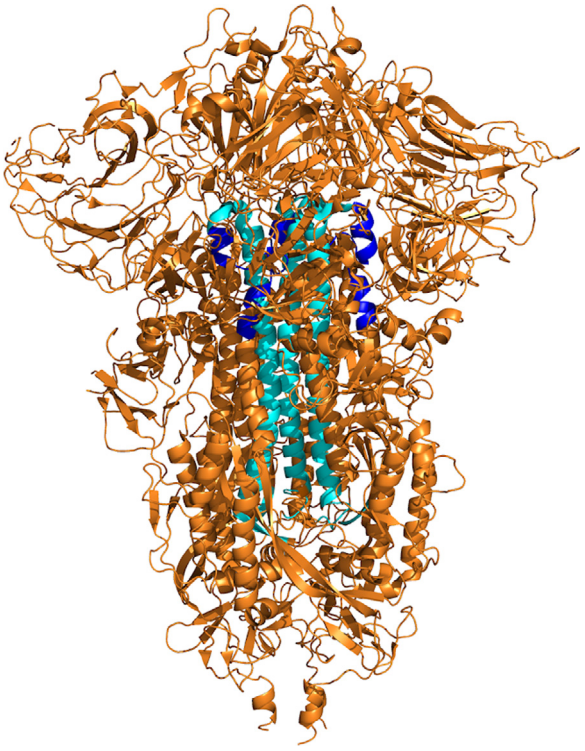

Figure 3. Regions of the spike protein with significant TM-align scores compared with myocarditis antigens (highlighted in cyan and blue).

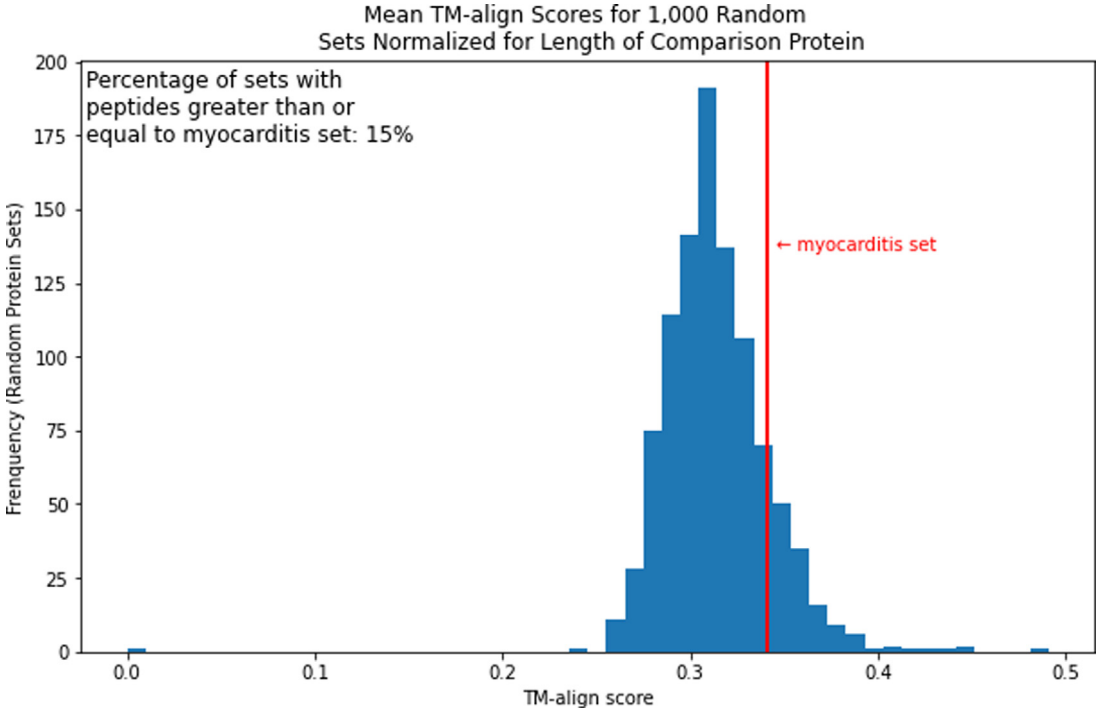

Figure 4. Mean TM-align score distribution of 1000 random protein sets.

those hospitalized, the median discharge was two days. By contrast, autoimmune diseases often progress over time through epitope spreading.[19] Overall, the incidents of myocarditis post-vaccination may not be T cell-mediated and perhaps are more compatible with a transient innate response.

However, the lack of statistical evidence of similarity between vaccine peptides and autoimmune antigens, in general, does not exclude that, in some individuals, there will be a cross-reactive response. Our analysis does not exclude cross-reactivity as a mechanism for post mRNA COVID-19 vaccine myocarditis, and more evidence is required to elucidate the mechanisms involved. Since median onset of myocarditis seems inconsistent with cross reactive adaptive immunity as a mechanism, future investigations might address additional mechanisms, for example, associated with activation of innate immunity, and employ experimental rather than sequence comparison methodologies.

Declaration of interests

DM has nothing to disclose.
JM has nothing to disclose.
AS has participated in a Moderna Advisory Board.
BP has nothing to disclose.

Acknowledgements

We wish to acknowledge the work of the entire IEDB team.

Contributors

Conceptualisation and study design: BP
Data curation: DM and AS
Formal analysis: DM, JM, and BP
Funding acquisition: AS and BP
Methodology: BP
Validation: DM, JM, and BP
Visualization: DM
Writing - original draft: DM, AS, and BP
Writing - review & editing: DM, JM, AS, and BP

Data Sharing Statement

Public data used for this work can be found at UniProt (https://www.uniprot.org/) and the Protein Data Bank (https://www.rcsb.org/). The code used for the analysis will be made available upon request to the corresponding authors.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.ebiom.2021.103807.

References

1 Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020;**395**(10223):497–506.
2 Han H, Xie L, Liu R, Yang J, Liu F, Wu K, et al. Analysis of heart injury laboratory parameters in 273 COVID-19 patients in one hospital in Wuhan, China. *J Med Virol* 2020;**92**(7):819–23.
3 WHO Coronavirus (COVID-19) Dashboard [Internet]. [cited 2021 Dec 15]. Available from: https://covid19.who.int
4 Diaz GA, Parsons GT, Gering SK, Meier AR, Hutchinson IV, Robicsek A. Myocarditis and pericarditis after vaccination for COVID-19. *JAMA* 2021. https://doi.org/10.1001/jama.2021.13443. [cited 2021 Sep 2]; Available from.
5 CDC. COVID-19 Vaccination [Internet]. Centers for Disease Control and Prevention. [cited 2021 Sep 13]. Available from https://www.cdc.gov/coronavirus/2019-ncov/vaccines/safety/myocarditis.html.
6 CDC. COVID-19 Vaccination [Internet]. Centers for Disease Control and Prevention. [cited 2021 Sep 12]. Available from https://www.cdc.gov/coronavirus/2019-ncov/vaccines/safety/JJUpdate.html.
7 Bracamonte-Baran W, Čiháková D. Cardiac autoimmunity: myocarditis. *Adv Exp Med Biol* 2017;**1003**:187–221.
8 Maisch B. Cardio-immunology of myocarditis: focus on immune mechanisms and treatment options. *Front Cardiovasc Med* 2019;**6**:48.
9 Tschöpe C, Ammirati E, Bozkurt B, Caforio ALP, Cooper LT, Felix SB, et al. Myocarditis and inflammatory cardiomyopathy: current evidence and future directions. *Nat Rev Cardiol* 2021;**18**(3):169–93.
10 Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, et al. The immune epitope database (IEDB): 2018 update. *Nucleic Acids Res.* 2019;**47**(D1):D339–43.
11 Mateus J, Grifoni A, Tarke A, Sidney J, Ramirez SI, Dan JM, et al. Selective and cross-reactive SARS-CoV-2 T cell epitopes in unexposed humans. *Science* 2020;**370**(6512):89–94.
12 Grifoni A, Voic H, Dhanda SK, Kidd CK, Brien JD, Buus S, et al. T Cell responses induced by attenuated flavivirus vaccination are specific and show limited cross-reactivity with other flavivirus species. *J Virol* 2020;**94**(10). e00089-20.
13 Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990;**215**(3):403–10.
14 Zhang Y, Skolnick J. TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res* 2005;**33**(7):2302–9.
15 Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature* 2021;**596**(7873):583–9.
16 Hubbard SJ. NACCESS: A Program for Calculating Accessibilities. Department of Biochemistry and Molecular Biology; 1993.
17 Grimaud M, Starck J, Levy M, Marais C, Chareyre J, Khraiche D, et al. Acute myocarditis and multisystem inflammatory emerging disease following SARS-CoV-2 infection in critically ill children. *Ann Intensive Care* 2020;**10**(1):69.
18 Bozkurt B, Kamat I, Hotez PJ. Myocarditis with COVID-19 mRNA vaccines. *Circulation* 2021;**144**(6):471–84.
19 Tuohy VK, Kinkel RP. Epitope spreading: a mechanism for progression of autoimmune disease. In: Górski A, Krotkiewski H, Zimecki M, eds. Autoimmunity [Internet], Dordrecht: Springer Netherlands; 2001:39–48. https://doi.org/10.1007/978-94-010-0981-2_4.