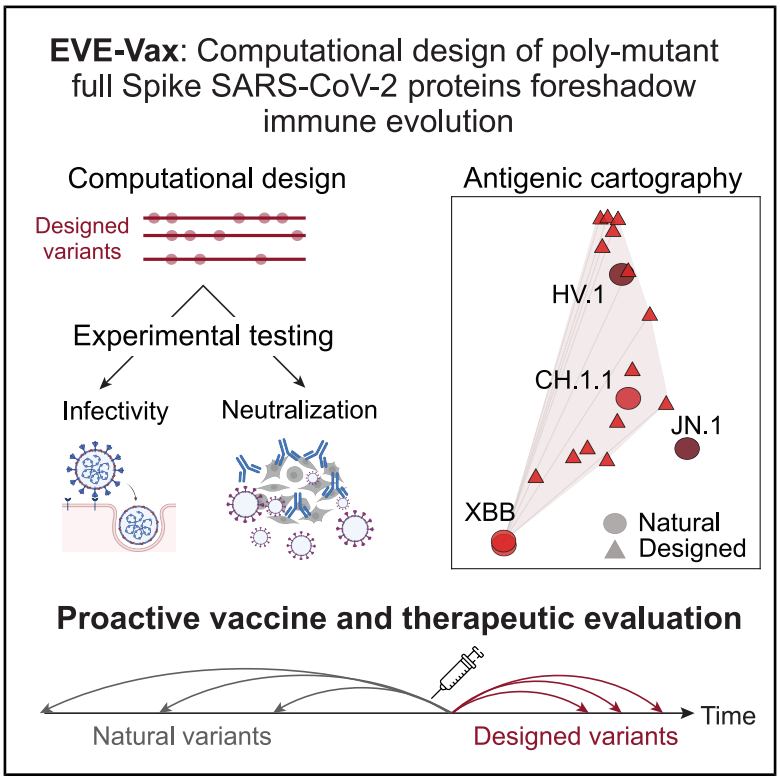
Article

Immunity

Computationally designed proteins mimic antibody immune evasion in viral evolution

Graphical abstract



Highlights

- Developed EVE-Vax, a computational method for designing antigens
- Designed and tested 83 poly-mutant SARS-CoV-2 spike proteins
- EVE-Vax designs foreshadowed immune escape observed in natural pandemic variants
- EVE-Vax designs offer an approach for proactive vaccine and therapeutic evaluation

Authors

Noor Youssef, Sarah Gurev, Fadi Ghantous, ..., Jeremy Luban, Michael S. Seaman, Debora S. Marks

Correspondence

lemieux@broadinstitute.org (J.E.L.), jeremy.luban@umassmed.edu (J.L.), mseaman@bidmc.harvard.edu (M.S.S.), debbie@hms.harvard.edu (D.S.M.)

In brief

Rapid viral evolution challenges the efficacy of vaccines and other medical interventions. Our ability to evaluate interventions is currently limited to assessments against past or circulating variants. Youssef et al. developed EVE-Vax, a computational method for designing poly-mutant proteins that foreshadow immune evasion seen in naturally evolving viral variants. Their findings offer an approach for proactive vaccine and therapeutic evaluations.

Youssef et al., 2025, Immunity *58*, 1–11
June 10, 2025 © 2025 The Authors. Published by Elsevier Inc.
https://doi.org/10.1016/j.immuni.2025.04.015

CellPress

Please cite this article in press as: Youssef et al., Computationally designed proteins mimic antibody immune evasion in viral evolution, Immunity (2025), https://doi.org/10.1016/j.immuni.2025.04.015

Immunity

CellPress
OPEN ACCESS

Article

Computationally designed proteins mimic antibody immune evasion in viral evolution

Noor Youssef,[1,2] Sarah Gurev,[1,2,3] Fadi Ghantous,[4] Kelly P. Brock,[1,2,14] Javier A. Jaimes,[5] Nicole N. Thadani,[1,2,15] Ann Dauphin,[5] Amy C. Sherman,[6] Leonid Yurkovetskiy,[5] Daria Soto,[5] Ralph Estanboulieh,[1] Ben Kotzen,[2,7] Pascal Notin,[1,2] Aaron W. Kollasch,[1,2] Alexander A. Cohen,[8] Sandra E. Dross,[9,10] Jesse Erasmus,[11] Deborah H. Fuller,[9,10] Pamela J. Bjorkman,[8] Jacob E. Lemieux,[2,7,12,*] Jeremy Luban,[2,5,12,13,*] Michael S. Seaman,[4,*] and Debora S. Marks[1,2,16,*]

[1]Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA
[2]Broad Institute of Harvard and MIT, Cambridge, MA 02139, USA
[3]Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA 02139, USA
[4]Center for Virology and Vaccine Research Beth Israel Deaconess Medical Center, Boston, MA 02215, USA
[5]Program in Molecular Medicine, University of Massachusetts Chan Medical School, Worcester, MA 01655, USA
[6]Division of Infectious Diseases, Department of Medicine, Brigham and Women's Hospital, Boston, MA 02115, USA
[7]Massachusetts General Hospital, Boston, MA 02114, USA
[8]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA
[9]Department of Microbiology, University of Washington, Seattle, WA 98195, USA
[10]National Primate Research Center, Seattle, WA 98109, USA
[11]HDT Bio, Seattle, WA 98109, USA
[12]Massachusetts Consortium on Pathogen Readiness, Boston, MA 02115, USA
[13]Ragon Institute of MGH, MIT, and Harvard, Cambridge, MA 02139, USA
[14]Present address: Kernal Biologics, Cambridge, MA 02142, USA
[15]Present address: Apriori Bio, Boston, MA 02142, USA
[16]Lead contact
*Correspondence: lemieux@broadinstitute.org (J.E.L.), jeremy.luban@umassmed.edu (J.L.), mseaman@bidmc.harvard.edu (M.S.S.), debbie@hms.harvard.edu (D.S.M.)
https://doi.org/10.1016/j.immuni.2025.04.015

SUMMARY

Recurrent waves of viral infection necessitate vaccines and therapeutics that remain effective against emerging viruses. Our ability to evaluate interventions is currently limited to assessments against past or circulating variants, which likely differ in their immune escape potential compared with future variants. To address this, we developed EVE-Vax, a computational method for designing antigens that foreshadow immune escape observed in future viral variants. We designed 83 SARS-CoV-2 spike proteins that transduced ACE2-positive cells and displayed neutralization resistance comparable to variants that emerged up to 12 months later in the COVID-19 pandemic. Designed spikes foretold antibody escape from B.1-BA.4/5 bivalent booster sera seen in later variants. The designed constructs also highlighted the increased neutralization breadth elicited by nanoparticle-based, compared with mRNA-based, boosters in non-human primates. Our approach offers targeted panels of synthetic proteins that map the immune landscape for early vaccine and therapeutic evaluation against future viral strains.

INTRODUCTION

The emergence of viral variants that evade protective immunity induced by prior infections, vaccines, and therapeutics is a challenge for the control of viral spread. This is exemplified by recurrent breakthrough infections observed with SARS-CoV-2 and influenza. Although SARS-CoV-2 vaccines and therapeutics have mitigated the severity of COVID-19, their efficacy has been progressively undermined by the evolution of new variants. For instance, the U.S. food and drug adminstration (FDA) emergency use approval for the majority of monoclonal antibody therapies has been revoked based on their loss of efficacy against emerging variants of concern (VOCs).[1] Similarly, first-generation vaccines and boosters exhibit reduced protection against more recent variants,[2,3] necessitating annual updates to vaccine formulations encoding the latest VOC circulating at the time of vaccine approval. The selected immunogen strain often differs from the dominant variant at the time the vaccine is administered,[4] leading to lower neutralizing antibody titers.[5] These limitations highlight the need for strategies that enable proactive, rather than reactive, responses to viral evolution.

Current preclinical evaluations of vaccine and therapeutic efficacy focus primarily on neutralization potency against previous or circulating viral strains or against related viruses, e.g., SARS-like

This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Please cite this article in press as: Youssef et al., Computationally designed proteins mimic antibody immune evasion in viral evolution, Immunity (2025), https://doi.org/10.1016/j.immuni.2025.04.015

betacoronaviruses (sarbecoviruses).[6–11] However, neutralizing antibodies that are broadly protective across sarbecoviruses or against circulating SARS-CoV-2 strains do not guarantee protection against emerging variants.[12] Such evaluations therefore do not address the question of how effectively interventions will protect against future viral evolution. By contrast, predictive frameworks that anticipate immune-evasive mutations could enable the design of viral proteins to proactively assess the breadth and potency of vaccine-elicited and therapeutic antibodies.

Both experimental and computational methods have identified future escape mutations from monoclonal antibodies and polyclonal sera.[2,3,6,11,13–30] Experimental approaches often utilize deep mutational scanning (DMS) to measure the effect of mutations on neutralization by patient antibodies or sera. However, DMS is typically limited to a subdomain of the antigen, measures the impact of single amino acid mutations or a limited number of combinations of mutations, and relies on patient sera that may not be available early in an outbreak.[2,3,6,11,13–29] Computational models offer a promising alternative because they bypass these limitations, could be generalizable across viral families, and may be available prior to a viral outbreak. For example, EVEscape, a computational deep learning model, accurately predicted immune-evading mutations across SARS-CoV-2, influenza, HIV, and Lassa virus.[30] However, although experimental approaches have been used to generate antigens with novel combinations of mutations that evade neutralization, it has yet to be shown that computational methods can similarly generate functional antigens that foreshadow immune escape.

In this work, we address this gap by computationally generating and experimentally testing SARS-CoV-2 spike proteins with novel combinations of mutations that are representative of future antigenic evolution using the EVE-Vax design pipeline. We generated 83 multi-mutant full-length spike constructs across 5 VOC backgrounds (B.1, BA.4/5, BA.2.12.1, BA.2.75, and XBB), which were engineered as single-cycle infection pseudotypes. We assessed neutralization susceptibility against polyclonal immune sera from individuals with diverse exposure histories throughout the COVID-19 pandemic. EVE-Vax-designed spikes replicated immune-escape profiles observed in emerging VOCs demonstrating its potential for mapping the immune landscape and guiding the evaluation and design of medical interventions that are robust to viral evolution.

RESULTS

A limitation in current vaccine and therapeutic evaluation approaches is that elicited protection is assessed against previous or currently circulating variants of a virus, which does not necessarily reflect efficacy against future viral evolution.[25] To address this challenge, we developed EVE-Vax, a computational pipeline for generating panels of antigenically diverse proteins, enabling the proactive evaluation of the breadth and future efficacy of vaccines and therapeutics.

EVE-Vax uses the EVEscape[30] framework to score the probability of antibody escape by combining 3 biologically relevant constraints: (1) impact on fitness, (2) accessibility to antibodies, and (3) disruption potential on antibody binding. Briefly, EVE-Vax first combines high-scoring single mutants to evaluate all possible double mutants, which are then further combined to generate multi-mutant full-length spike constructs (Figures 1A and S1; STAR Methods).

We designed 83 multi-mutant spike proteins on 5 different VOC backgrounds: B.1 (wild type variant with D614G mutation), BA.4/5, BA.2.12.1, BA.2.75, and XBB (Figures 1B and S1; Table S1). Constructs contained up to 10 novel combinations of mutations relative to the background VOC and up to 46 mutations relative to the ancestral B.1 strain (Table S2). In total, 37 unique mutations, across 30 positions, were introduced in different combinations and on different spike backgrounds (Figure 1C). Most mutations were located in primary antigenic regions despite no explicit inclusion of this constraint in the EVE-Vax pipeline—17 (57%) of the mutated residues were in the receptor binding domain (RBD) and 12 (40%) were in the N-terminal domain (NTD).

Computationally designed multi-mutant spike proteins retain infectivity *in vitro*

Generating multi-mutant proteins is challenging as functionality often decreases with increasing mutational load. For example, in experiments that measured RBD expression using error-prone PCR libraries,[31] ~47% of all single mutations were expressed in contrast to less than 2% of those with 8 or more mutations (Figure 2A). The reduction in functional sequences with increasing mutational depth is generally observed across viral (Lassa glycoprotein complex, GPC[32]) and non-viral (*Clytia gregaria* green fluorescent protein, GFP[33]) proteins (Figure 2A). Furthermore, full spike proteins containing combinations of commonly observed pandemic mutations, rather than introducing mutations at random using error-prone PCR, resulted in less than 30% of proteins with 8 or more mutations maintaining infectivity.[29]

Among the 83 designed constructs, 90% (75/83) were infectious using replication-incompetent lentiviral-based pseudoviruses in single-round biosafety-level-2 infection assays[34] (Figures 2A, 2B, and S2; Table S3). The 10% of our designs that were not infectious may be explained post hoc by two observations: (1) 4 of the 8 non-infectious designs contained a triplet of mutations (L452R, F490R, and Q493S) that were closer in the three-dimensional structure than any triplets seen in the pandemic (Figure S2), and (2) the remaining 4 constructs were designed using a model trained exclusively on pre-pandemic sequences and included a mutation (R403T) that likely disrupted ACE2 binding.[35] Together, these results highlight that the 90% success rate observed in our computationally designed proteins exceeds the expected rates for randomly introduced mutations or combinations of individually non-deleterious mutations and provide insights for refining the EVE-Vax design algorithm.

Spikes designed on early SARS-CoV-2 variants exhibit neutralization resistance similar to subsequent variants

To evaluate the relevance of the EVE-Vax-designed spike proteins as proxies for future viral variation, we performed pseudovirus neutralization assays using nine panels of human polyclonal serum pools representing diverse SARS-CoV-2 exposure histories: convalescent sera from unvaccinated individuals infected with (1) B.1 or (2) Delta infections; (3) recipients of two primary doses of mRNA vaccines (mRNA-1273 or BNT162b2); (4) recipients of three doses of an mRNA vaccine; (5) vaccinated

Please cite this article in press as: Youssef et al., Computationally designed proteins mimic antibody immune evasion in viral evolution, Immunity (2025), https://doi.org/10.1016/j.immuni.2025.04.015

CellPress
OPEN ACCESS

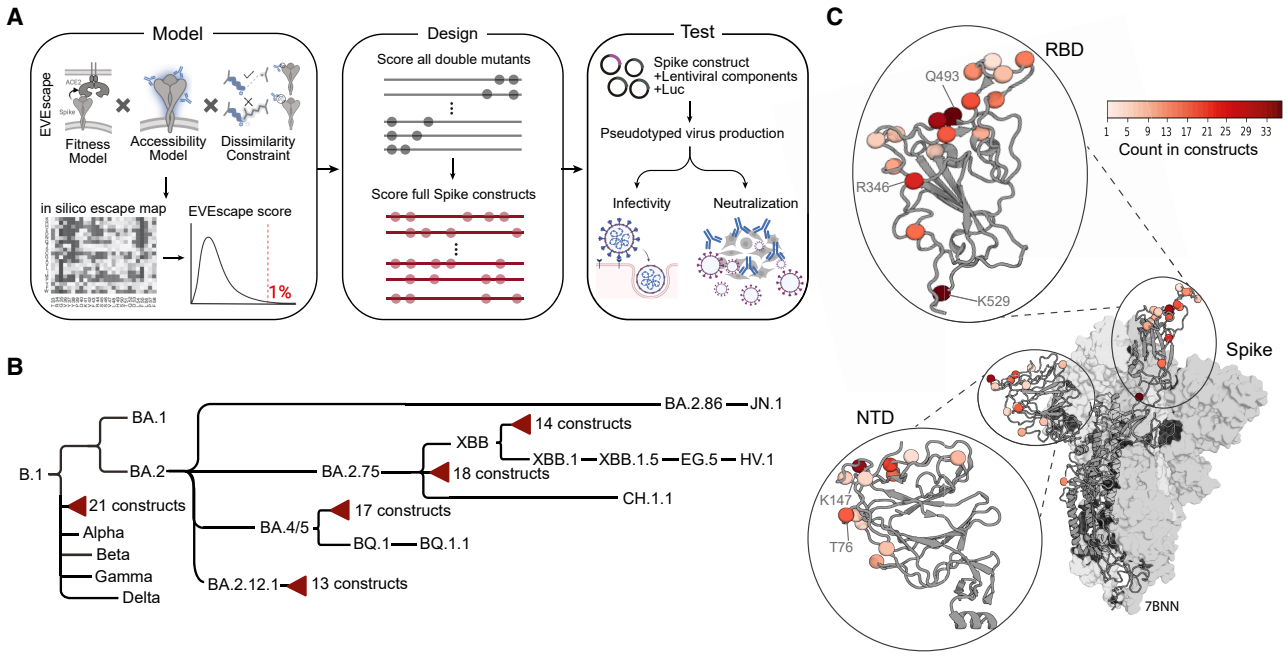


Figure 1. Computationally designed spike proteins allow for a proactive assessment of vaccines and therapeutics

(A) Schematic overview of EVE-Vax for designing antigenic proteins. Single mutants within the top 1% of highest-predicted escape scores were combined to generate all possible double mutants. Double mutants were scored and further combined to create multi-mutant constructs. Designed constructs were subsequently evaluated for infectivity and neutralization sensitivity using pseudotyped virus assays. Parts of the figure were created with BioRender.

(B) Cladogram depicting VOCs and computationally designed constructs (red triangles). Branch lengths are proportional to the temporal order of variant emergence.

(C) Mutations across the 83 designed spike constructs mapped onto a representative 3D structure (PDB: 7BNN). Coloring indicates the frequency with which a given residue was mutated across all designed constructs. See also Figure S1 and Tables S1 and S2.

individuals boosted with BA.4/5 bivalent mRNA vaccines; and vaccinated patients who experienced breakthrough infections (BTIs) with (6) Delta, (7) BA.1, (8) BA.2.12.1, or (9) BA.4/5 variants (Figure 2C; Table S4). Neutralization assays were conducted on pseudoviruses expressing 66 of the designed spike constructs, 20 SARS-CoV-2 variants (B.1, alpha, beta, delta, gamma, BA.1, BA.2, BA.2.12.1, BA.2.75, BA.4/5, BQ.1, BQ.1.1, XBB, XBB.1, XBB.1.5, CH.1.1, EG.5, HV.1, BA.2.86, and JN.1), and the SARS-CoV-1 spike protein (Figure 2D; Table S4).

To characterize the degree of immune evasion observed throughout the pandemic, we compared the half-maximal neutralizing antibody titers (ID_{50}) for each variant relative to the parent variant from which it evolved (Figure 2E). As expected, most variants were less susceptible to neutralization (had higher antibody escape) compared with the parental variant that preceded it, with the exception of BQ.1.1, XBB.1, and XBB.1.5 variants, which retained similar titers to their parent variants, BQ.1, XBB, and XBB.1, respectively ($p > 0.2$, Wilcoxon rank-sum test). The XBB (7.2-fold, $p < 0.01$, Wilcoxon rank-sum test) and CH.1.1 (14.2-fold, $p < 0.01$, Wilcoxon rank-sum test) variants displayed the highest neutralization resistance relative to their parent variant BA.2.75. On average, emerging variants exhibited a 3.9-fold reduction in geometric mean ID_{50} titers relative to their parent variant and a range from 0.67 to 14.2 (gray region, Figure 2E). Variants with higher antibody escape generally exhibited reduced infectivity compared with their parent variants (Figure 2B).

Pseudoviruses expressing designed spikes had comparable neutralization titers as SARS-CoV-2 variants that evolved throughout the pandemic. On average, EVE-vax-designed spikes had a 1.9-fold reduction (range of 0.5 to 5.31) in geometric mean ID_{50} titer compared with their relative parent variant (Figure 2E). More specifically, designs on specific backgrounds had similar neutralization resistance as SARS-CoV-2 variants evolved from those same backgrounds. For instance, the B.1-4a-designed construct (B.1 + K147N + S494R + F490R + R683N) had a 3.9-fold reduction in geometric mean ID_{50} titer, exceeding the neutralization resistance of Alpha, Delta, and Gamma (1.5-, 1.7-, and 2.0-fold reduction relative to B.1, respectively) and comparable to Beta (4.8-fold reduction relative to B.1). Similarly, the BA.4/5-2a construct (BA.4/5 + R346T + S494R) exhibited a 1.9-fold reduction in geometric mean ID_{50} titers relative to BA.4/5, comparable to BQ.1 (2.4-fold reduction). The BA.2.75-4c design (BA.2.75 + G339D + L452R + Q493R + K529L) showed a 5.3-fold reduction in geometric mean ID_{50} titer relative to BA.2.75, below the relative neutralization resistance of XBB (7.2-fold). Neutralization profiles of EVE-vax designed constructs on the XBB background remain consistent with the most recent SARS-CoV-2 variants (JN.1.7.1, XDV, KQ.1, KP.1.1, KP.2, KP.1.1.1, KP.3, KP.3.1.1, and LB.1), which emerged nearly two years after the designs (Figure S3). Together, these results demonstrate that EVE-Vax designed constructs exhibit a similar antibody escape capacity as variants that have naturally evolved under immune

Please cite this article in press as: Youssef et al., Computationally designed proteins mimic antibody immune evasion in viral evolution, Immunity (2025), https://doi.org/10.1016/j.immuni.2025.04.015

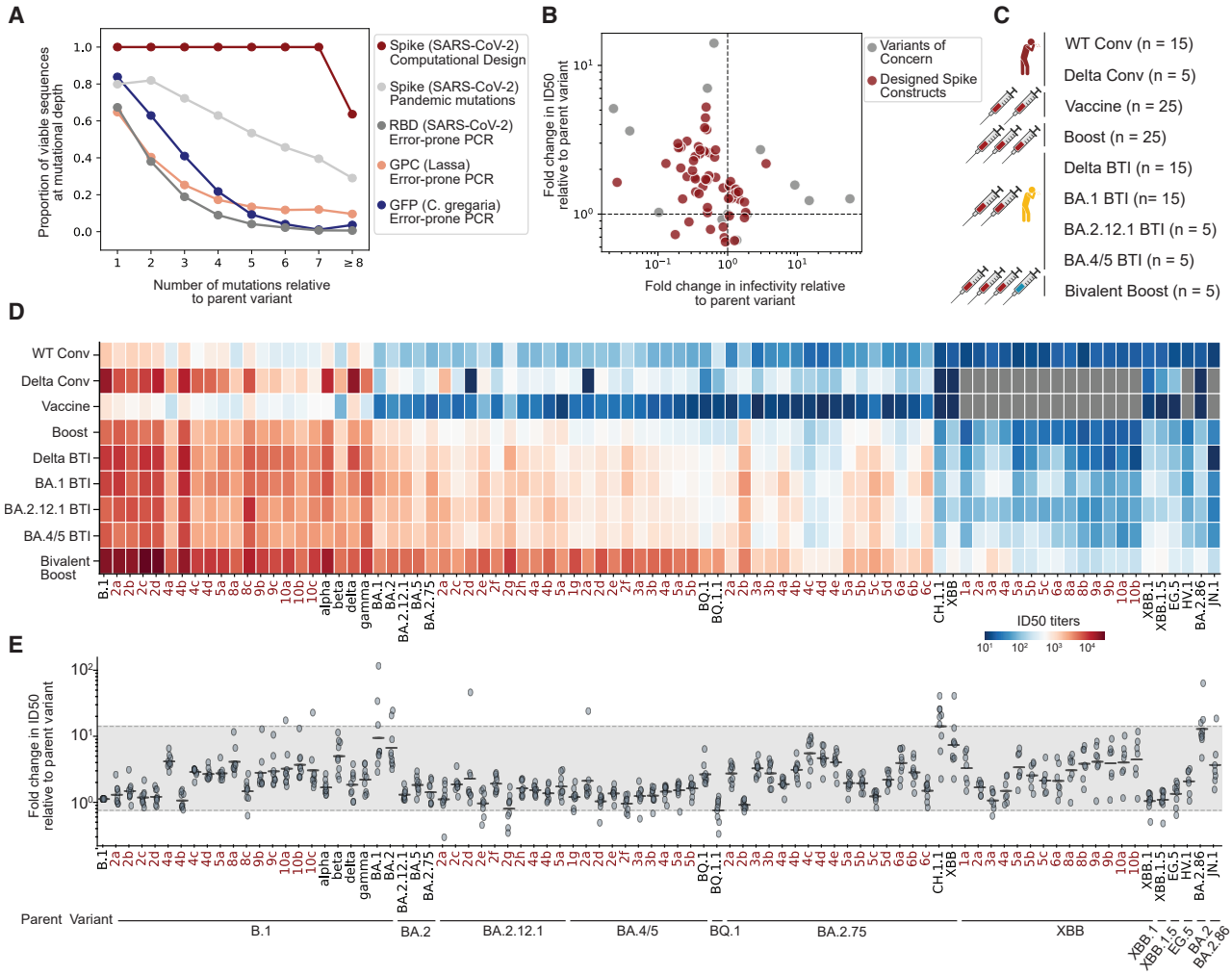


Figure 2. EVE-Vax-designed constructs demonstrate neutralization escape comparable to VOCs

(A) Proportion of viable sequences across mutational depths. EVE-Vax-designed constructs exhibit higher viability compared with sequences generated by error-prone PCR within the SARS-CoV-2 RBD (assayed for expression[31]), the Lassa glycoprotein complex (GPC) (assayed for cell entry),[32] the *Clytia gregaria* green fluorescent protein (GFP, assayed for fluorescence activity[33]), or from combinations of frequently observed pandemic mutations in spike (assayed for infectivity[29]).

(B) Relationship between infectivity and neutralization. EVE-Vax-designed constructs and VOCs with higher escape from sera typically showed reduced infectivity relative to their parent variant.

(C) Summary of 9 serum panels used for neutralization assays. Panels were from individuals who were convalescent (Conv.), vaccinated, boosted, or experienced breakthrough infection (BTI). A total of 23 serum pools were created, each pooling samples from 5 individuals.

(D) Neutralizing ID_{50} titers for VOCs and designed constructs across 9 serum panels. Reported are the geometric mean titers across serum pools within a panel. Gray boxes indicate variants that were not tested in that serum panel. EVE-Vax constructs are highlighted in red text.

(E) Fold change in neutralizing ID_{50} titers relative to the parent variant. Points represent the geometric mean fold change for each serum panel, with lines indicating the overall geometric mean across all 9 panels. The gray band represents the range of neutralization sensitivity observed among VOCs. See also Figure S2 and Tables S3 and S4.

pressure and can therefore serve as useful proxies for future SARS-CoV-2 evolution.

Designed variants foreshadow antigenic evolution

To investigate whether EVE-Vax designed constructs recapitulate the antigenic evolution observed during the pandemic, we constructed antigenic maps where the distance between any pair of variants reflects the similarity in their neutralization profiles across diverse serum pools (Figures 3A and S3). Constructs designed on earlier variants exhibited an antigenic resemblance to variants that emerged later in the pandemic (Figure 3A). The BA.2.12.1-5a designed construct mimics the neutralizability of later pandemic variant BA.2.75 with both variants mutating residue K147 (N in the design, E in the later variant). The functional impact of K147 mutations can be seen for the NTD antibody S2X303, whose epitope partially overlaps the NTD antigenic supersite[36] (Figure 3B). XBB designs that contained L452R (XBB-8a,9a,9b,10a,10b) or S494R (XBB-1a) closely matched

Please cite this article in press as: Youssef et al., Computationally designed proteins mimic antibody immune evasion in viral evolution, Immunity (2025), https://doi.org/10.1016/j.immuni.2025.04.015

CellPress
OPEN ACCESS

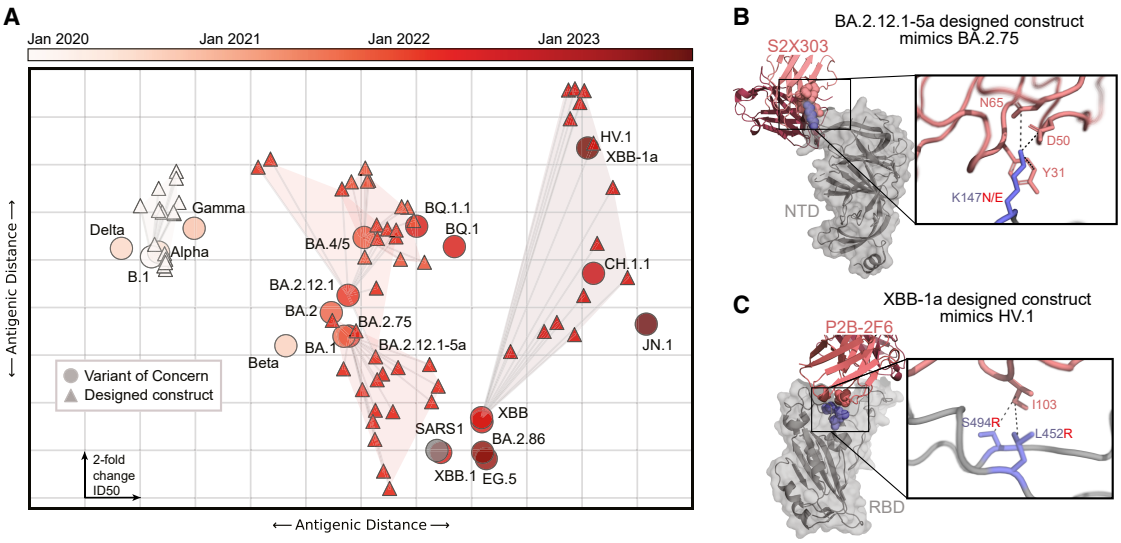


Figure 3. EVE-Vax-designed constructs foreshadow future SARS-CoV-2 immune-escape evolution

(A) Antigenic cartography depicts a two-dimensional representation of the antigenic distance based on neutralization by 23 serum pools. Each grid box corresponds to a 2-fold change in neutralizing ID_{50} titers. Variants (circles) are colored according to their month of emergence since the onset of the pandemic. Designed constructs (triangles) on earlier variants foreshadow the antigenicity of later pandemic variants, as indicated by proximity in the antigenic map.

(B) Designed construct BA.2.12.1-5a mimics the neutralizability of later pandemic variant BA.2.75. Both contain a mutation at K147 (N in the design, E in the later variant). The impact of K147 mutations can be seen for NTD antibody S2X303, with interactions to N65, D50, and Y31 on the antibody (PDB: 7SOF).

(C) Many XBB designs containing L452R (XBB-8a,9a,9b,10a,10b) or S494R (XBB-1a) closely resemble the neutralization profile of the later HV.1 variant, which also harbors L452R. Arginine mutations in either S494 or L452 result in escape from RBD class 2 antibody P2B-2F6[2,37] (PDB: 8DCC). See also Figure S3 and Table S4.

the neutralization profile of HV.1, which also contained L452R, with notable escape from BA.1 BTI sera compared with other serum panels (Table S4). The impacts of these mutations can be seen in RBD class 2 antibody P2B-2F6,[37] where mutation to an arginine of either S494 or L452 (both contacting I103 on the antibody) results in escape[2] (Figure 3C). These observations align with the hypothesis that later Omicron subvariants evolved L452R to enhance immune evasion following BA.1 infection.[3] These trends were consistent across other rounds of design: constructs designed on BA.2.75 were antigenically similar to the XBB lineages that emerged nine months later, constructs designed on the XBB variant exhibited antigenic responses akin to the CH.1.1 and HV.1 variants that emerged three and 12 months later, constructs based on BA.4/5 anticipated BQ.1.1, and designs on B.1 presaged the alpha and gamma variants (Figure 3A, shaded area). These findings underscore the capability of EVE-Vax designs to recapitulate antigenic profiles akin to those of future variants using only data available at the time of VOC emergence.

Evaluating neutralizing antibody responses elicited by mRNA vaccines

To demonstrate the utility of our approach in the proactive evaluation of vaccine efficacy, we retrospectively used the designed constructs to evaluate the B.1-BA.4/5 bivalent booster vaccine. Considering variants circulating in the four months preceding the bivalent booster campaign (BA.2.75, BQ.1, BQ.1.1, and XBB), the geometric mean ID_{50} titers were 1,931 for serum panels representative of bivalent boosting (four shots; Figure 4A; Table S1 and S4). The high titers are indicative of adequate protection against these prior variants. However, constructs designed on BA.2.75 and XBB demonstrated a range of antibody escape, with ID_{50} titers ranging from 193 to 4,029. The broad range of titers, specifically the low titers, were reflective of the escape potential of variants that evolved post booster vaccine implementation—the subsequent variants, XBB.1, XBB.1.5, and CH.1.1, had ID_{50} titers of 538, 587, and 308, respectively (Figure 4A). Moreover, 8 XBB-based designs displayed lower ID_{50} titers than CH.1.1, demonstrating the potential for further escape from neutralizing antibodies elicited by bivalent booster vaccines. These results highlight the utility of EVE-Vax designs for early vaccine evaluation.

Evaluating neutralizing antibody responses elicited by nanoparticle vaccines

Cohen et al.[8,9] recently developed two RBD nanoparticle vaccines: mosaic-8b nanoparticles displaying RBDs from eight different sarbecoviruses and a homotypic nanoparticle displaying the RBD from the Beta SARS-CoV-2 variant. Both nanoparticle vaccines elicited neutralizing antibodies against SARS-CoV-2 variants when tested in mouse and non-human primate (NHP) animal models.[8,9] Fourteen NHPs were primed with a mixture of nucleotide vaccines and boosted with either a bivalent mRNA vaccine (n = 4), a mosaic-8b nanoparticle (n = 5), or a homotypic nanoparticle (n = 5) vaccine.[38] To evaluate the neutralizing potential of these booster vaccines against future SARS-CoV-2 evolution, we examined differences in neutralization activity across five SARS-CoV-2 variants (B.1, BA.2.12.1, BA.2.75, BA.4/5, and XBB), as well as four designed spike constructs (BA.5-2a, BA.2.12.1-2d, BA.2.75-4c, and XBB-10b; Table S4).

Please cite this article in press as: Youssef et al., Computationally designed proteins mimic antibody immune evasion in viral evolution, Immunity (2025), https://doi.org/10.1016/j.immuni.2025.04.015

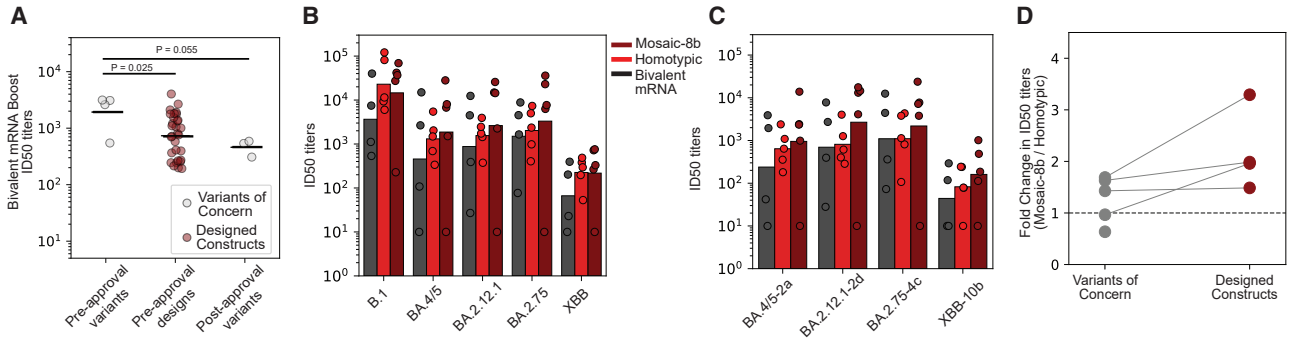


Figure 4. Evaluating the breadth and potency of vaccines using designed spike constructs

(A) Neutralizing ID_{50} titers against sera from individuals who received a BA.4/5 bivalent booster vaccination after 3 doses of an mRNA vaccine. Lines indicate geometric mean across variants or designs. The BA.4/5 bivalent booster vaccine was evaluated on its ability to neutralize variants circulating in the 4 months prior to the creation of the booster vaccines (BA.2.75, BQ.1, BQ.1.1, and XBB). Designed spike constructs on the background of variants circulating prior to the vaccine foreshadowed the antibody escape observed against variants that emerged following the bivalent booster approval (XBB.1, XBB.1.5, and CH.1.1). Statistical significance was assessed using the Wilcoxon rank-sum test.

(B) Neutralizing potency of sera collected from non-human primates primed with a mixture of nucleic acid vaccines and boosted with either a bivalent mRNA (n = 4), a homotypic nanoparticle (n = 5), or a mosaic 8b nanoparticle (n = 5). Bar charts represent the geometric mean ID_{50} titers for each variant.

(C) Bar charts represent the geometric mean ID_{50} titers for each designed construct. Designed constructs highlight the improvement of protection from antibodies elicited by the mosaic-8b booster compared with the homotypic nanoparticle.

(D) Neutralization sensitivity of a given variant by the mosaic-8b boost sera relative to the homotypic boost sera. Higher values indicate the mosaic-8b vaccine better neutralizes a given variant. Lines connect designed constructs to their parent variants. See also Table S4.

Sera from bivalent mRNA-boosted NHPs demonstrated the lowest neutralization titers against the tested SARS-CoV-2 variants (mean titer of 678) compared with sera from NHPs, which were boosted with the homotypic nanoparticle (1,843, P = 0.166 Wilcoxon rank-sum test) or the mosaic-8b nanoparticle (2,199, p = 0.060, Wilcoxon rank-sum test, Figure 4B). Neutralizing titers were consistently lower against the EVE-Vax designed constructs compared with the SARS-CoV-2 variants, with on average a 2.3, 4.0, and 2.3-fold decrease in the bivalent mRNA, homotypic, and mosaic-8b sera, respectively (Figure 4C). Although the neutralization titers among the natural SARS-CoV-2 variants were comparable between the mosaic-8b and homotypic-boosted NHPs, the designed constructs showcase the improved performance of the mosaic-8b nanoparticle booster vaccine (Figure 4D). Neutralizing titers against the designed constructs BA.2.12.1-2d, BA.2.75-4c, BA.5-2a, and XBB-10b were 3.3, 2.0, 1.5, and 2.0-fold higher in sera from mosaic-8b boosted NHPs compared with homotypic boosted animals. Overall, the results suggest that nanoparticle booster vaccines lead to higher neutralizing titers against future SARS-CoV-2 variants as compared with bivalent mRNA booster vaccines and that the mosaic-8b elicits more cross-reactive antibodies compared with the homotypic nanoparticle.

EVE-Vax designs have comparable escape to assay-derived designs

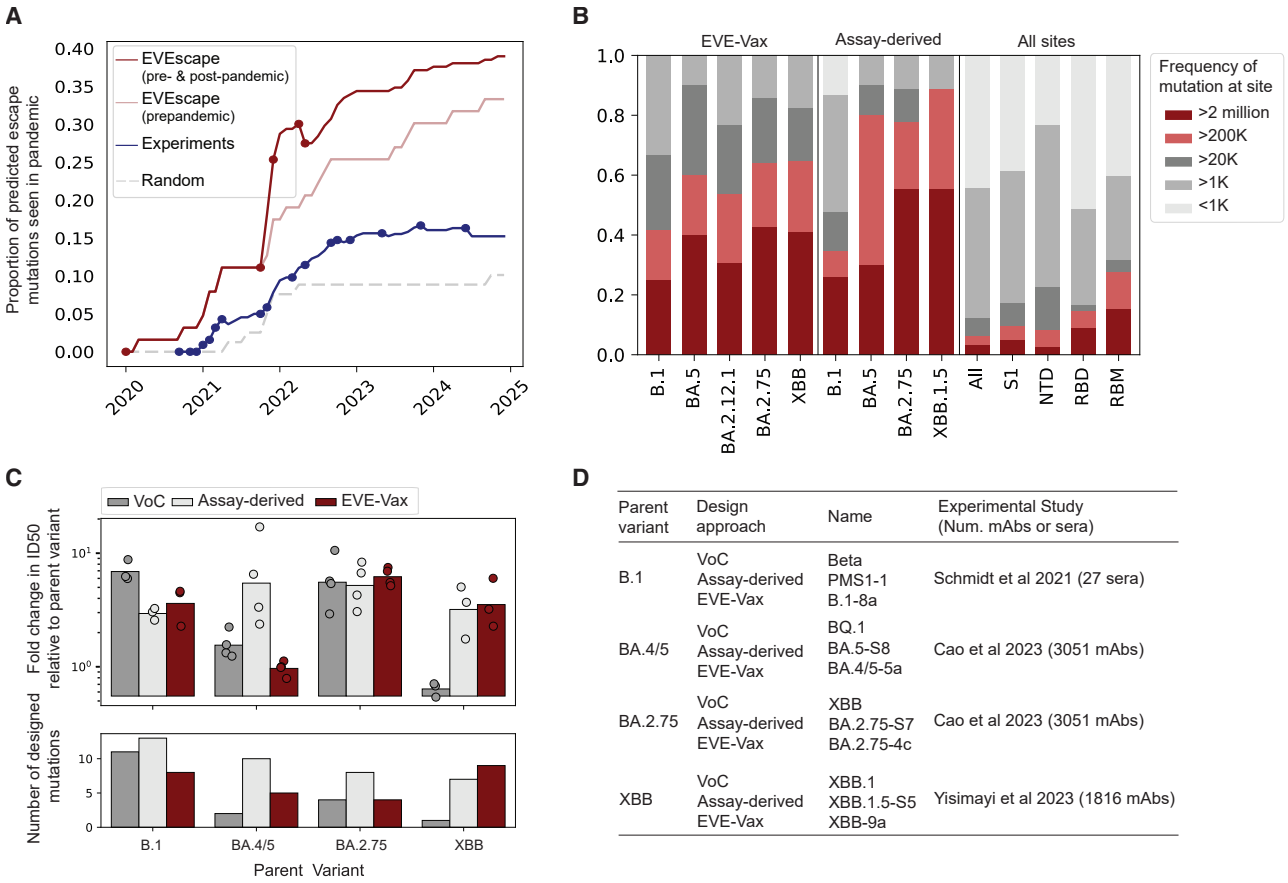
Lastly, we compared the ability of computational and experimental methods to forecast pandemic mutations and generate constructs that foreshadow immune escape. Prior work demonstrated that EVEscape, trained on data available prior to the onset of the COVID-19 pandemic, more accurately predicted escape mutations observed during the pandemic than the earliest mutational scanning experiments.[30] To date, 33% of escape mutations predicted by *pre-pandemic* EVEscape (mutations in the top 5% of scores) were observed in over 1,000 viral strains during the pandemic. In contrast, only 15% of escape mutations predicted by early mutational scanning experiments met this threshold (Figures 5A and S4; Table S5). Updating the computational models with SARS-CoV-2 sequences available through July 2022 (models developed in this study) increased the proportion of observed escape mutations to ~40% (85 of the 218 computationally predicted escape mutations appeared in more than 1,000 sequences). By comparison, updating experimental datasets to reflect contemporary immune pressures by including escape mutations from 20 high-throughput experimental studies[2,3,6,11,13–29] resulted in ~15% of escape mutations being observed during the pandemic (64 of the 420 experimentally predicted escape mutations).

We compare the ability of EVE-Vax constructs versus "assay-derived" constructs to foreshadow immune escape (assay-derived constructs refer to previously published constructs that were developed using experimental methods). We compared our B.1 designs to PMS1-1 and PMSD4 from Schmidt et al.,[26] which included 13 mutations in each construct selected based on plasma neutralization sensitivity and distribution across the spike protein (Figure S5). For BA.2.75 and BA.4/5 backgrounds, we compared our constructs to designs from Cao et al.,[22] where RBD mutations were selected based on criteria derived from deep mutational scans assaying neutralization resistance against over three thousand BA.2-neutralizing antibodies (Figures S5). We compare our XBB designs to the closest available experimental designs from Yisimayi et al.,[23] who designed XBB.1.5 RBD constructs using the same criteria as Cao et al.[22] but measured neutralizing activity across 1,816 RBD-targeting antibodies (Figure S5).

Both computational and experimental approaches identified positions that were frequently mutated during the pandemic at greater rates than random selection of mutations in

Please cite this article in press as: Youssef et al., Computationally designed proteins mimic antibody immune evasion in viral evolution, Immunity (2025), https://doi.org/10.1016/j.immuni.2025.04.015

CellPress
OPEN ACCESS



Parent variant	Design approach	Name	Experimental Study (Num. mAbs or sera)
B.1	VoC Assay-derived EVE-Vax	Beta PMS1-1 B.1-8a	Schmidt et al 2021 (27 sera)
BA.4/5	VoC Assay-derived EVE-Vax	BQ.1 BA.5-S8 BA.4/5-5a	Cao et al 2023 (3051 mAbs)
BA.2.75	VoC Assay-derived EVE-Vax	XBB BA.2.75-S7 BA.2.75-4c	Cao et al 2023 (3051 mAbs)
XBB	VoC Assay-derived EVE-Vax	XBB.1 XBB.1.5-S5 XBB-9a	Yisimayi et al 2023 (1816 mAbs)

Figure 5. EVE-Vax constructs have comparable polyclonal escape to constructs designed using pandemic sera and antibodies

(A) Proportion of escape mutations seen in over 1,000 strains in Global Initiative on Sharing All Influenza Data (GISAID), as predicted by EVEscape (with and without SARS-CoV-2 sequences), high-throughput DMS experiments across 20 studies (Table S5), or randomly selected RBD mutations. Predicted escape mutations were defined as nonsynonymous, single-nucleotide RBD mutations in the top 5% of escape scores for computational and experimental methods (top 1% and 10% in Figure S4). Points indicate updates to the set of escape mutations based on the availability of a more recent DMS data (blue) or model (red, which includes SARS-CoV-2 sequences available prior to that date). Predictions for the full spike protein (computational) and for each independent experiment and model are provided in Figure S4.

(B) Proportion of designed sites mutated at different frequencies in GISAID, categorized by computational or experimental design rounds or by spike regions. Designed mutations are enriched at high-frequency sites compared with immunodominant spike regions (e.g., RBD or NTD). Spike regions analyzed: all (1–1,273), S1 (1–686), NTD (13–305), RBD (319–541), and RBM (437–508).

(C) Fold change in ID_{50} titers of assay-derived constructs, EVE-Vax constructs, or descendant VOC relative to the respective parent variant. The construct with the highest neutralization escape was selected for each round of computational and assay-derived designs. Individual points represent fold change for matched sera pool between computational and experimental studies; bars indicate geometric mean fold reduction.

(D) Table of selected EVE-Vax construct, assay-derived construct, or descendant VOC for each parent variant. For experimental and computational methods, the construct with the highest average escape per round is reported. RBD, receptor-binding domain; NTD, N-terminal domain; RBM, receptor-binding motif. See also Figures S4 and S5 and Table S5.

immunodominant regions of spike (Figures 5B and S5). Additionally, EVE-Vax constructs had comparable escape from polyclonal sera to assay-derived constructs while incorporating fewer mutations per spike protein ($p > 0.05$, Mann-Whitney rank-sum test, Bonferroni correction; Figures 5C and 5D). Furthermore, the majority of the escape mutations included in assay-derived designs were high scoring in the computational model scores (Figure S5), highlighting that these mutations could have been identified computationally if the threshold used for identifying escape mutations was lowered in the EVE-Vax pipeline. Conversely, the mutations in our EVE-Vax constructs would not have been identified from the experimental data (Figure S5). These findings indicate that EVE-Vax provides an alternative to high-throughput experimental methods for designing antigenic proteins that mimic immune-escape evolution.

DISCUSSION

Driven by the need for vaccines and therapeutics that maintain efficacy against viral evolution, hundreds of vaccine candidates are currently undergoing preclinical or clinical evaluations.[39] However, it remains unclear the extent to which these candidate vaccines will effectively combat future viral variants. To date, neutralization potential of SARS-CoV-2 booster vaccines,[40]

Please cite this article in press as: Youssef et al., Computationally designed proteins mimic antibody immune evasion in viral evolution, Immunity (2025), https://doi.org/10.1016/j.immuni.2025.04.015

pan-variant therapeutics,[6,41] and other antivirals[42] have been evaluated against previous or contemporary SARS-CoV-2 variants or against distantly related sarbecoviruses.[8,9] Such approaches are limited to retrospectively testing efficacy against emerging pandemic variants or throughout evolution, which do not guarantee effectiveness against future variants.[12] We present a vaccine and therapeutic evaluation approach that addresses these issues by designing panels of diverse antigens that can be used, in safe non-replicative pseudotype assays, to assess the breadth and potency of candidate vaccines and therapeutics to neutralize future viral variants.

We highlight the utility of this approach through analyses of the bivalent mRNA booster vaccine and pan-sarbecovirus nanoparticle-based vaccines. We demonstrate that the immune escape observed by SARS-CoV-2 variants that emerged following the bivalent booster vaccination campaign, such as XBB and CH.1.1, could have been anticipated at the time that booster vaccines were implemented. Additionally, EVE-Vax-designed constructs demonstrated the potential for further immune escape, which has since been realized by JN.1 and other variants in the BA.2.86 lineage that continue to emerge. Alternatively, nanoparticle-based vaccines offer a promising vaccine modality that has the potential to elicit broadly neutralizing antibodies. The conceptual premise posits that vaccines designed to counteract sarbecoviruses, with broad taxonomic distances, would elicit antibodies targeting conserved regions, thus decreasing the likelihood of escape under immune pressures. Supporting the advantage of the mosaic nanoparticle vaccine, we find higher neutralizing titers against our EVE-Vax constructs compared with the homotypic nanoparticle or mRNA vaccines. Taken together, these findings suggest that, when evaluating vaccines against a panel of designed variants, the range of neutralizing titers, especially the lowest observed values, may signal future escape potential.

A key advantage of the computational approach is the ability to predict escape mutations independent of antibody or sera availability, enabling proactive design efforts before the emergence of novel variants or spillover events. Furthermore, the approach is generalizable across viral families and is likely to be of relevance for mapping the immune landscape of key antigenic proteins and developing countermeasures against highly diverse or poorly characterized viruses. For example, addressing the high diversity of Lassa virus strains presents a significant challenge for vaccine development. EVE-Vax can be used to generate panels of diverse strains to assess the cross-protection potential of vaccines currently in clinical and preclinical trials.[43] By leveraging non-replicative pseudovirus assays, even biosafety-level 4 pathogens can be evaluated under safer conditions, consistent with established practices.[11,24] In addition to generating panels of synthetic proteins for vaccine evaluation, evidence presented in this paper suggests that the EVE-Vax pipeline could more directly inform vaccine design by identifying epitopes likely to escape prior vaccine-induced or therapeutic antibodies.

Limitations of the study

The extent to which the method presented here is generalizable to other viral antigens will depend on many factors—not least, whether the available evolutionary sequence record is sufficient for a computational approach to learn functional constraints. Heuristics for determining the number and diversity of sequences needed to accurately predicti functional constraints have been explored to a limited extent.[44] More quantitative methods are needed, especially those tailored to viruses, as many viruses with a high pandemic and spillover risk remain understudied and undersequenced. Although the promise of protein language models, trained on a broad set of proteins, may mitigate some of these limitations, their applicability for designing *de novo* proteins for vaccine evaluation and design remains to be explored. Another limitation is the focus in this study on antibody neutralization. Although important for initial clinical efficacy,[45] other components of the immune response, such as T-cell-mediated immunity, are often essential for long-term protection. Lastly, forecasting future variants, whether experimental or computational, could, in theory, be misused; however, these risks can be minimized by the responsible development and sharing of such information.[46]

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to, and will be fulfilled by, the lead contact, Debora Marks (debbie@hms.harvard.edu).

Materials availability

Plasmids generated in this study have been deposited to Addgene (https://www.addgene.org/Jeremy_Luban/; itemized in Table S2). Full plasmid sequences are available from the respective repositories. The remainder of the plasmids used here were previously described and are also listed with code numbers in the key resources table.

Data and code availability

All data are available in the main text or the supplementary materials. All code used for analysis in this study is publicly available on GitHub (https://github.com/debbiemarkslab/Vax_design). Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

The authors would like to thank Stephen Walsh and Lindsey Baden at Brigham and Women's Hospital (Boston, MA) for assisting in the collection of serum from vaccinated individuals for this study, Chris Sander, Andrew Hebbeler, and members of the Luban, Lemieux, Seaman and Marks labs for discussions. We thank Drs. Michael Farzan and Huihui Ma at The Scripps Research Institute for providing the HEK 293T/ACE2 cell line and Dr. Barney Graham at the NIH Vaccine Research Center for providing the pCMV-R8.2 and pHR'CMV-Luc plasmids. The cartoon diagrams in Figures 1A and 2C and the graphical abstract were created in part with Biorender.com. The authors gratefully acknowledge the POSITIVES study team for providing access to serum samples. This work was supported by grants from the Coalition for Epidemic Preparedness Innovations (CEPI) and the Massachusetts Consortium on Pathogen Readiness (MassCPR). J.L. was also supported by an NIH grant (R37 AI147868). We gratefully acknowledge all data contributors, i.e., the authors and their originating laboratories responsible for obtaining the specimens and their submitting laboratories for generating the genetic sequence and metadata and sharing via the GISAID Initiative, on which some of this research is based.

AUTHOR CONTRIBUTIONS

Conceptualization, N.Y., D.S.M., J.E.L., J.L., and M.S.S.; methodology, N. Y., S.G., and K.P.B.; investigation, N.Y., S.G., F.G., K.P.B., J.A.J., N.N.T., A.D., L.Y., D.S., R.E., B.K., P.N., and A.W.K.; resources, A.A.C., S.E.D., J.

Please cite this article in press as: Youssef et al., Computationally designed proteins mimic antibody immune evasion in viral evolution, Immunity (2025), https://doi.org/10.1016/j.immuni.2025.04.015

E., D.H.F., P.J.B., J.E.L, J.L., and M.S.S.; data curation, N.Y. and S.G.; writing – original draft, N.Y., S.G., and D.S.M.; writing – review & editing, N.Y., S.G., D.S.M. J.E.L., J.L., and M.S.S.; visualization, N.Y. and S.G.; project administration, N.Y.; funding acquisition, N.Y., S.G., D.S.M., J.E.L., J.L., and M.S.S.

DECLARATION OF INTERESTS

D.S.M. is an advisor for Dyno Therapeutics, Octant, Jura Bio, Tectonic Therapeutic, and Genentech and is a cofounder of Seismic Therapeutic.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Cell lines
  - SARS-CoV-2 convalescent and vaccine human sera
  - Sarbecovirus non-human primate vaccine sera
- METHOD DETAILS
  - Multiple Sequence Alignments for fitness models
  - EVE-Vax protein design algorithm
  - Recombinant DNA sequence plasmid design
  - Infectivity assays with SARS-CoV-2 spike-pseudotyped lentiviral particles
  - Neutralization assays with SARS-CoV-2 spike-pseudotyped lentiviral particles
  - Analysis of proportion of viable sequences at different mutational depths
  - Analysis of escape mutation frequencies in the public database
  - Comparison to assay-derived constructs designed using high throughput experiments
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.immuni.2025.04.015.

Received: May 21, 2024
Revised: October 9, 2024
Accepted: April 14, 2025
Published: May 8, 2025

REFERENCES

1. Kip, K.E., McCreary, E.K., Collins, K., Minnier, T.E., Snyder, G.M., Garrard, W., McKibben, J.C., Yealy, D.M., Seymour, C.W., Huang, D.T., et al. (2023). Evolving Real-World Effectiveness of Monoclonal Antibodies for Treatment of COVID-19: A Cohort Study. Ann. Intern. Med. *176*, 496–504. https://doi.org/10.7326/M22-1286.
2. Cao, Y., Wang, J., Jian, F., Xiao, T., Song, W., Yisimayi, A., Huang, W., Li, Q., Wang, P., An, R., et al. (2022). Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. Nature *602*, 657–663. https://doi.org/10.1038/s41586-021-04385-3.
3. Cao, Y., Yisimayi, A., Jian, F., Song, W., Xiao, T., Wang, L., Du, S., Wang, J., Li, Q., Chen, X., et al. (2022). BA.2.12.1, BA.4 and BA.5 escape antibodies elicited by Omicron infection. Nature *608*, 593–602. https://doi.org/10.1038/s41586-022-04980-y.
4. Scheaffer, S.M., Lee, D., Whitener, B., Ying, B., Wu, K., Liang, C.-Y., Jani, H., Martin, P., Amato, N.J., Avena, L.E., et al. (2023). Bivalent SARS-CoV-2 mRNA vaccines increase breadth of neutralization and protect against the BA.5 Omicron variant in mice. Nat. Med. *29*, 247–257. https://doi.org/10.1038/s41591-022-02092-8.
5. Kaku, Y., Uriu, K., Kosugi, Y., Okumura, K., Yamasoba, D., Uwamino, Y., Kuramochi, J., Sadamasu, K., Yoshimura, K., Nagashima, M., et al. (2024). Virological characteristics of the SARS-CoV-2 KP.2 variant. Lancet Infect Dis. *24*, e416. https://doi.org/10.1016/S1473-3099(24)00298-6.
6. Liu, Z., Zhou, J., Wang, X., Xu, W., Teng, Z., Chen, H., Chen, M., Zhang, G., Wang, Y., Huang, J., et al. (2023). A pan-sarbecovirus vaccine based on RBD of SARS-CoV-2 original strain elicits potent neutralizing antibodies against XBB in non-human primates. Proc. Natl. Acad. Sci. USA *120*, e2221713120. https://doi.org/10.1073/pnas.2221713120.
7. Li, D., Martinez, D.R., Schäfer, A., Chen, H., Barr, M., Sutherland, L.L., Lee, E., Parks, R., Mielke, D., Edwards, W., et al. (2022). Breadth of SARS-CoV-2 neutralization and protection induced by a nanoparticle vaccine. Nat. Commun. *13*, 6309. https://doi.org/10.1038/s41467-022-33985-4.
8. Cohen, A.A., van Doremalen, N., Greaney, A.J., Andersen, H., Sharma, A., Starr, T.N., Keeffe, J.R., Fan, C., Schulz, J.E., Gnanapragasam, P.N.P., et al. (2022). Mosaic RBD nanoparticles protect against challenge by diverse sarbecoviruses in animal models. Science *377*, eabq0839. https://doi.org/10.1126/science.abq0839.
9. Cohen, A.A., Gnanapragasam, P.N.P., Lee, Y.E., Hoffman, P.R., Ou, S., Kakutani, L.M., Keeffe, J.R., Wu, H.-J., Howarth, M., West, A.P., et al. (2021). Mosaic nanoparticles elicit cross-reactive immune responses to zoonotic coronaviruses in mice. Science *371*, 735–741. https://doi.org/10.1126/science.abf6840.
10. Baum, A., Fulton, B.O., Wloga, E., Copin, R., Pascal, K.E., Russo, V., Giordano, S., Lanza, K., Negron, N., Ni, M., et al. (2020). Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. Science *369*, 1014–1018. https://doi.org/10.1126/science.abd0831.
11. Schmidt, F., Weisblum, Y., Muecksch, F., Hoffmann, H.-H., Michailidis, E., Lorenzi, J.C.C., Mendoza, P., Rutkowska, M., Bednarski, E., Gaebler, C., et al. (2020). Measuring SARS-CoV-2 neutralizing antibody activity using pseudotyped and chimeric viruses. J. Exp. Med. *217*, e20201181. https://doi.org/10.1084/jem.20201181.
12. Jian, F., Wec, A.Z., Feng, L., Yu, Y., Wang, L., Wang, P., Yu, L., Wang, J., Hou, J., Berrueta, D.M., et al. (2024). A generalized framework to identify SARS-CoV-2 broadly neutralizing antibodies. Preprint at bioRxiv. https://doi.org/10.1101/2024.04.16.589454.
13. Greaney, A.J., Starr, T.N., Gilchuk, P., Zost, S.J., Binshtein, E., Loes, A.N., Hilton, S.K., Huddleston, J., Eguia, R., Crawford, K.H.D., et al. (2021). Complete Mapping of Mutations to the SARS-CoV-2 spike Receptor-Binding Domain that Escape Antibody Recognition. Cell Host Microbe *29*, 44–57.e9. https://doi.org/10.1016/j.chom.2020.11.007.
14. Starr, T.N., Greaney, A.J., Addetia, A., Hannon, W.W., Choudhary, M.C., Dingens, A.S., Li, J.Z., and Bloom, J.D. (2021). Prospective mapping of viral mutations that escape antibodies used to treat COVID-19. Science *371*, 850–854. https://doi.org/10.1126/science.abf9302.
15. Greaney, A.J., Loes, A.N., Crawford, K.H.D., Starr, T.N., Malone, K.D., Chu, H.Y., and Bloom, J.D. (2021). Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. Cell Host Microbe *29*, 463–476.e6. https://doi.org/10.1016/j.chom.2021.02.003.
16. Dong, J., Zost, S.J., Greaney, A.J., Starr, T.N., Dingens, A.S., Chen, E.C., Chen, R.E., Case, J.B., Sutton, R.E., Gilchuk, P., et al. (2021). Genetic and structural basis for SARS-CoV-2 variant neutralization by a two-antibody cocktail. Nat. Microbiol. *6*, 1233–1244. https://doi.org/10.1038/s41564-021-00972-2.
17. Starr, T.N., Greaney, A.J., Dingens, A.S., and Bloom, J.D. (2021). Complete map of SARS-CoV-2 RBD mutations that escape the monoclonal antibody LY-CoV555 and its cocktail with LY-CoV016. Cell Rep. Med. *2*, 100255. https://doi.org/10.1016/j.xcrm.2021.100255.
18. Greaney, A.J., Starr, T.N., Barnes, C.O., Weisblum, Y., Schmidt, F., Caskey, M., Gaebler, C., Cho, A., Agudelo, M., Finkin, S., et al. (2021). Mapping mutations to the SARS-CoV-2 RBD that escape binding by different classes of antibodies. Nat. Commun. *12*, 4196. https://doi.org/10.1038/s41467-021-24435-8.

Please cite this article in press as: Youssef et al., Computationally designed proteins mimic antibody immune evasion in viral evolution, Immunity (2025), https://doi.org/10.1016/j.immuni.2025.04.015

19. Starr, T.N., Czudnochowski, N., Liu, Z., Zatta, F., Park, Y.-J., Addetia, A., Pinto, D., Beltramello, M., Hernandez, P., Greaney, A.J., et al. (2021). SARS-CoV-2 RBD antibodies that maximize breadth and resistance to escape. Nature *597*, 97–102. https://doi.org/10.1038/s41586-021-03807-6.

20. Greaney, A.J., Loes, A.N., Gentles, L.E., Crawford, K.H.D., Starr, T.N., Malone, K.D., Chu, H.Y., and Bloom, J.D. (2021). Antibodies elicited by mRNA-1273 vaccination bind more broadly to the receptor binding domain than do those from SARS-CoV-2 infection. Sci. Transl. Med. *13*, eabi9915. https://doi.org/10.1126/scitranslmed.abi9915.

21. Greaney, A.J., Starr, T.N., Eguia, R.T., Loes, A.N., Khan, K., Karim, F., Cele, S., Bowen, J.E., Logue, J.K., Corti, D., et al. (2022). A SARS-CoV-2 variant elicits an antibody response with a shifted immunodominance hierarchy. PLoS Pathog. *18*, e1010248. https://doi.org/10.1371/journal.ppat.1010248.

22. Cao, Y., Jian, F., Wang, J., Yu, Y., Song, W., Yisimayi, A., Wang, J., An, R., Chen, X., Zhang, N., et al. (2023). Imprinted SARS-CoV-2 humoral immunity induces convergent Omicron RBD evolution. Nature *614*, 521–529. https://doi.org/10.1038/s41586-022-05644-7.

23. Yisimayi, A., Song, W., Wang, J., Jian, F., Yu, Y., Chen, X., Xu, Y., Yang, S., Niu, X., Xiao, T., et al. (2024). Repeated Omicron exposures override ancestral SARS-CoV-2 immune imprinting. Nature *625*, 148–156. https://doi.org/10.1038/s41586-023-06753-7.

24. Dadonaite, B., Brown, J., McMahon, T.E., Farrell, A.G., Figgins, M.D., Asarnow, D., Stewart, C., Lee, J., Logue, J., Bedford, T., et al. (2024). Spike deep mutational scanning helps predict success of SARS-CoV-2 clades. Nature *631*, 617–626. https://doi.org/10.1038/s41586-024-07636-1.

25. Jian, F., Wang, J., Yisimayi, A., Song, W., Xu, Y., Chen, X., Niu, X., Yang, S., Yu, Y., Wang, P., et al. (2025). Evolving antibody response to SARS-CoV-2 antigenic shift from XBB to JN.1. Nature *637*, 921–929. https://doi.org/10.1038/s41586-024-08315-x.

26. Schmidt, F., Weisblum, Y., Rutkowska, M., Poston, D., DaSilva, J., Zhang, F., Bednarski, E., Cho, A., Schaefer-Babajew, D.J., Gaebler, C., et al. (2021). High genetic barrier to SARS-CoV-2 polyclonal neutralizing antibody escape. Nature *600*, 512–516. https://doi.org/10.1038/s41586-021-04005-0.

27. Greaney, A.J., Eguia, R.T., Starr, T.N., Khan, K., Franko, N., Logue, J.K., Lord, S.M., Speake, C., Chu, H.Y., Sigal, A., et al. (2022). The SARS-CoV-2 Delta variant induces an antibody response largely focused on class 1 and 2 antibody epitopes. PLoS Pathog. *18*, e1010592. https://doi.org/10.1371/journal.ppat.1010592.

28. Starr, T.N., Greaney, A.J., Stewart, C.M., Walls, A.C., Hannon, W.W., Veesler, D., and Bloom, J.D. (2022). Deep mutational scans for ACE2 binding, RBD expression, and antibody escape in the SARS-CoV-2 Omicron BA.1 and BA.2 receptor-binding domains. PLoS Pathog. *18*, e1010951. https://doi.org/10.1371/journal.ppat.1010951.

29. Dadonaite, B., Crawford, K.H.D., Radford, C.E., Farrell, A.G., Yu, T.C., Hannon, W.W., Zhou, P., Andrabi, R., Burton, D.R., Liu, L., et al. (2023). A pseudovirus system enables deep mutational scanning of the full SARS-CoV-2 spike. Cell *186*, 1263–1278.e20. https://doi.org/10.1016/j.cell.2023.02.001.

30. Thadani, N.N., Gurev, S., Notin, P., Youssef, N., Rollins, N.J., Ritter, D., Sander, C., Gal, Y., and Marks, D.S. (2023). Learning from prepandemic data to forecast viral escape. Nature *622*, 818–825. https://doi.org/10.1038/s41586-023-06617-0.

31. Starr, T.N., Greaney, A.J., Hilton, S.K., Ellis, D., Crawford, K.H.D., Dingens, A.S., Navarro, M.J., Bowen, J.E., Tortorici, M.A., Walls, A.C., et al. (2020). Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. Cell *182*, 1295–1310. e20. https://doi.org/10.1016/j.cell.2020.08.012.

32. Carr, C.R., Crawford, K.H.D., Murphy, M., Galloway, J.G., Haddox, H.K., Matsen, F.A., 4th, Andersen, K.G., King, N.P., and Bloom, J.D. (2024). Deep mutational scanning reveals functional constraints and antibody-escape potential of Lassa virus glycoprotein complex. Immunity *57*, 2061–2076.e11. https://doi.org/10.1016/j.immuni.2024.06.013.

33. Gonzalez Somermeyer, L., Fleiss, A., Mishin, A.S., Bozhanova, N.G., Igolkina, A.A., Meiler, J., Alaball Pujol, M.-E., Putintseva, E.V., Sarkisyan, K.S., and Kondrashov, F.A. (2022). Heterogeneity of the GFP fitness landscape and data-driven protein design. eLife *11*, e75842. https://doi.org/10.7554/eLife.75842.

34. Yurkovetskiy, L., Wang, X., Pascal, K.E., Tomkins-Tinch, C., Nyalile, T.P., Wang, Y., Baum, A., Diehl, W.E., Dauphin, A., Carbone, C., et al. (2020). Structural and Functional Analysis of the D614G SARS-CoV-2 spike Protein Variant. Cell *183*, 739–751.e8. https://doi.org/10.1016/j.cell.2020.09.032.

35. Zech, F., Schniertshauer, D., Jung, C., Herrmann, A., Cordsmeier, A., Xie, Q., Nchioua, R., Prelli Bozzo, C., Volcic, M., Koepke, L., et al. (2021). Spike residue 403 affects binding of coronavirus spikes to human ACE2. Nat. Commun. *12*, 6855. https://doi.org/10.1038/s41467-021-27180-0.

36. McCallum, M., Walls, A.C., Sprouse, K.R., Bowen, J.E., Rosen, L.E., Dang, H.V., De Marco, A., Franko, N., Tilles, S.W., Logue, J., et al. (2021). Molecular basis of immune evasion by the Delta and Kappa SARS-CoV-2 variants. Science *374*, 1621–1626. https://doi.org/10.1126/science.abl8506.

37. Ju, B., Zhang, Q., Ge, J., Wang, R., Sun, J., Ge, X., Yu, J., Shan, S., Zhou, B., Song, S., et al. (2020). Human neutralizing antibodies elicited by SARS-CoV-2 infection. Nature *584*, 115–119. https://doi.org/10.1038/s41586-020-2380-z.

38. Cohen, A.A., Keeffe, J.R., Schiepers, A., Dross, S.E., Greaney, A.J., Rorick, A.V., Gao, H., Gnanapragasam, P.N.P., Fan, C., West, A.P., et al. (2024). Mosaic sarbecovirus nanoparticles elicit cross-reactive responses in pre-vaccinated animals. Cell *187*, 5554–5571. https://doi.org/10.1016/j.cell.2024.07.052.

39. WHO. COVID-19 vaccine tracker and landscape. https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines.

40. Chalkias, S., Harper, C., Vrbicky, K., Walsh, S.R., Essink, B., Brosz, A., McGhee, N., Tomassini, J.E., Chen, X., Chang, Y., et al. (2022). A Bivalent Omicron-Containing Booster Vaccine against Covid-19. N. Engl. J. Med. *387*, 1279–1291. https://doi.org/10.1056/NEJMoa2208343.

41. Tan, C.-W., Chia, W.-N., Young, B.E., Zhu, F., Lim, B.-L., Sia, W.-R., Thein, T.-L., Chen, M.I.-C., Leo, Y.-S., Lye, D.C., et al. (2021). Pan-Sarbecovirus Neutralizing Antibodies in BNT162b2-Immunized SARS-CoV-1 Survivors. N. Engl. J. Med. *385*, 1401–1406. https://doi.org/10.1056/NEJMoa2108453.

42. Pitts, J., Li, J., Perry, J.K., Du Pont, V., Riola, N., Rodriguez, L., Lu, X., Kurhade, C., Xie, X., Camus, G., et al. (2022). Remdesivir and GS-441524 Retain Antiviral Activity against Delta, Omicron, and Other Emergent SARS-CoV-2 Variants. Antimicrob. Agents Chemother. *66*, e0022222. https://doi.org/10.1128/aac.00222-22.

43. Sulis, G., Peebles, A., and Basta, N.E. (2023). Lassa fever vaccine candidates: A scoping review of vaccine clinical trials. Trop. Med. Int. Health *28*, 420–431. https://doi.org/10.1111/tmi.13876.

44. Notin, P., Kollasch, A.W., Ritter, D., Van Niekerk, L., Paul, S., Spinner, H., Rollins, N.J., Shaw, A., Orenbuch, R., Weitzman, R., et al. (2023). ProteinGym: Large-Scale Benchmarks for Protein Design and Fitness Prediction.. Preprint at bioRxiv.

45. Pascal, K.E., Coleman, C.M., Mujica, A.O., Kamat, V., Badithe, A., Fairhurst, J., Hunt, C., Strein, J., Berrebi, A., Sisk, J.M., et al. (2015). Pre- and postexposure efficacy of fully human antibodies against spike protein in a novel humanized mouse model of MERS-CoV infection. Proc. Natl. Acad. Sci. USA *112*, 8738–8743. https://doi.org/10.1073/pnas.1510830112.

46. Carter, S.R., Curtis, S., Emerson, C., Gray, J., Haydon, I.C., Hebbeler, A., Qureshi, C., Randolph, N., Rives, A., and Stuart, L. (2024). Responsible AI x Biodesign. Responsible AI x Biodesign. https://responsiblebiodesign.ai.

47. Boucau, J., Marino, C., Regan, J., Uddin, R., Choudhary, M.C., Flynn, J.P., Chen, G., Stuckwisch, A.M., Mathews, J., Liew, M.Y., et al. (2022).

Please cite this article in press as: Youssef et al., Computationally designed proteins mimic antibody immune evasion in viral evolution, Immunity (2025), https://doi.org/10.1016/j.immuni.2025.04.015

Duration of Shedding of Culturable Virus in SARS-CoV-2 Omicron (BA.1) Infection. N. Engl. J. Med. *387*, 275–277. https://doi.org/10.1056/NEJMc2202092.

48. Seaman, M.S., Siedner, M.J., Boucau, J., Lavine, C.L., Ghantous, F., Liew, M.Y., Mathews, J.I., Singh, A., Marino, C., Regan, J., et al. (2022). Vaccine breakthrough infection leads to distinct profiles of neutralizing antibody responses by SARS-CoV-2 variant. JCI Insight *7*, e159944. https://doi.org/10.1172/jci.insight.159944.
49. Li, Y., Choudhary, M.C., Regan, J., Boucau, J., Nathan, A., Speidel, T., Liew, M.Y., Edelstein, G.E., Kawano, Y., Uddin, R., et al. (2024). SARS-CoV-2 viral clearance and evolution varies by type and severity of immunodeficiency. Sci. Transl. Med. *16*, eadk1599. https://doi.org/10.1126/scitranslmed.adk1599.
50. Frazer, J., Notin, P., Dias, M., Gomez, A., Min, J.K., Brock, K., Gal, Y., and Marks, D.S. (2021). Disease variant prediction with deep generative models of evolutionary data. Nature *599*, 91–95. https://doi.org/10.1038/s41586-021-04043-8.
51. Wilks, S. (2024). Racmacs: Antigenic Cartography Macros. https://acorg.github.io/Racmacs/.
52. Johnson, L.S., Eddy, S.R., and Portugaly, E. (2010). Hidden Markov model speed heuristic and iterative HMM search procedure. BMC Bioinformatics *11*, 431. https://doi.org/10.1186/1471-2105-11-431.
53. Mou, H., Quinlan, B.D., Peng, H., Liu, G., Guo, Y., Peng, S., Zhang, L., Davis-Gardner, M.E., Gardner, M.R., Crynen, G., et al. (2021). Mutations derived from horseshoe bat ACE2 orthologs enhance ACE2-Fc neutralization of SARS-CoV-2. PLoS Pathog. *17*, e1009501. https://doi.org/10.1371/journal.ppat.1009501.
54. Gangavarapu, K., Latif, A.A., Mullen, J.L., Alkuzweny, M., Hufbauer, E., Tsueng, G., Haag, E., Zeller, M., Aceves, C.M., Zaiets, K., et al. (2023). Outbreak.info. Outbreak.info genomic reports: Scalable and dynamic surveillance of SARS-CoV-2 variants and mutations. Nat. Methods *20*, 512–522. https://doi.org/10.1038/s41592-023-01769-3.
55. Suzek, B.E., Huang, H., McGarvey, P., Mazumder, R., and Wu, C.H. (2007). UniRef: comprehensive and non-redundant UniProt reference clusters. Bioinformatics *23*, 1282–1288. https://doi.org/10.1093/bioinformatics/btm098.
56. Khare, S., Gurry, C., Freitas, L., Schultz, M.B., Bach, G., Diallo, A., Akite, N., Ho, J., Lee, R.T., Yeo, W., et al. (2021). GISAID's Role in Pandemic Response. China CDC Wkly. *3*, 1049–1051. https://doi.org/10.46234/ccdcw2021.255.
57. Edgar, R.C. (2022). Muscle5: High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. Nat. Commun. *13*, 6968. https://doi.org/10.1038/s41467-022-34630-w.
58. Eisenberg, D., Weiss, R.M., and Terwilliger, T.C. (1984). The hydrophobic moment detects periodicity in protein hydrophobicity. Proc. Natl. Acad. Sci. USA *81*, 140–144. https://doi.org/10.1073/pnas.81.1.140.
59. Takebe, Y., Seiki, M., Fujisawa, J., Hoy, P., Yokota, K., Arai, K., Yoshida, M., and Arai, N. (1988). SR alpha promoter: an efficient and versatile mammalian cDNA expression system composed of the simian virus 40 early promoter and the R-U5 segment of human T-cell leukemia virus type 1 long terminal repeat. Mol. Cell. Biol. *8*, 466–472. https://doi.org/10.1128/mcb.8.1.466-472.1988.
60. Ujike, M., Huang, C., Shirato, K., Makino, S., and Taguchi, F. (2016). The contribution of the cytoplasmic retrieval signal of severe acute respiratory syndrome coronavirus to intracellular accumulation of S proteins and incorporation of S protein into virus-like particles. J. Gen. Virol. *97*, 1853–1864. https://doi.org/10.1099/jgv.0.000494.
61. Lai, R.P.J., Yan, J., Heeney, J., McClure, M.O., Göttlinger, H., Luban, J., and Pizzato, M. (2011). Nef decreases HIV-1 sensitivity to neutralizing antibodies that target the membrane-proximal external region of TMgp41. PLoS Pathog. *7*, e1002442. https://doi.org/10.1371/journal.ppat.1002442.
62. Tortorici, M.A., Czudnochowski, N., Starr, T.N., Marzi, R., Walls, A.C., Zatta, F., Bowen, J.E., Jaconi, S., Di Iulio, J., Wang, Z., et al. (2021). Broad sarbecovirus neutralization by a human monoclonal antibody. Nature *597*, 103–108. https://doi.org/10.1038/s41586-021-03817-4.
63. Yisimayi, A., Song, W., Wang, J., Jian, F., Yu, Y., Chen, X., Xu, Y., Yang, S., Niu, X., Xiao, T., et al. (2024). Repeated Omicron infection alleviates SARS-CoV-2 immune imprinting. Nature *625*, 148–156. https://doi.org/10.1038/s41586-023-06753-7.

Please cite this article in press as: Youssef et al., Computationally designed proteins mimic antibody immune evasion in viral evolution, Immunity (2025), https://doi.org/10.1016/j.immuni.2025.04.015

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Human polyclonal sera - B.1 convalescent infection	Serum blood donation. This paper.	N/A
Human polyclonal sera - Delta convalescent infection	Serum blood donation. This paper.	N/A
Human polyclonal sera - B.1 (Moderna) Primary Vaccination	Serum blood donation. This paper.	N/A
Human polyclonal sera - B.1 (Pfizer) Primary Vaccination	Serum blood donation. This paper.	N/A
Human polyclonal sera - B.1 (Moderna) Boost	Serum blood donation. This paper.	N/A
Human polyclonal sera - B.1 (Pfizer) Boost	Serum blood donation. This paper.	N/A
Human polyclonal sera - B.1 & BA.4/5 Bivalent Boost	Serum blood donation. This paper.	N/A
Human polyclonal sera - Delta Vaccine Breakthrough Infection	Collected from patients[47–49]	N/A
Human polyclonal sera - BA.1 Vaccine Breakthrough Infection	Collected from patients[47–49]	N/A
Human polyclonal sera - BA.2.12.1 Vaccine Breakthrough Infection	Collected from patients[47–49]	N/A
Human polyclonal sera - BA.4/5 Vaccine Breakthrough Infection	Collected from patients[47–49]	N/A
NHP polyclonal sera - bivalent mRNA Vaccine Boost	Collected from NHP (Cohen et al.[38])	N/A
NHP polyclonal sera - mosaic-8b nanoparticle Boost	Collected from NHP (Cohen et al.[38])	N/A
NHP polyclonal sera - homotypic nanoparticle Boost	Collected from NHP (Cohen et al.[38])	N/A
Bacterial and virus strains		
See Table S2 for 104 viral strains used	This paper	See Table S2 for addgene identifiers
Chemicals, peptides, and recombinant proteins		
Filters, 0.45 μM	Avantor	Cat#76479-020
Critical commercial assays		
RT Assay	Yurkovetskiy et al.[34]	N/A
Steady-Glo Luciferase Assay System	Promega	Cat#E2550
Bright-Glo Luciferase Assay System	Promega	Cat#E2650
HIV pseudotyping assay	Yurkovetskiy et al.[34]	N/A
Deposited data		
SARS-CoV-2 Sequences	GISAID[49]	N/A
Spike structures in open and closed conformations	PDB	PDB:6VXX; PDB:6VYB; PDB:7CAB; PDB:7BNN
Raw and analyzed data	This paper	Tables S3 and S4 and https://github.com/debbiemarkslab/Vax_design
Experimental models: Cell lines		
HEK 293T cells with stably overexpressed ACE2	Drs. Michael Farzan and Huihui Ma (The Scripps Research Institute)	N/A
HEK 293T/17 cells	ATCC	Cat#CRL-11268

(*Continued on next page*)

Please cite this article in press as: Youssef et al., Computationally designed proteins mimic antibody immune evasion in viral evolution, Immunity (2025), https://doi.org/10.1016/j.immuni.2025.04.015

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
HEK293T cells with stably expressed ACE2/TMPRSS2	Yurkovetskiy et al.[34]	N/A
HEK293T cells	ATCC	Cat#CRL-3216
Recombinant DNA		
Plasmid: pDMJ2	This paper	N/A
Plasmid: pCMV-R8.2	Dr. Barney Graham (NIH Vaccine Research Center)	N/A
Plasmid: pHR'CMV-Luc	Dr. Barney Graham (NIH Vaccine Research Center)	N/A
Software and algorithms		
EVE-Vax	This paper	https://github.com/debbiemarkslab/Vax_design
EVEscape	Thadani et al.[30]	https://github.com/OATML-Markslab/EVEscape
EVE	Frazer et al.[50]	https://github.com/OATML-Markslab/EVE
Racmacs	Wilks[51]	https://acorg.github.io/Racmacs/
Jackhmmer	Johnson et al.[52]	http://hmmer.org/download.html

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Cell lines

HEK293T cells (derived from HEK293 Homo sapiens, female, embryonic kidney cells) used here were obtained from the American Type Culture Collection or ATCC (CRL-3216 for 293T or CRL-11268 for 293T/17). HEK293T cells with stably expressed human ACE2 were obtained from the same cell clone previously described by Mou et al.[53] and were kindly provided by Drs. Michael Farzan and Huihui Ma (The Scripps Research Institute). HEK293T cells were transduced with vesicular stomatitis virus (VSV) G protein-pseudotyped murine leukemia viruses to stably express the human ACE2 cell surface receptor protein. All cells were cultured in humidified incubators with 5% CO2 at 37°C, and monitored for mycoplasma contamination using the Mycoplasma Detection kit (Lonza LT07-318), were cultured in DMEM supplemented with 10% heat-inactivated FBS, 1 mM sodium pyruvate, 20 mM GlutaMAX, 1 × MEM non-essential amino acids, and 25 mM HEPES, pH 7.2. HEK293T cells with stably expression of the human angiotensin converting enzyme 2 (ACE2) and the transmembrane serine protease 2 (TMPRSS2) (HEK293T-ACE2-TMPRSS2 cells) were developed by the Luban laboratory.[34]

SARS-CoV-2 convalescent and vaccine human sera

We analyzed 115 serum samples collected from individuals who were either convalescent, vaccinated, boosted, or experienced breakthrough infection (Table S4). Informed consent was obtained from participants in the Post-vaccination Viral Characteristics Study (POSITIVES) study.[47–49] Serum samples were acquired approximately 2-4 weeks after vaccination or convalescence. For patients experiencing breakthrough infection, whole genome sequencing was performed on a nasal swab collected at the time of diagnosis to confirm infecting SARS-CoV-2 variant.

In order to maximize the number of designed spike variant pseudoviruses that can be tested against the same immune serum panels, a total of 23 serum pools were created by grouping 5 individual patient samples per pool. The pooled sera are representative of the different populations that have existed throughout the pandemic: natural infection, primary vaccination, boosted, bivalent boosted, and breakthrough infection. When sufficient samples were available, individual serum samples were pooled based on neutralization titers resulting in low, medium, and high neutralization pools.

Sarbecovirus non-human primate vaccine sera

Fourteen NHPs (5-year-old male and female cynomolgus macaques of Mauritian origin) received four doses of DNA or repRNA SARS-CoV-2 vaccines 64 and 30 weeks prior to boosting with either a BA.1 bivalent mRNA vaccine (n = 4), a mosaic-8b sarbecovirus RBD nanoparticle (n = 5), or a homotypic sarbecovirus RBD nanoparticle (n = 5) vaccine.[38] Samples were collected 2-weeks post booster. The mixed immune history in these groups of pre-vaccinated NHPs is representative of a complex immune history in people who have been vaccinated and/or infected multiple times.

Please cite this article in press as: Youssef et al., Computationally designed proteins mimic antibody immune evasion in viral evolution, Immunity (2025), https://doi.org/10.1016/j.immuni.2025.04.015

METHOD DETAILS

Multiple Sequence Alignments for fitness models

We designed spike constructs on the backgrounds of five VOCs: the canonical wildtype SARS-CoV-2 spike protein (Uniprot ID: P0DTC2) with D614G mutation (B.1), and the BA.4/5, BA.2.12.1, BA.2.75, and XBB spike sequences. We define the pandemic lineages based on the list of mutations described in outbreak.info.[54] For each round of construct design, we built multiple sequence alignments for spike with both pre-pandemic sequences and pandemic spike sequences seen before the emergence of the variant used as the background for the design.

To generate the pre-pandemic alignment, we used 5 iterations of the jackhmmer iterative HMM-based homology search alignment tool,[52] searching against the Uniref100 dataset[55] and removed all sequences deposited after January 2020. To account for the growing number of pandemic sequences, we included unique pandemic sequences seen more than 100 times in the Global Initiative on Sharing All Influenza Data (GISAID) EpiCoV database (www.gisaid.org)[56] at the time of emergence of the variant (Table S1). The resulting pandemic sequences and evolutionary alignments were combined and aligned with super5,[57] version 5.1. We remove sequences that are aligned to less than 50% of the query sequence and remove residue positions that contain more than 70% gap characters. We down weighted redundant sequence clusters by assigning each protein sequence i a weight $w(i) = 1/T$, where T is the number of sequences in the alignment within a given hamming distance cutoff of t (0.01).

For the B.1 variant, two models were trained to assess the impact of incorporating SARS-CoV-2 sequences. The first model was trained exclusively on sequences available from pre-pandemic coronaviruses (n = 4,577 non-SARS-CoV-2 spike sequences; Table S1). For comparison purposes, the second model was updated to include SARS-CoV-2 spike sequences seen prior to May 2022 (n = 5,868, with 1,291 sequences from SARS-CoV-2 which represented ~9 independent sequences in model training due to high similarity) and used for a much smaller design round (6 of the 21 B.1 designs were using this model). For all other variants, model training included all SARS-CoV-2 spike sequences available up to the emergence of the respective variant (Tables S1). Due to the high similarity across the SARS-CoV-2 sequences, they at most represented 1.4% of the training dataset by effective number of weighted-redundant sequences.

EVE-Vax protein design algorithm

We use the EVEscape framework,[30] which combines fitness, accessibility, and dissimilarity, to predict a mutant's ability to escape antibody recognition. For the fitness (f) component we trained an EVE model,[50] a Bayesian variational autoencoder (VAE) on the multiple sequence alignments. We estimate the relative fitness of each sequence as the log likelihood ratio between a mutated sequence and the wild type sequence. This ratio is itself approximated as the difference in Evidence Lower Bound (ELBO). We estimate ELBO values using twenty thousand Monte Carlo samples of the latent space. For the dissimilarity (d) component, we use both charge and hydrophobicity[58] differences between each pair of amino acids and then assign each pair a dissimilarity value equal to the sum of the standard-scaled differences. For the accessibility (a) component, we selected spike structures from the RCSB Protein Data Bank (PDB) representative of both ''open'' and ''closed'' configurations (PDB IDs: 6VXX, 6VYB, 7CAB, and 7BNN). Using these structures, we calculated the weighted contact number (WCN) for each residue position i as $WCN_i = \sum_{j \neq i} (r_{ij}^{-2})$ where r_{ij} is the distance between the geometric centers of the side chain of the residues occupying sites i and j. Each position is then assigned the minimum WCN across all structures. We use the negative WCN as a measure of residue accessibility to antibodies.

For each mutation, we combined the fitness, accessibility, and dissimilarity scores using a temperature scaled logistic function to get a single escape score for each individual mutation. The score of a mutation m is calculated as

$$EVEscape(m) = logistic\left(T_f^{-1} * F_f^m\right) * logistic\left(T_a^{-1} * F_a^m\right) * logistic\left(T_d^{-1} * F_d^m\right)$$

where T_i is the temperature scaling and F_i is the standardized vector for factor i. We then take the log transform of the product. We modified the temperature parameters compared to the values used previously[30] by increasing the contribution of the fitness component, $T_f = 1, T_a = 8$, and $T_d = 16$. The modular design of this framework enhances model interpretability by allowing us to assess the relative contributions of each of the three components to a mutation's overall escape score (Figure S1).

We designed 83 constructs on the backgrounds of five different spike VOC sequences: B.1 (ancestral strain with D614G), BA.4/5, BA.2.12.1, BA.2.75, and XBB. For each background variant we first scored all possible single amino acid substitutions to spike. Note that we consider amino acid mutations involving any number of nucleotide changes rather than focusing on only single nucleotide changes which are by far the most common (more than 99%) mutations seen in GISAID.[56] To design constructs, we considered single amino acid mutations in the top 1% of EVEscape score. We then generated and scored all possible double mutants. Lastly, we combined highly fit double mutants and scored them to create higher order constructs.

Recombinant DNA sequence plasmid design

All recombinant DNA work was conducted according to protocols approved by the University of Massachusetts Chan Medical School Institutional Biosafety Committee. We designed a pDMJ2 plasmid where spike expression is under the control of a cytomegalovirus immediate early promoter (CMV-IE) fused to the human T lymphotropic virus type 1 (HTLV-1) 5′ UTR.[59] Transcription termination is directed by a bovine growth hormone polyadenylation signal (bGH-PolyA). The pDMJ2 plasmid carries an origin of

Please cite this article in press as: Youssef et al., Computationally designed proteins mimic antibody immune evasion in viral evolution, Immunity (2025), https://doi.org/10.1016/j.immuni.2025.04.015

CellPress
OPEN ACCESS

replication (ori) from ColE1 and an Ampicillin resistance gene. spike sequences were inserted between the *XbaI* and *BamHI* restriction sites (Table S2; Figure S1).

An ancestral B.1 spike sequence was codon optimized by Twist Biosciences with a deletion of the terminal 18 amino acids to remove the endoplasmic reticulum retention signal.[60] This sequence was used as the base sequence on which other constructs were designed to minimize effects of nucleotide sequence on protein synthesis efficiency. Specifically, to generate spike sequences of VOCs and designed constructs, we mutated individual codon positions to the most commonly used codon in human cells encoding the mutant amino acid. We repeated this for all mutations in a given sequence. If a restriction site was introduced, the second most common codon was used.

Infectivity assays with SARS-CoV-2 spike-pseudotyped lentiviral particles

All infectivity assays utilized replication-incompetent lentiviral pseudotypes and were conducted according to protocols approved by the University of Massachusetts Chan Medical School Institutional Biosafety Committee. HIV-1-derived virions bearing a luciferase reporter gene, and pseudotyped with a SARS-CoV-2 spike variant, were produced by transfection of HEK293T cells, as previously reported.[34] Each S protein expression plasmid was separately transfected into HEK293 cells with plasmids encoding HIV-1 structural proteins and enzymes. Separate plasmids were transfected that encode RNAs with HIV-1 *cis*-acting signals for packaging and replication and either GFP or luciferase (Luc) reporter cassettes. Fresh media was added after spinfection and cells were incubated prior to analysis. The supernatant from the transfected cells was collected 72 hours post-transfection and filtered (0.45 μM, Avantor™) to remove cellular debris. For each condition tested, multiple virus stocks were produced, and each stock was tested in triplicate after vector particle normalization by using reverse-transcriptase activity. Virion yield in each transfection supernatant was normalized using our in-house exogenous reverse transcriptase (RT) activity assay.[34,61] As in Yurkovetskiy et al.,[34] 5 μL transfection supernatant was mixed with 5 μL 0.25% Triton X-100, 50 mM KCl, 100 mM Tris-HCl pH 7.4, and 0.4 U/μL RiboLock RNase inhibitor, and then diluted 1:100 in 5 mM $(NH_4)_2SO_4$, 20 mM KCl, and 20 mM Tris-HCl pH 8.3. 10 μL of this was then added to a single-step, RT-PCR assay with 35 nM MS2 RNA (IDT) as template, 500 nM of each primer (5′-TCCTGCTCAACTTCCTGTCGAG-3′ and 5′-CAC AGGTCAAACCTCCTAGGAATG-3′), and 0.1 μL hot-start Taq DNA polymerase (Promega, Madison, WI) in 20 mM Tris-Cl pH 8.3, 5 mM $(NH_4)_2SO_4$, 20 mM KCl, 5 mM $MgCl_2$, 0.1 mg/mL BSA, 1/20,000 SYBR Green I (Invitrogen), and 200 μM dNTPs in total 20 μL reaction. The RT-PCR reaction was carried out in a Biorad CFX96 real-time PCR detection system with the following parameters: 42°C for 20 minutes, 95°C for 2 minutes, and 40 cycles [95°C for 5 seconds, 60°C for 5 seconds, 72°C for 15 seconds, and acquisition at 80°C for 5 seconds].

Supernatant containing pseudovirions was then added to HEK293T cells that were previously transduced with puromycin- and blasticidin-resistant lentivectors to stably express the human angiotensin converting enzyme 2 (ACE2) and the transmembrane serine protease 2 (TMPRSS2),[34] respectively (HEK293T-ACE2-TMPRSS2). 16 hours prior to transduction, HEK293T-ACE2-TMPRSS2 cells were seeded in 96 well plates. HEK293T cells were plated at 7.5×10^3 cells per well. Cells were incubated in virus-containing media for 16 hours at 37°C when fresh medium was added to cells and returned to incubation for another 56 hours.. Luciferase activity on target cells was measured as an indication of transduction efficiency, using the Steady-Glo® Luciferase Assay System (Promega Corporation) read on a Promega GloMax Discover machine. All plasmids were deposited to Addgene (Table S2). All experiments were performed in triplicate. Sequences were considered non-functional if they had the same measured transduction (Relative Light Units; RLU) as the negative control (pDMJ2 without spike inserted, $<1 \times 10^4$ RLU).

Neutralization assays with SARS-CoV-2 spike-pseudotyped lentiviral particles

Neutralizing activity against SARS-CoV-2 pseudovirus was measured using a single-round infection assay in 293T/ACE2 target cells. Pseudotyped virus particles were produced in 293T/17 cells (ATCC) by co-transfection of pDMJ2 plasmids encoding codon-optimized SARS-CoV-2 spike variant, packaging plasmid pCMV R8.2, and luciferase reporter plasmid pHR' CMV-Luc. Packaging and luciferase plasmids were kindly provided by Dr. Barney Graham (NIH, Vaccine Research Center). The 293T cell line stably overexpressing the human ACE2 cell surface receptor protein was kindly provided by Drs. Michael Farzan and Huihui Ma (The Scripps Research Institute). For neutralization assays, serial dilutions of patient serum samples were performed in duplicate followed by addition of pseudovirus. Pooled serum samples from convalescent COVID-19 patients or pre-pandemic normal healthy serum (NHS) were used as positive and negative controls, respectively. Plates were incubated for 1 hour at 37 C followed by addition of 293/ACE2 target cells ($1x10^4$ /well). Wells containing cells and pseudovirus (without sample) or cells alone acted as positive and negative infection controls, respectively. Assays were harvested on day 3 using Promega BrightGlo luciferase reagent and luminescence detected with a Promega GloMax luminometer. Titers are reported as the dilution of serum that inhibited 50% or 80% virus infection (ID_{50} and ID_{80} titers, respectively).

Analysis of proportion of viable sequences at different mutational depths

We used the raw data from four deep mutation scans (viral and non-viral) to quantify the likelihood of obtaining functional sequences with multiple mutations compared to a wild type sequence. We analyzed the SARS-CoV-2 RBD data from Starr et al.,[31] the full spike SARS-CoV-2 data from Dadonaite et al.,[29] the Lassa glycoprotein complex (GPC) data from Carr et al.,[32] and the *Clytia gregaria* green fluorescent protein (GFP) data from Somermeyer et al.[33]

In the RBD dataset, mutations were introduced using error-prone PCR resulting in 135,386 unique mutant RBDs with up to 11 amino acid mutations. The expression of each sequence was measured as the change in mean fluorescence intensity, Δlog(MFI),

Please cite this article in press as: Youssef et al., Computationally designed proteins mimic antibody immune evasion in viral evolution, Immunity (2025), https://doi.org/10.1016/j.immuni.2025.04.015

relative to the unmutated SARS-CoV-2 RBD. We used sequences containing synonymous mutations (labeled viable) and sequences with premature stop codons (labeled nonviable) to train a logistic regression for classifying missense variants as either viable or nonviable. The fitted model had an intercept of 4.66 and a coefficient of 1.39. The decision threshold estimated for expression scores was -1.56 Δlog(MFI) with variants having expression values less than this boundary classified as non-viable, and variants with expression scores greater than the boundary classified as viable. When analysing this same data, Greaney et al.[15] used an expression threshold of -1 Δlog(MFI) to classify variants as nonviable since that was the expression deficit for mutations to core disulfide residues. More recently, Greaney et al.[13] used an expression threshold of -1.5 Δlog(MFI). Our estimated decision threshold of -1.56 Δlog(MFI) is therefore in-line with previous biologically-informed cutoffs.

We repeated the logistic regression analysis described above on the other three datasets. In the full spike dataset from Dadonaite et al.,[29] only mutations seen in the pandemic were introduced to the full spike Delta variant in higher-order combinations. The dataset included 191,418 unique variants containing up to 38 mutations. For each variant, a functional score was estimated based on its relative frequency in the spike- versus VSV-G-pseudotyped libraries. The fitted model had an intercept of 6.11 and a coefficient of 3.91. The decision threshold was estimated to be a functional score of -3.34, with variants having values lower than this cutoff classified as non-viable and variants with higher values classified as viable.

In the Lassa GPC data from Carr et al.,[32] the libraries were created using error-prone PCR. The dataset included 66,793 unique variants containing up to 34 mutations. For each variant, a functional score was estimated based on cell entry measurements in 293T cells. The fitted model had an intercept of 2.5 and a coefficient of 0.8. The decision threshold was estimated to be a functional score of -3.15, with variants having values lower than this cutoff classified as non-viable and variants with higher values classified as viable.

In the *Clytia gregaria* GFP dataset from Somermeyer et al.,[33] the libraries were created using error-prone PCR. The dataset included 26,165 unique variants containing up to 230 mutations. For each variant, a functional score was estimated based on the log of the green fluorescence intensity. The fitted model had an intercept of -46.6 and a coefficient of 10.9. The decision threshold was estimated to be a functional score of 4.25, with variants having values lower than this cutoff classified as non-viable and variants with higher values classified as viable.

Analysis of escape mutation frequencies in the public database

We analyzed experimental data from 20 high-throughput deep mutational scanning studies (Table S5) that measured escape mutations against monoclonal antibodies (mAbs) or sera.[2,3,6,13–25,27–29,62] For each study, mutations were assigned an escape score, calculated as either the maximum or mean score across conditions (antibodies or sera). Escape mutations were defined as nonsynonymous mutations within one nucleotide of the wildtype amino acid and ranked in the top 5% of escape scores for each study. The proportion of escape mutations observed more than 1,000 times throughout the pandemic (using GISAID[56] downloaded on Dec 19 2024) is reported. Proportions for each study independently and with different thresholds (top 1%, 5% and 10%) are shown in Figure S4.

To estimate the cumulative proportion of escape mutations across studies (Figure 5A), escape mutations were initially defined based on the earliest available experimental data. When subsequent studies became available, we updated the set of escape mutations to include any additional mutations identified by the most recent study which were not in the prior set of escape mutations. This iterative approach accounts for variability between studies conducted in different laboratories, at different times, and under differing experimental conditions, while maintaining the utility of earlier studies performed on previous variants of concern. The same protocol was applied to define escape mutations predicted by computational models.

Comparison to assay-derived constructs designed using high throughput experiments

We compare the EVE-Vax designed constructs to multi-mutant constructs designed from large-scale experimental studies of escape from polyclonal sera or thousands of monoclonal antibodies.[11,22,63] For B.1 designs, we compare against designs from Schmidt et al.,[11] with matched sera: High Titer Convalescent Wildtype (ours: High convalescent WT titer pool (n=5), theirs: RU27 (n=27)), Convalescent Wildtype (ours: Convalescent WT (n=15), theirs: Ran21, (n=21)), and Vaccine (ours: Primary Vaccination (n=25), theirs: Vac14 (n =14); Figure S5). Their designs included mutations chosen from passaging a (rVSV)/SARS-CoV-2 chimeric virus in the presence of plasmas for up to six passages, and mutations were chosen based on their effects on plasma neutralization sensitivity and distribution throughout the spike protein. The constructs with the highest average fold reduction were B.1-8a (ours) and PSM1-1 (theirs).

For BA.5, we compare against the 8 BA.5 RBD designs generated by Cao et al.,[22] with matched sera: Boost (ours: Boost (n=25), theirs: CoronaVac 3 doses (n= 40)), BA.1 BTI (ours: BA.1 BTI (n=15), theirs: BA.1 BTI post 3 doses of CoronaVac, (n= 50)), BA.2.12.1 BTI (ours: BA.2.12.1 BTI (n=5), theirs: BA.2 BTI post 3 doses of CoronaVac (n=50)) and BA.5 BTI (ours: BA.5 BTI (n=5), theirs: BA.5 BTI post 3 doses of CoronaVac, (n=36); Figure S5). Their designs included mutations chosen from measuring high-throughput scans of human-ACE2 binding affinity, RBD expression and neutralizing activity across 3,051 BA.2-neutralizing antibodies from varied immune backgrounds, as well as codon constraints on each residue. The constructs with the highest average fold reduction were BA.5-5a (ours) and BA.5-S8 (theirs).

For BA.2.75, we compare against the 8 BA.2.75 RBD designs generate by Cao et al.,[22] with matched sera: Boost (ours: Boost (n=25), theirs: CoronaVac 3 doses (n= 40)), BA.1 BTI (ours: BA.1 BTI (n=15), theirs: BA.1 BTI post 3 doses of CoronaVac, (n= 50)), BA.2.12.1 BTI (ours: BA.2.12.1 BTI (n=5), theirs: BA.2 BTI post 3 doses of CoronaVac (n=50)) and BA.5 BTI (ours: BA.5 BTI (n=5),

Please cite this article in press as: Youssef et al., Computationally designed proteins mimic antibody immune evasion in viral evolution, Immunity (2025), https://doi.org/10.1016/j.immuni.2025.04.015

CellPress
OPEN ACCESS

theirs: BA.5 BTI post 3 doses of CoronaVac, (n=36); Figure S5). Their designs included mutations chosen from measuring high-throughput scans of human-ACE2 binding affinity, RBD expression and neutralizing activity across 3,051 BA.2-neutralizing antibodies from varied immune backgrounds, as well as codon constraints on each residue. The constructs with the highest average fold reduction were BA.2.75-4c (ours) and BA.2.75-S7 (theirs).

For XBB, we compare against the 7 XBB.1.5 RBD designs generated by Yisimayi et al.,[63] with matched sera: BA.1 BTI (ours: BA.1 BTI (ours: BA.1 BTI (n=15), theirs: BA.1 BTI prior to BA.5/BF.7 reinfection (n=26)), BA.2.12.1 BTI (ours: BA.2.12.1 BTI (n=5), theirs: BA.2 BTI prior to BA.5/BF.7 reinfection, (n=19)), and BA.5 BTI (ours: BA.5 BTI (n=5), theirs: Reinfection with BA.5 or BF.7 after BA.1 or BA.2 infection without vaccination (n=12); Figure S5). Their designs included mutations chosen from measuring high-throughput scans of human-ACE2 binding affinity, RBD expression and neutralizing activity across 1,816 SARS-CoV-2 RBD-targeting antibodies from varied immune backgrounds, as well as codon constraints on each residue. The constructs with the highest average fold reduction were XBB-9a relative to XBB (ours) and XBB.1.5-S5 relative to XBB.1.5 (theirs).

QUANTIFICATION AND STATISTICAL ANALYSIS

We use the Wilcoxon signed rank test from SciPy for comparing differences in neutralization, with the p-value indicated and fold reduction in geometric mean ID_{50} titers shown. For XBB designs the Delta convalescent and primary vaccine sera pool were not included since the serum pools were depleted at the time these constructs were made. The statistical test in this case only considered the remaining sera.

For antigenic cartography, we use Racmacs[51] to perform multidimensional scaling (MDS) on our titration data measuring the neutralization sensitivity of all VOCs and designs for each serum pool. MDS is used to create a 2-dimensional representation of the pairwise antigenic distance between each VOC or design, as dictated by how each is neutralized by the different sera pools. MDS pairwise distances are consistent, which can be seen across pairs of variant-variant, variant-design, and design-design (Figure S3). This allows us to visualize the trajectory of evolution in antigenic space.