



Regulatory Approved Monoclonal Antibodies Contain Framework Mutations Predicted From Human Antibody Repertoires

OPEN ACCESS

Brian M. Petersen¹, Sophia A. Ulmer¹, Emily R. Rhodes¹, Matias F. Gutierrez-Gonzalez², Brandon J. Dekosky^{2,3}, Kayla G. Sprenger^{1*} and Timothy A. Whitehead^{1*}

Edited by:

Hiromi Kubagawa,
German Rheumatism Research
Center (DRFZ), Germany

Reviewed by:

Mats Ohlin,
Lund University, Sweden
Claude-Agnes Reynaud,
Université Paris Descartes,
France

*Correspondence:

Timothy A. Whitehead
timothy.whitehead@colorado.edu
Kayla G. Sprenger
Kayla.Sprenger@colorado.edu

Specialty section:

This article was submitted to
B Cell Biology,
a section of the journal
Frontiers in Immunology

Received: 22 June 2021

Accepted: 03 September 2021

Published: 27 September 2021

Citation:

Petersen BM, Ulmer SA, Rhodes ER,
Gutierrez-Gonzalez MF, Dekosky BJ,
Sprenger KG and Whitehead TA
(2021) Regulatory Approved
Monoclonal Antibodies Contain
Framework Mutations Predicted From
Human Antibody Repertoires.
Front. Immunol. 12:728694.
doi: 10.3389/fimmu.2021.728694

¹ Department of Chemical and Biological Engineering, University of Colorado, Boulder, CO, United States, ² Department of Pharmaceutical Chemistry, University of Kansas, Lawrence, KS, United States, ³ Department of Chemical Engineering, University of Kansas, Lawrence, KS, United States

Monoclonal antibodies (mAbs) are an important class of therapeutics used to treat cancer, inflammation, and infectious diseases. Identifying highly developable mAb sequences *in silico* could greatly reduce the time and cost required for therapeutic mAb development. Here, we present position-specific scoring matrices (PSSMs) for antibody framework mutations developed using baseline human antibody repertoire sequences. Our analysis shows that human antibody repertoire-based PSSMs are consistent across individuals and demonstrate high correlations between related germ lines. We show that mutations in existing therapeutic antibodies can be accurately predicted solely from baseline human antibody sequence data. We find that mAbs developed using humanized mice had more human-like FR mutations than mAbs originally developed by hybridoma technology. A quantitative assessment of entire framework regions of therapeutic antibodies revealed that there may be potential for improving the properties of existing therapeutic antibodies by incorporating additional mutations of high frequency in baseline human antibody repertoires. In addition, high frequency mutations in baseline human antibody repertoires were predicted *in silico* to reduce immunogenicity in therapeutic mAbs due to the removal of T cell epitopes. Several therapeutic mAbs were identified to have common, universally high-scoring framework mutations, and molecular dynamics simulations revealed the mechanistic basis for the evolutionary selection of these mutations. Our results suggest that baseline human antibody repertoires may be useful as predictive tools to guide mAb development in the future.

Keywords: monoclonal antibodies, antibody therapeutics, antibody repertoires, deep sequencing, protein stability, affinity maturation, somatic hypermutation, developability

INTRODUCTION

Monoclonal antibodies (mAbs) are now ubiquitous as therapeutics, with over \$100 billion in sales worldwide in 2020 (1) and applications ranging from oncology (2) and inflammation (3) to infectious diseases (4). mAbs are engineered not only to have potent and specific binding to a given target but also to have favorable drug properties, including *in vivo* stability, manufacturability, immunogenicity, solubility, and polyspecificity (5). Identifying highly developable mAb sequences *in silico* could greatly reduce the time and costs of therapeutic mAb development.

Antibody sequences sourced from baseline human antibody repertoires could inform our ability to engineer therapeutic mAbs by 'borrowing' consensus mutations (6, 7). This premise rests on the successful use of sequence conservation in protein engineering for improving the functional properties of enzymes (8–10), nanobodies (11), and membrane proteins (12). Antibodies in particular contain great potential for sequence optimization because every human body contains an estimated 10^{11} B cells with highly diverse antibody sequences (13), providing a rich space from which to glean important insights that could be used to guide future engineering efforts.

Using sequence conservation for improving antibody properties was first explored by Steipe et al. (14), who used known antibody sequences from the Kabat Database (15) to identify consensus positions within mouse V_{κ} repertoires. Mutation to the amino acids at these consensus positions resulted in improved thermodynamic stability for the majority of the antibody sequences tested. However, the power of any sequence-based method relies on the size of the database. It is now possible to sequence tens of millions of antibodies from a single individual. Studies evaluating such human antibody repertoires have focused on cataloging the immune response to vaccination or infection (16–19). Recently, the Great Repertoire Project conducted the most extensive attempt to sequence entire baseline human antibody repertoires to date, acquiring a total of 364 million antibody sequences by sequencing full Leukopaks from ten healthy, HIV-negative adults (20).

We revisited the idea that sequence conservation predicts developable antibody sequences using this much more comprehensive database of baseline human antibody sequences, applied towards the analysis of FDA-approved mAbs. Specifically, we sought to answer the following questions: (i) are there mutations from germline (GL) sequences that are highly prevalent in baseline human antibody repertoires and if so, are these also found in FDA-approved mAbs, given the generally favorable developability properties of these mAbs?; and more broadly (ii) can sequence information alone predict more developable from less developable mAbs? We restricted our analysis to the framework regions (FRs) of the variable heavy (V_H) domain as antibody FRs impact *in vivo* stability, solubility, and immunogenicity (6) while also contributing significantly less than complementarity determining regions (CDRs) for binding antigen. We also explored some of the dynamics of peptide-MHC-II interactions using computational binding

predictions (21), as the MHC-II peptide epitope contained within antibodies and other protein drugs has been recognized as an important component of clinical success (22, 23). As a result, sequence information for FR regions can be applied to a broad array of antibodies with varying applications.

In this study, we present position-specific substitution profiles (PSSM for position-specific scoring matrix) for antibody FR mutations using the most complete dataset of baseline antibody repertoire sequences to date (20). We show that antibody repertoire-based PSSMs are consistent across subjects and produce high correlations between GL V_H genes with expected differences based on sequence similarity and familial relationships. Our analysis shows that mutations in existing therapeutic antibodies can be accurately predicted solely from repertoire sequence data. We then quantitatively assessed entire FRs of these therapeutic antibodies and compared them to their baseline human repertoire counterparts. These data suggest that there may be potential for improving existing therapeutic antibody properties through incorporation of additional mutations of high-frequency in human repertoires. In addition, we found that high frequency repertoire mutations tended to reduce the affinity of germline-encoded peptides that bound to MHC-II epitopes. Several therapeutic mAbs have common, universally high scoring FR mutations, and simulations revealed a mechanistic basis for the favorable drug properties of some mutations. Overall, our results suggest that human antibody repertoires are useful as predictive tools that will facilitate engineering mAbs by improving drug-like properties.

METHODS

Selection of Human Repertoire Antibody Sequences

In our analysis, we considered only the FR regions of immunoglobulin G (IgG) V_H antibody segments. IgM sequences are the other common isotype included in the database but were ignored because they typically have low levels of mutation due to their role in the early stages of the immune response (24). D and J segments were excluded since these segments are highly variable, which would inhibit our ability to achieve significant coverage of all possible mutations at these positions.

Out of the 51 total V_H genes in the Great Repertoire Project database (20) (https://github.com/briney/grp_paper) (data accessed November 2020), we analyzed 25 V_H genes that represented inferred precursors for the human/humanized FDA-approved mAbs that we were able to collect sequences for. Additionally, these 25 V_H genes represent many of the most common V_H genes in this dataset. A minimum of 100,000 repertoire IgG variable heavy segments were extracted from the database and analyzed for each GL. The cutoff value of 100,000 was deemed necessary for the creation of reliable scoring matrices and was determined by a random subsampling analysis where score differences were compared across multiple sampling depths (**Figure S1**). It should be noted that the resulting

compiled sequences span several different alleles within each GL gene, though these differences are later reconciled by adjusting the scoring matrices to remove this bias, as described in detail in a subsequent section. All GL VH genes excluded from this analysis are in **Table S1**. GL VH genes were excluded either due to low sequence counts or if no FDA-approved mAb precursors were represented in the panel.

Selection of FDA-Approved Antibody Sequences

FDA-approved antibodies included in the analysis were selected according to three criteria: (i.) the sequences had to be human/humanized antibodies (name ending in -umab), (ii.) the sequences had to be available in the DrugBank database (25) at the time of sequence collection (accessed October 28th 2020), and (iii.) the Great Repertoire Project database (20) had to contain more than 100,000 sequences from the same inferred GL gene. It is critical that the analyzed antibody is human/humanized since this allows for the antibody to be matched to human GL genes.

In addition, to determine whether we can effectively identify developability issues in antibodies, we selected sequences for a second panel of engineered antibodies that were determined to have biophysical properties that fall outside of historically accepted limits. We used data from Jain et al. (5), which identified 12 biophysical properties that were indicative of an antibody's likelihood of progressing through all three stages of FDA approval. From this dataset, we selected all antibodies that were found to have two or more biophysical properties that fell outside of the described acceptable ranges. In addition, the antibodies had to be human/humanized and the Great Repertoire Project database had to contain at least 100,000 sequences in the repertoire database from the same inferred GL gene. IMGT BlastSearch v1.2.0 (26) was used to infer the GL gene for the V_H sequence of each of the analyzed FDA-approved antibodies, and the hit with the lowest E value was assigned as the GL V_H gene. As described below, position-specific scoring matrices (PSSMs) for each of the 25 GL genes were then constructed (**Table S1**).

Generation of Position-Specific Scoring Matrices for Human VH Genes

A multiple sequence alignment was performed on each set of sequences derived from a common GL gene using a lightweight version of the MAFFT algorithm (27) designed for large numbers of similar sequences and using the FFT-NS-2 progressive alignment method in low-memory mode with the gap opening penalty set to max (5.0). All other alignment parameters were set to default. Each alignment file was collapsed into a single table of mutational counts by tallying the number of observations of each amino acid at each position in the sequence. This count output was modified to remove instances of insertions, indicated by positions that contained under 10% occupancy (i.e., more than 90% gaps). In addition, individual counts of wild-type (WT) and allelic variants were removed (**Table S4**). Amino acid substitutions with zero counts were replaced with a pseudocount of one to circumvent an undefined score (acting

as a lower bound for scores). Then, the tables were manually aligned to IMGT human GL V_H FR reference sequences (26) to associate each column of the count table with a real position in the FR sequence.

The counts for all amino acid substitutions within the FR regions were log transformed into a score *via* the following equation:

$$S_{ij} = \log_{1.26} \left(\alpha \cdot \frac{N_{ij}}{\sum_{j=A}^Y N_{ij}} \right) \quad (1)$$

Here S_{ij} is the score of amino acid substitution j at position i and α is a V_H-specific constant (Range: 726-2228, s.d.: 320) that centers the mean of the score distribution for V_H-specific FR mutations at zero. For comparisons across GL families, this α normalization term results in a max difference of 4.9 between GL with a difference of <1.0 at 1 s.d. N_{ij} is the count of amino acid substitution j at position i , the sum of which, over all amino acids from alanine (A) to tyrosine (Y), is equal to the total number of sequences without gaps in the multiple sequence alignment at that position. We used a log base of 1.26 to adhere to previous convention used by Dayhoff et al. for log-odds matrices (28). PSSMs for the 25 analyzed GL VH genes can be found in **Data S1**. This same protocol was then applied to the FDA-approved antibody sequences, which were first aligned to their corresponding GL sequences and then nonsynonymous FR mutations and their respective scores were recorded.

Generation of Framework Scores for Individual Antibody Sequences

We then derived an overall FR region score that we term an "FR score", a metric which can be used to compare sequences of antibodies with different numbers of mutations across different GL VH genes. We assumed that each FR mutation has an additive effect and is independent of one another (no epistatic interactions). With this assumption, the FR score is then defined as:

$$FR = \begin{cases} 1 & \text{for } m = 0 \\ \frac{\sum_{k=1}^m S_{ijk}}{c_0 m + c_1 m^2} & \text{for } m > 0 \end{cases} \quad (2)$$

where S_{ijk} is the score of the k^{th} sequential FR mutation of amino acid substitution j at position i as determined by Eqn. 1. The sum of m (the number of FR mutations) mutation scores is normalized by constants c_0 and c_1 which are specific to that GL gene (**Table S3**), derived from a least squares regression. The denominator of the FR score is a normalization constant which gives an estimate of a typical score of a randomly sampled Ab from the repertoire with the same number of FR mutations (**Figure S6**). The FR score as derived gives an estimation of how different an antibody's framework mutations are as compared to baseline human repertoire antibodies. As such, if an antibody has framework mutations with scores similar to what we would expect of a human antibody with an equivalent number of mutations, this ratio becomes one. Thus, we assign all antibodies with no framework mutations an FR score of one.

Prediction and Analysis of MHC-II Peptide Epitope Affinity

MHC-II peptide epitope affinity prediction was carried out for a set of 38 representative HLA-DRB1 alleles (21) using netMHCIIpan 4.0 (29). Predictions were performed using 15-mer peptide scans for each peptide containing mutations with an associated PSSM, and the netMHCIIpan flag -BA was selected to obtain predicted MHC-II binding affinities. Predictions were also carried out for germline (GL) peptides using the same settings, for the 38 representative HLA-DRB1 alleles. Next, netMHCIIpan predictions for FDA-approved mAb peptides were matched to GL peptides and their own binding predictions. To track affinity changes from high affinity GL peptides, peptides matching a germline peptide with peptide: MHC-II $K_D < 1,000$ nM were assigned to three different bins based on the ratio of (mutated peptide K_D /GL peptide K_D). This was done for K_D s predicted for the same allele. Three bins were created: K_D fold-change < 0.5 , $0.5 \leq K_D$ fold-change ≤ 2 , and K_D fold-change > 2 . We also considered the total set of peptides matching any high affinity GL peptides ($K_D < 1,000$ nM), and the total set of peptides matching any low affinity GL peptide ($K_D > 1,000$ nM). Next, PSSM values were averaged by DRB1 allele and K_D group using standard mean. This was done to reduce redundancy, since the same peptide can have different affinities for different HLA-DRB1 alleles and thus be in different K_D groups at the same time. As a result, average PSSMs per DRB1 and K_D group take into account peptide contributions based on differential DRB1 binding. Next, PSSM distributions were compared between K_D bins using a Welch Two Sample t-test, and p-values were adjusted for multiple comparisons using the Benjamini-Hochberg method. All data processing steps were carried out in R.

Molecular Dynamics Simulations

Molecular dynamics (MD) simulations were carried out using the GROMACS 2021 (30, 31) MD engine and the TIP3P (32) and AMBER99SB-ILDN (33) force fields to explicitly model water and the antibodies, respectively. Simulations were initiated from crystal structures of the following antibodies: Atezolizumab (PDB code 5X8L) (34), Daratumumab (PDB code 7DUN) (35), and Omalizumab (PDB code 4X7S) (36), with all antigens/ligands removed beforehand. Additional simulations were performed with a single germline reversion mutation incorporated into each mAb sequence (A54S for Atezolizumab and Omalizumab, and F103Y for Daratumumab), using Visual Molecular Dynamics (VMD) (37). Counterions in the form of Na^+ or Cl^- , also described by the AMBER99SB-ILDN force field, were added to the systems as needed to neutralize any net charge. Each system contained between 67,000 and 94,000 atoms.

The initial coordinates for each system were minimized for a maximum of 5,000 steps using a steepest descent energy minimization. This was followed by 0.5 ns each of NVT and then NPT equilibration simulations performed at 310 K using the Bussi–Donadio–Parrinello thermostat (38), and at 1.0 bar using the Berendsen barostat (39) (same temperature and thermostat) for the NPT simulations. The time constant for

coupling in both the NVT and NPT simulations was 0.1 ps. NPT production simulations were then performed at the same temperature and pressure and using the same thermostat and Parrinello–Rahman barostat (40). Particle mesh Ewald (PME) summations (41) were used to calculate long-range electrostatic interactions with a cutoff of 1.0 nm, and Lennard Jones interactions were calculated over 1.0 nm and shifted beyond this distance. Neighbor lists were updated every 10 steps with a cutoff of 1.0 nm and all simulations utilized full periodic boundary conditions. Production simulations were carried out for 0.3 μs each, for a total of 1.8 μs across the six simulations.

Standard GROMACS tools were used to compute the root mean squared deviation (RMSD) of the antibodies from their respective energy-minimized structures as a function of simulation time. For all RMSD plots, data was plotted every 10th frame. We also computed the root mean squared fluctuation (RMSF) of the heavy chains of the antibodies over the last half (150 ns) of the simulations, during which time all six simulations were deemed converged based on RMSD. All simulation images were rendered in VMD.

Statistical Calculations

Unless otherwise noted, all p-values were calculated using Welch's t-test. Correlations were analyzed using a Pearson product-moment correlation coefficient. Hierarchical clustering for GL comparisons was performed using SciPy in Python (www.scipy.org) using the unweighted pair group method with arithmetic mean (UPGMA).

RESULTS

Generation of Unbiased Framework PSSMs From Antibody Repertoires

We created an efficient method to score the set of individual FR mutations in an FDA-approved mAb, which is shown schematically in **Figure 1** and is carried out as follows. First, we identified the set of human or humanized FDA-approved mAbs for which sequences were available in the DrugBank database (25). From this list, we used IMGT BlastSearch (26) to infer a GL V_H gene for each mAb (**Table S1**). We restricted our analysis to the FR positions of V_H only (IMGT Numbering FR16/24/25-26 depending on the V_H ; FR39-55; FR66-104), as the Great Repertoire Project reported only heavy chain sequences. Additionally, we considered only IgG sequences, as IgM sequences typically have low levels of mutation due to their role in the early stages of the immune response (23).

The collected IgG sequences from the Great Repertoire Project come from the eight individuals available on GitHub (accessed November 2020); our analysis of this dataset shows that the extent to which a B cell population is mutated varies widely across V_H genes and individuals (**Tables S2** and **S3**, respectively). Our expectation is that this dataset is composed of a minimally-biased representation of the baseline human antibody repertoire that includes both naïve and mature B cell sequences contained in the peripheral circulation. This

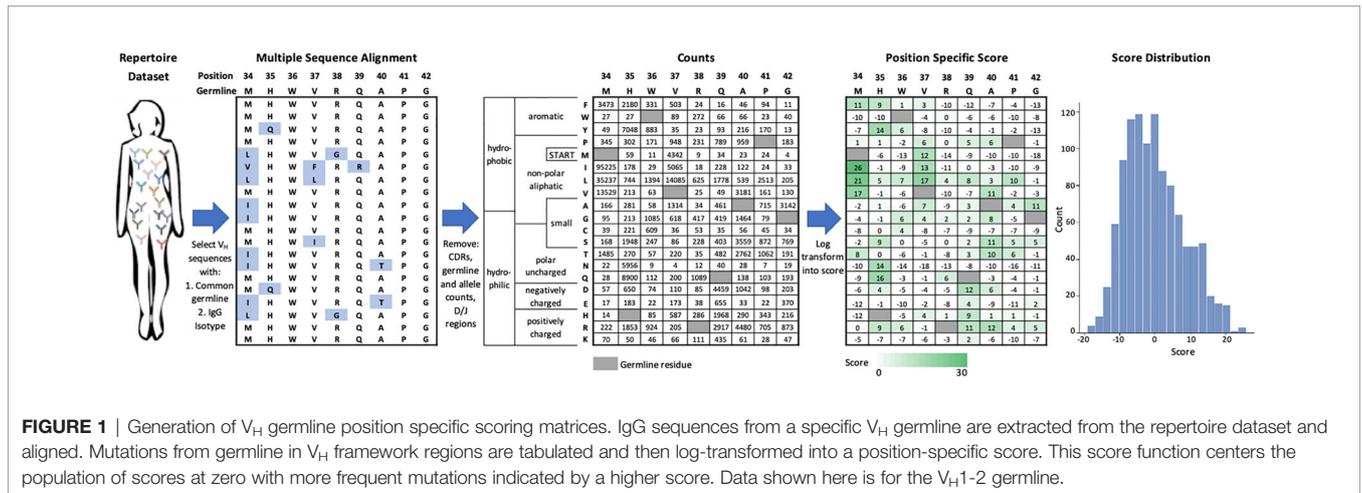


FIGURE 1 | Generation of V_H germline position specific scoring matrices. IgG sequences from a specific V_H germline are extracted from the repertoire dataset and aligned. Mutations from germline in V_H framework regions are tabulated and then log-transformed into a position-specific score. This score function centers the population of scores at zero with more frequent mutations indicated by a higher score. Data shown here is for the V_H1-2 germline.

expectation is informed by an analysis of the dataset, which reveals an average mutation frequency per position of 9.8%. Additionally, we analyzed the entire set of IgG sequences for synonymous (dN) and synonymous (dS) substitutions. Taking the ratio of the total number of substitutions for each of these we observe an overall dN/dS ratio of 2.87, indicating that these sequences have undergone targeted mutagenesis and selection. Further evidence for mature B cell sequences exposed to antigen is found in comparison of mutation rates for IgG vs. IgM sequences. Of the collected IgG sequences, 17.4% were found to have >98% nucleotide identity with their corresponding germline precursor and 3.3% are unmutated. These numbers are considerably lower than those of the collected IgM sequences, for which 54.6% were found to have >98% nucleotide identity with their corresponding germline precursor and 14.0% are unmutated (Table S3). Another piece of evidence for this repertoire including B cells exposed to antigen result from examination of the mutational rate of the CDRs relative to FR. The frequency of mutation by position is plotted in Figure S2 and shows an average mutation frequency of 18.8% in the CDRs versus 7.4% in the FRs.

We aligned a set of relevant IgG sequences by multiple sequence alignment (27) in order to generate a FR position-specific scoring matrix (PSSM) for each GL V_H gene considered (see Methods). For each FR position, we counted the number of sequences observed to have a mutation at this position and log-transformed this frequency to a score term. This scoring term has one adjustable parameter per GL V_H gene such that the mean score for a randomly selected mutation would be zero. According to how we defined this scoring term (see Methods), higher scores represent more commonly observed amino acid mutations and the gross majority of scores range between -10 and 10.

One complication arising from statistical analysis of antibody mutations is that humans possess significant allelic diversity in V_H genes, even in FR regions. For example, consider V_H3-9 which has a T99M FR allelic variation (Table S4). In generating a PSSM at position 99, one could identify the appropriate allele for each patient (either T or M) and then generate separate PSSMs for each allelic variation. Alternatively, we chose the simpler

approach of removing all allelic variants from our PSSM since these only affect a small minority of possible substitutions.

The actual number of IgG GL V_H-specific sequences in the Great Repertoire Project database varies tremendously, from around 2,500 sequences (V_H1-45) to more than 2 million sequences (V_H3-7). Given this range, we asked what number of sequences would be sufficient for accurate recapitulation of GL PSSMs. We chose to perform this analysis on V_H5-51 by subsampling between 1,000 and 300,000 sequences. For each mutation, we computed the absolute difference between the score derived from the full dataset (~700,000 sequences) and the score from the subsample. Increasing the number of sequences from 1,000 to 25,000 led to increasing agreement with the full dataset, with essentially no difference observed in the median displacement at or above 25,000 sequences for the large majority of scores (i.e., scores greater than -15) (Figure S1). However, minor differences in displacement at 2σ (95%) were observed between 25,000 and 100,000 sequences. Thus, we conservatively restricted our analysis of GL V_H segments to those with 100,000 or more IgG sequences present in the Great Repertoire Project database. The GL V_H genes analyzed and resulting sequence counts are denoted in Table S1.

We then asked whether mutation frequencies would be highly correlated between individuals and if the site-specific preferences would be repeatable. To test this, we generated patient-specific V_H3-23 PSSMs for all 8 individuals for which data was available which contained the same V_H3-23*01 allele. The V_H3-23 GL was selected to ensure that a sufficient number of sequences was available for generating these PSSMs. We observed a high degree of correlation between patient scores with correlation coefficients between 0.86 and 0.91 for all pairwise comparisons (Figure 2). We then repeated our analysis considering only amino acid mutations encoded by 1-nucleotide (nt) substitution, as these points contain the highest sequence coverage in our dataset and, therefore, are presumably most accurate. The results were essentially unchanged, with pairwise correlation coefficients between individuals ranging between 0.84 and 0.92 (Figure S3A). These correlation coefficients are only slightly below the basal noise level (0.93-0.94) calculated by computing

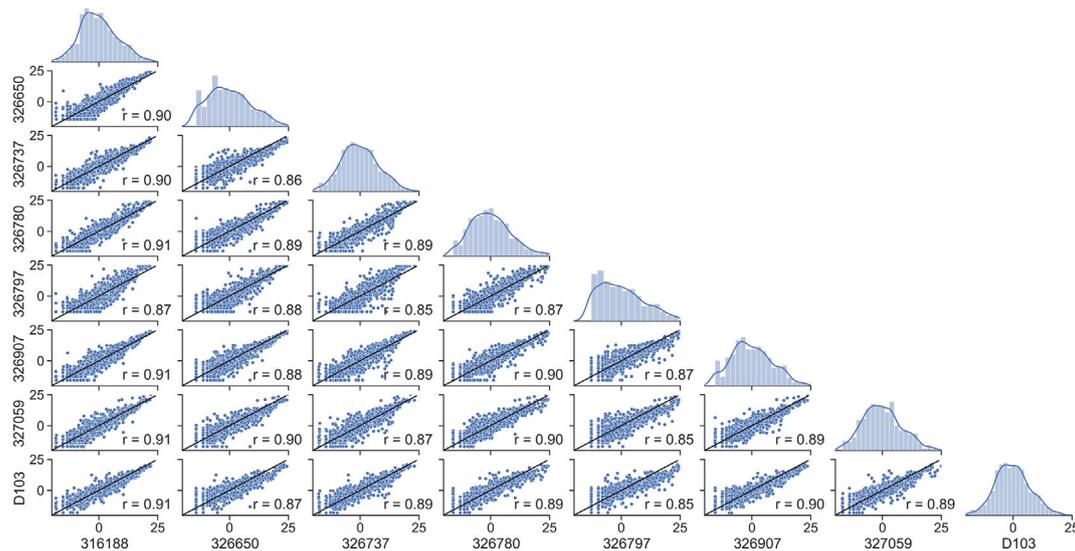


FIGURE 2 | Framework scores are repeatable between individuals. Distributions and correlations of patient-specific PSSM scores for the V_H -3-23*01 germline. Alphanumeric characters on the y- and x-axes represent de-identified subjects.

the correlation coefficient between two independent random samples for each of two patients (Figure S3B). These results show that GL V_H -specific FR mutations are largely repeatable between individuals. This finding is not new, as Sheng et al. showed in their analysis that affinity maturation somatic hypermutation (SHM) seems to produce highly consistent substitutions for antibody V segments across different donors, even in the absence of functional selection (16). However, our independent analysis on a unique dataset corroborates their results, which together suggest underlying biophysical mechanisms are responsible for the reproducibility of the selection of certain amino acid mutations at different FR positions. These results strongly suggest that the process of affinity maturation results in replicable frequency distributions of FR mutations at the amino acid level.

Such replicable frequencies can arise if a given FR mutation is selected during affinity maturation for a general, antigen-independent reason. For example, a given FR mutation may improve the surface density of B cell receptors or could remove an aggregation-prone patch on the surface of the protein. An alternative but non-exclusive explanation is as follows. Sequences from baseline human antibody repertoires tend to be enriched in mutations which can be generated by a single nucleotide substitution, while mutations that would require di- or tri-nucleotide substitutions tend to occur less frequently. This is because activation-induced cytidine deaminase (AID) primarily generates single nucleotide point mutations in B cell receptor sequences during affinity maturation (42), and as this mutational frequency is $\sim 10^{-3}$ per base per generation (43), it is far less probable that the same codon will be targeted twice or thrice than once. Consistent with this notion, analysis of our PSSMs shows that FR mutation scores that are one nucleotide away from the GL sequence are significantly higher (mean score of 8.42) than

those requiring two or three substitutions (mean scores of -2.37 and -8.18, respectively) (Figure S4).

FDA-Approved mAbs Contain Framework Mutations That Can Be Predicted From Human Antibody Repertoires

We asked whether FR mutations in FDA-approved mAbs could be predicted from native antibody repertoires. 39 FDA-approved human/humanized mAbs were found to have sufficient depth of coverage (i.e., > 100,000 sequences) in their GL V_H family in the Great Repertoire Project database and were thus subjected to further analysis (Table 1). While the precise developmental pathways for all of these FDA-approved mAbs are not publicly available, it is known that most of the mAbs were discovered using either hybridoma technology, transgenic mice, or phage display.

On average, the 39 FDA-approved mAbs have a mean of six V_H FR mutations, and a range from zero to 16 FR mutations. The degree to which these mutations might align with (and thus may be predicted by) those found in human antibody repertoires is difficult to know *a priori*. On the one hand, the use of phage display or transgenic mice could result in a sufficiently different selection environment than what exists during affinity maturation in human germinal centers, leading to an altogether different set of FR mutations than are evolved in human antibody repertoires. On the other hand, the clinical approval process eliminates any antibodies that may have been selected ineffectively for drug-like properties and failed in manufacturing or pre-clinical and clinical trials. The affinity maturation process may indirectly select for some of the same qualities of antibodies that have gone through regulatory approval, including high expression yield, thermodynamic stability, low non-specific binding, little to no aggregation

TABLE 1 | Statistical significance for framework mutations for FDA-approved mAbs.

Germline Gene	FDA-approved Antibodies	Number of V _H Framework Mutations	p-value	Number of Repertoire Sequences Analyzed
VH1-2	Pembrolizumab	16	7.0E-06	300,000
VH1-3	Vedolizumab	9	4.7E-02	129,913
VH1-46	Benralizumab	24	1.1E-06	451,920
	Ravulizumab			
	Burosumab			
VH1-69	Risankizumab	23	1.8E-05	300,000
	Ixekizumab			
	Galcanezumab			
VH2-5	Palivizumab	7	5.1E-03	141,372
VH3-7	Secukinumab	10	7.4E-05	2,003,693
	Fremanezumab			
	Durvalumab			
VH3-9	Adalimumab	9	7.6E-07	403,612
	Ofatumumab			
	Sarilumab			
VH3-23	Pertuzumab	38	9.3E-07	300,000
	Denosumab			
	Daratumumab			
	Avelumab			
	Dupilumab			
	Emicizumab			
	Lanadelumab			
	Atezolizumab			
VH3-30	Ipilimumab	5	1.3E-04	300,000
	Erenumab			
VH3-33	Canakinumab	4	2.1E-02	710,048
	Nivolumab			
VH3-48	Certolizumab pegol	11	2.1E-03	300,000
VH3-66	Omalizumab	27	7.6E-08	321,361
	Trastuzumab			
	Eptinezumab			
VH3-74	Efalizumab	26	7.7E-05	1,355,540
	Elotuzumab			
	Atezolizumab			
VH4-4	Alemtuzumab	7	2.2E-04	1,003,945
VH4-30-4	Tocilizumab	11	2.6E-06	160,757
	Necitumumab			
VH5-51	Ustekinumab	7	2.6E-05	702,327
	Guselkumab			

propensity, and low immunogenicity (44, 45). In this latter case, FR mutations in FDA-approved mAbs could indeed be predicted by baseline antibody repertoire abundance.

To answer this question, we used our GL V_H-specific PSSMs to score each FR mutation contained in the FDA-approved mAbs. Strikingly, every one of the 16 GL V_H families that form the set of inferred precursors for these mAbs showed significantly higher (at 95% confidence level) FR mutation scores for FDA-approved mAbs relative to the set of all possible mutation scores (Figures 3A, B, and S5, Table 1). The statistical significance of this result ranged from a p-value of 1.8e-5 (V_H1-69) to 7.6e-8 (V_H3-66) (Figure 3B). However, the distributions in Figure 3B of the set of all possible mutation scores do not reflect the number of repertoire antibody FR mutations with given scores. Indeed, generating random samples of scores for 1,000 repertoire antibody sequences from each of these distributions shows that the majority of observed antibody mutation scores in human repertoires are above 10 (Figure 3C). Plotting the individual mutation scores for the

FDA-approved mAbs on top of these new distributions, we see that the distributions match much more closely, with scores of the FDA-approved mAbs now slightly below those of the baseline antibody repertoires (Figure 3C). The slightly lower scores of the FDA-approved mAbs largely result from the fact that in each case, there are at least two mutations below a score of zero, and for the entire dataset of FDA-approved FR mutations 10.3% had a score below zero. These results suggest that while repertoire datasets can predict mutations in mAb V_H FR regions with reasonable accuracy, as indicated by the large overlap in the distributions, there is also considerable room for engineering mAb sequences to mimic baseline human repertoire antibodies with increased fidelity.

FDA-Approved mAbs Have Lower FR Scores Than Typical Repertoire Antibodies

Motivated by the previous results, we wondered how the cumulative effect of all FR mutations for a given FDA-approved mAb compares to that of baseline repertoire

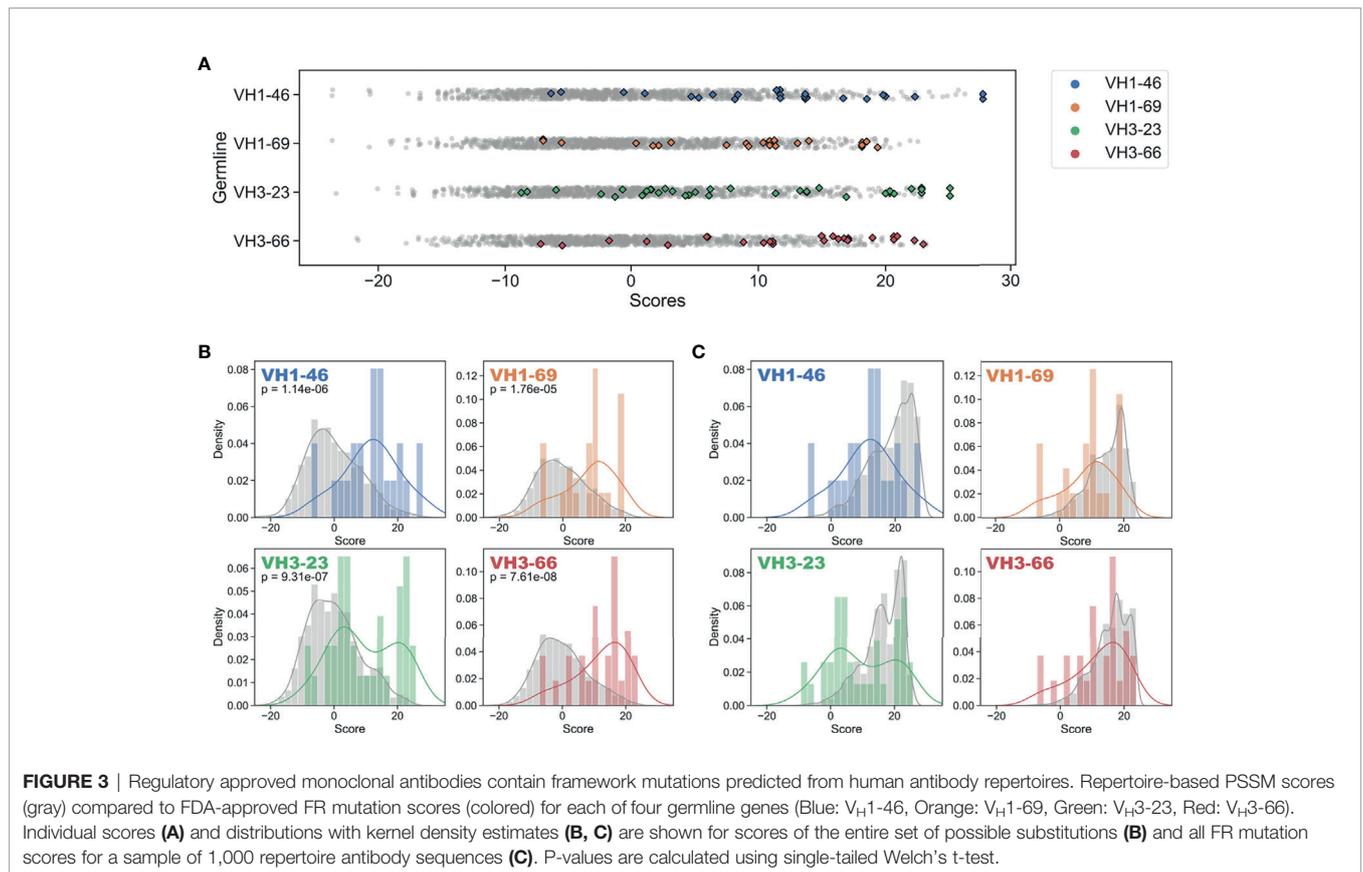


FIGURE 3 | Regulatory approved monoclonal antibodies contain framework mutations predicted from human antibody repertoires. Repertoire-based PSSM scores (gray) compared to FDA-approved FR mutation scores (colored) for each of four germline genes (Blue: V_H1-46 , Orange: V_H1-69 , Green: V_H3-23 , Red: V_H3-66). Individual scores (A) and distributions with kernel density estimates (B, C) are shown for scores of the entire set of possible substitutions (B) and all FR mutation scores for a sample of 1,000 repertoire antibody sequences (C). P-values are calculated using single-tailed Welch's t-test.

antibodies. To assess this, we assumed, for simplicity, that individual mutations are independent, such that the cumulative effect of FR mutations is best represented as a sum of individual mutation scores. Due to the V_H -specific centering parameter included in the score function, individual mutation scores cannot be directly compared across V_H families. To account for these differences, the score sum was normalized by a regression function which estimates the typical score sum of a randomly sampled repertoire antibody from the same V_H family with the same number of FR mutations (Figure S6 and Table S5). The resulting “FR score” is defined such that a score of one represents an antibody that has FR mutations similar to that of a typical repertoire antibody framework sequence (see Methods). Antibody sequences with no FR mutations were assigned a pseudo-score of one since the calculated FR score is otherwise undefined.

We then compared FDA-approved mAbs to repertoire antibodies using FR scores. A simple random sample with replacement of 1,000 repertoire antibody sequences across all GL genes was generated and an FR score for each sampled antibody was computed (Data S2). As expected, FR scores for the repertoire antibodies center at a value of one (Figure S7). Next, FR scores were calculated for all FDA-approved mAbs used in the analysis above and the results are plotted along with the repertoire averages in Figure 4A as a function of the number of FR mutations. All individual mutation scores and FR scores for FDA-approved mAbs are tabulated in Table S6.

Thirty of the 39 FDA-approved mAbs have FR scores below the baseline repertoire antibody average of one (p -value $1.4e-4$). Figure 4A highlights the five mAbs with the highest and lowest FR scores. Denosumab and Dupilumab are the two highest scoring FDA-approved mAbs with FR scores of 1.45 and 1.23, respectively. Both mAbs have a small set of apolar/polar FR mutations (A55G, A25G, S40T, A55S), and interestingly, both mAbs share the same V_H3-23 GL and were produced from transgenic mice. Nivolumab is the lowest scoring FDA-approved mAb with an FR score of 0.12 and only one mutation requiring a 2-nt change from the GL sequence (A24K: score=2.0). Emicizumab and Pertuzumab, both from the V_H3-23 GL and both developed using hybridoma technology, also have low FR scores of 0.32 and 0.26, respectively. However, the low scores of these mAbs are only somewhat explained by the fact that 14 mutations (out of 17 total) are more than 1-nt away from the GL codon. The solved complex of Pertuzumab and human epidermal growth factor receptor 2 (ErbB2 or HER2) (46) partially reveals the source of Pertuzumab’s low FR score. The complex is unusual in that FR positions surrounding CDRH2 are directly contacting ErbB2, including mutations for Y66I, A68N, D69Q, and N82R. Thus, rare FR mutations are necessary in this specific mAb for Erb2 recognition.

As shown above, the methods used to develop mAbs can influence the selective pressure and, therefore, the mutations that are incorporated into therapeutic mAbs. On average, we find that the number of FR mutations in FDA-approved mAbs discovered

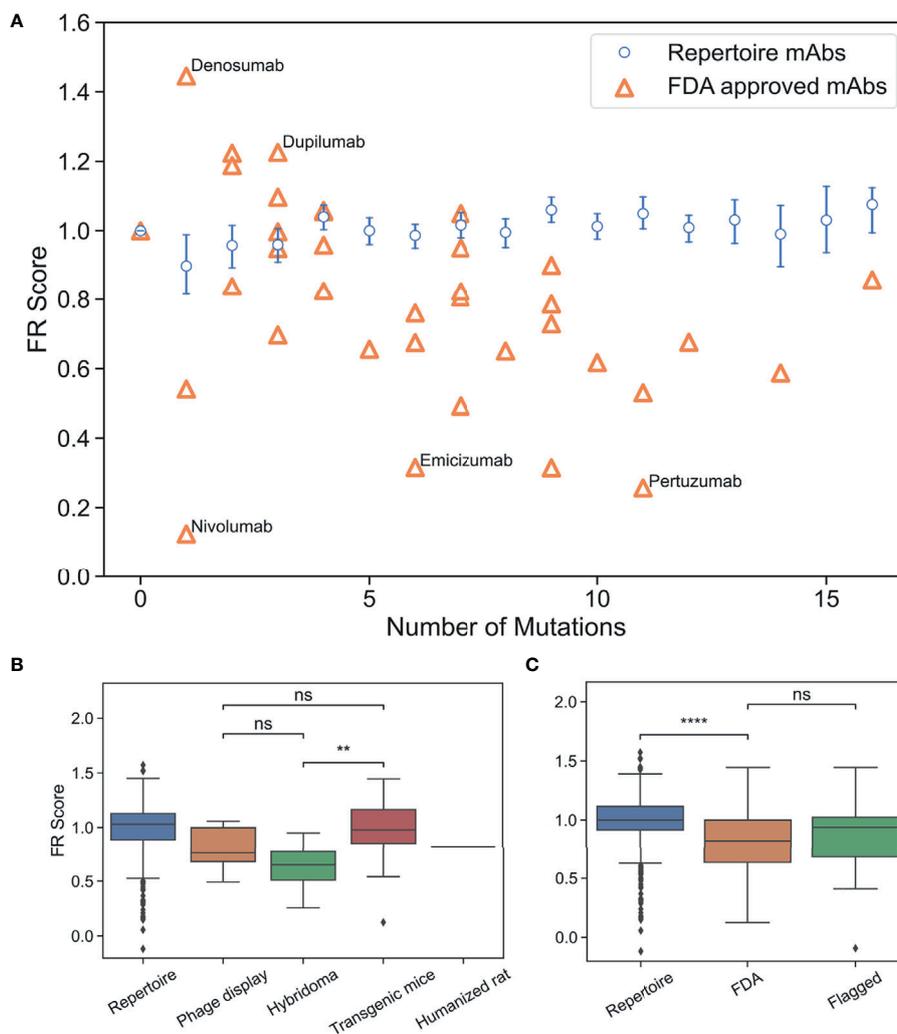


FIGURE 4 | FR scores of FDA-approved and flagged mAbs are lower than typical human antibodies. **(A)** Comparison of FR scores for repertoire Abs and FDA-approved mAbs. Error bars on FR scores represent 95% confidence intervals. **(B)** FR scores of FDA-approved mAbs by development method. **(C)** FR scores of repertoire, FDA-approved, and flagged mAbs. (ns, $p > 0.05$, ** $p < 1e-2$, **** $p < 1e-4$).

by hybridoma technology is higher than in FDA-approved mAbs discovered by transgenic mice or phage display (p -value $\leq 4.6e-5$) (**Figure S8**). Even though FR scores are independent of number of FR mutations, we find that mAbs developed using hybridoma technology scored significantly lower than transgenic mice (p -value $5.5e-3$), although there is not a statistically significant difference between mAbs developed *via* hybridoma *versus* phage display (p -value 0.13) (**Figure 4B**). Phage display and transgenic mice technologies produce mAb FRs with similar FR scores (p -value 0.29). Thus, for existing mAbs, those produced by transgenic mice are the most representative of the baseline human antibody repertoire in terms of their FR mutations.

We next questioned whether FR scores could be used to diagnose mAb developability issues. Collecting an appropriate dataset for this analysis is difficult, as there are few databases of therapeutic antibodies with developability issues available in the

publicly accessible literature. As a proxy, we chose to use the set of “flagged” antibodies originally described in Jain et al. (5). Jain et al. examined 12 different biophysical properties of a set of mAbs in various stages of clinical trials, with outliers in any of these properties considered “flags”. Two or more of these flags identified a mAb as a developability risk. We compiled all human/humanized mAb sequences from this analysis that contained at least two flagged properties and sufficient repertoire sequence data (39 in total). It should be noted that five of the flagged Abs that we analyzed have since received FDA approval and are therefore also contained within the FDA-approved dataset. FR scores were calculated for flagged mAbs and compared to our previously described panel of FDA-approved mAbs (**Figures 4C, S9** and **Table S7**). Six of the flagged mAbs contained no FR mutations. As with the FDA-approved mAbs, we see that flagged mAbs also contain

significantly lower FR scores than a population of randomly selected repertoire Abs (p-value 7.6e-3). However, there is no statistically significant difference between the distribution means (p-value 0.37) for regulatory-approved vs. flagged mAbs. Thus, our FR score is too coarse-grained to differentiate between antibody groups at such a late stage of development.

Biophysical Basis of Highly Prevalent FR Mutations

Why is the frequency of a given FR mutation largely repeatable between individuals? It could be the case that AID preferentially encodes given FR mutations due to sequence hotspots (42, 47), with minimal selection of the amino acid sequence. Alternatively, there may be underlying biophysical bases for the selection of certain sequences. However, uncovering the rationale for every FR mutation is challenging as any individual mutation is pleiotropic, thus impacting many qualities *in vivo* including aggregation propensity, proteolytic stability, B cell receptor expression, antigen-binding affinity, and polyspecificity. Beyond these traits, it has also been hypothesized that engineering antibodies with FR regions derived from GL gene-preferred substitution patterns would result in lower immunogenicity as these antibodies would more closely reflect native antibody responses, which are readily recognized by the immune system (6, 21). As an initial effort, we decided to focus on two properties which could reasonably explain the molecular basis behind these high scores: stability and MHC-II peptide epitope content. These biophysical properties can be addressed sufficiently using *in vitro* biochemical assays or *in silico*.

Other research groups have published the effects of FR mutations on stability and heterologous expression yields (Table 2) (48–51). Across these four studies, thermodynamic stability was measured (as $\Delta\Delta G$, ΔT_{50} , ΔT_m) for nine point mutants. All seven point mutants that required a 1-nt substitution and had an individual mutation score above zero. These seven showed neutral or improved stability. Additionally, the two mutants with the highest mutation scores (V_H1-69: V76L; V_H6-1: S74G) also showed the highest thermodynamic

stability (measured as ΔT_{50}). Together, these results hint at a relationship between stability and repertoire abundance of a given FR mutation, although the small size of this dataset does not allow for sufficient statistical power to definitively draw a link between these factors.

FR mutations could also be selected to decrease the MHC-II epitope content of a given antibody sequence. Previous research has shown that unmutated antibody sequences have germline-encoded peptides that bind to and are presented by MHC-II proteins. These peptide epitopes are targeted for removal by SHM during B cell development *in vivo* (21). We sought to determine whether individual FR PSSM scores calculated here correlate with differential peptide:MHC-II binding affinities. To address this question, MHC-II binding affinities (K_D) for all 15-mer peptides in a sliding window (i.e., fragments of 15 consecutive amino acids) sourced from the 39 FDA approved mAbs were calculated using a custom computational pipeline (see *Methods*). For each peptide containing FR mutations, mutation scores were binned by the K_D fold-change from the corresponding unmutated peptide's K_D . Interestingly, peptides with $> 2 K_D$ fold-change from their unmutated germline peptides showed higher PSSM scores than mutations with $< 0.5 K_D$ fold-change, and these data were statistically significant. This analysis suggests that prevalent FR mutations are more likely to decrease the affinity of mAb peptide fragments for the MHC-II groove than to increase the peptide:MHC-II affinity at a given mutation site; these data were consistent with prior reports of human repertoire antibodies using paired heavy and light chain human sequence datasets (21) (Figure 5). As expected, mutations with moderate or no change in binding ($0.5 \leq K_D$ fold change ≤ 2) showed the highest PSSM scores across all groups because the majority of mutations do not affect the ~5-8 heavy chain variable region MHC-II peptide epitopes for a given MHC-II gene. Strikingly, we found that the high-affinity MHC-II peptide epitopes showed higher PSSM scores than low-affinity MHC-II peptide epitopes, and the differences were statistically significant. Together, these data demonstrate that high-affinity MHC-II peptide epitopes inside the antibody heavy chain variable

TABLE 2 | Differential stabilities for individual framework point mutants.

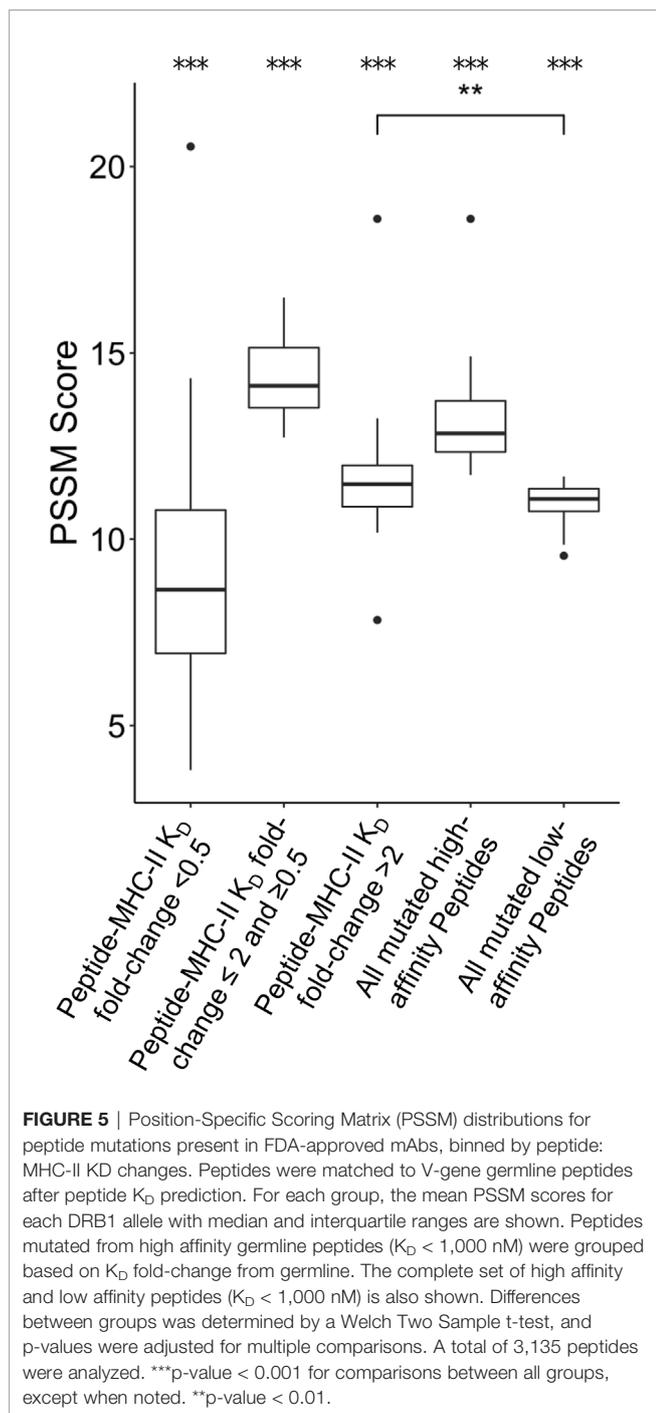
VH Gene	Mutation (paper)	Mutation (IMGT)	Minimum nt	Score	Property*	Reference
VH3-23	Q81H	Q90H	1	16.1	Mean $\Delta T_m = 0.0^\circ\text{C}^\ddagger$	(48)
	A84P	A96P	1	15.1	$\Delta T_{50} = 0.7^\circ\text{C}$	(49)
VH1-69	M48G	M53G	2	-1.4	$\Delta T_{50} = 3^\circ\text{C}$	(50)
	M48I	M53I	1	11.4	$\Delta T_{50} = 2^\circ\text{C}$	
	V67I	V76I	1	14.6	$\Delta T_{50} = 4^\circ\text{C}$	
	V67L	V76L	1	19.1	$\Delta T_{50} = 7^\circ\text{C}$	
VH6-1	V72D	V69D	2	10.5	320% ^a , 180% ^b yield	(51)
	S76G	S74G	1	26.8	0.1 ^a , 2.2 ^b kJ/mol stability	
	S90Y	S88Y	1	4.4	210% ^a , 150% ^b yield	
					3.7 ^a , 3.5 ^b kJ/mol stability	
					180% ^a , 230% ^b yield	
					-0.1 ^a , 1.4 ^b kJ/mol stability	

[‡]Mean ΔT_m was calculated over all genotypes presented in Table 3 in ref (48).

*Yield is calculated as a relative percentage to the original antibody sequence.

^aAntibody 2C2.

^bAntibody 6B3.



region are preferentially targeted for mutations, in both human repertoires and in clinically approved mAbs (Figure 5).

Position-Specific Framework Mutation Scores Are Highly Correlated Between Germline Families

It has been established that affinity maturation produces antibodies that have GL-specific substitution patterns (16). What is the exact correlation of substitutions between GL V_H genes? To address this

question, we calculated the Pearson correlation coefficient between each pair of GL PSSMs (Figure 6). Overall, correlation coefficients were quite high between GL PSSMs, ranging from 0.52 to 0.93. Hierarchical clustering was performed, revealing that correlations within subfamilies were higher than those between subfamilies. For example, all of the V_{H1} GL members clustered tightly together, as did those in V_{H3} , and in V_{H4} . The resulting clusters may, to an extent, be reflective of sequence similarity: GL members within a subfamily will have a higher pairwise amino acid identity than between subfamilies, which could skew the results. To account for this, we also calculated Pearson correlation coefficients only between GL PSSMs for which a common DNA codon is shared (Figure S10). As expected, higher correlations were observed between subfamilies than before, but the overall clustering within subfamilies was conserved. Thus, FR mutations are largely, though not wholly, repeatable between GL families and the correlation is stronger within subfamilies than between subfamilies.

Certain Highly Prevalent FR Mutations Are Observed Across Many Germline Families

We next looked at ‘universal’ FR mutations, or highly abundant mutations that are seen across all subfamilies considered. 18 mutations at 14 FR positions were enriched across all GL members scanned in this work (Table S8). Overall, ‘universal’ mutations were largely chemically similar and/or polar substitutions at positions predominantly on the protein surface and distal to the CDRs. ‘Universal’ mutations also tended to be the WT residue for at least some of the GL families. For example, serine at FR85 is highly prevalent across GL families and is the

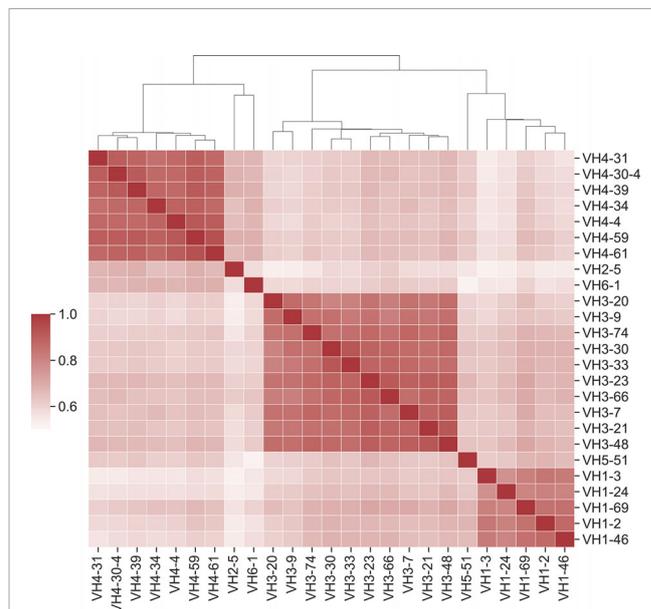


FIGURE 6 | Position-specific framework mutation scores are highly correlated within VH germline gene families. Pearson correlation coefficients were calculated for each pair of PSSM scores. Germline families are grouped by hierarchical clustering of Pearson correlation coefficients. Dendrogram indicates similarity between germline members.

WT residue for several V_H GL families (V_H1-2, V_H1-3, V_H1-46, V_H1-69, V_H5-51).

We noticed a specific mutation (S54A) that shows up in several FDA-approved mAbs (Pertuzumab, Atezolizumab, Omalizumab, and Trastuzumab) with high scores across multiple GL V_Hs (V_H3-23 and V_H3-66 with scores of 20 and 17, respectively). This mutation, which was observed at a frequency of 8.1% in the analyzed VH3-23 sequences, has also been observed at similar

frequencies in another antibody evolution study (19). **Figure 7A** shows the predicted scores for the mutation of residue 54 to five different amino acids (A, C, G, S, and T) across all 25 GL VH genes studied. We observe that this specific mutation also shows up in multiple other GL V_Hs for which there is no corresponding FDA-approved mAb (V_H3-20, V_H3-21, V_H3-48, V_H3-74, and V_H3-9, with scores ranging from 20-21). To understand the mechanistic basis for S54A selection, we analyzed the four FDA-approved

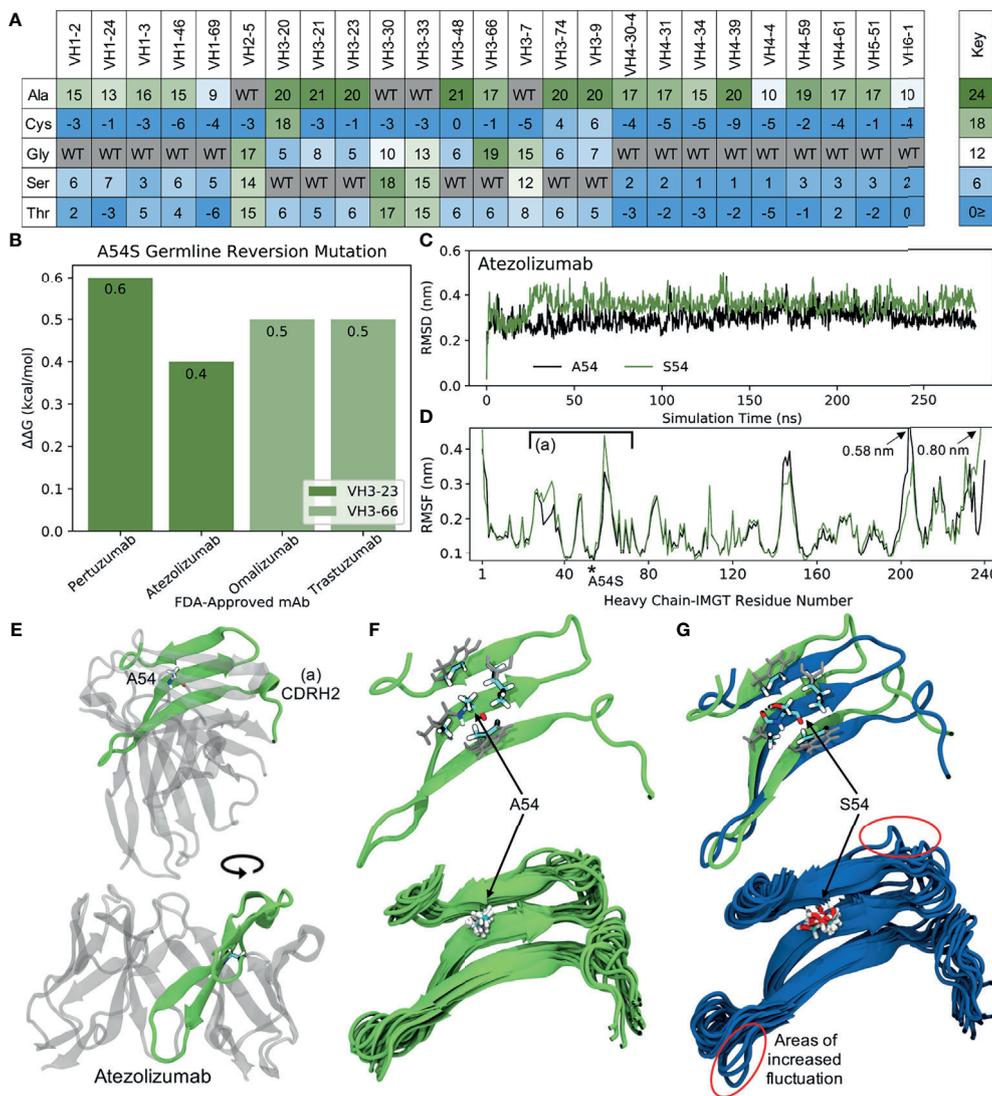


FIGURE 7 | Molecular dynamics (MD) simulations indicate the ‘universal’ FR mutation S54A improves mAb stability. **(A)** Heatmap of FR scores for the mutation of residue 54 to different amino acids for each of the 25 germline genes. **(B)** Change in stability as predicted by PremPS upon introducing the A54S germline-reverted mutation into the sequences of four FDA-approved mAbs from two different germlines. **(C)** Root mean square deviation (RMSD) of Atezolizumab with and without the germline-reverted mutation A54S, referenced to the respective energy-minimized structures, as a function of MD simulation time. **(D)** Root mean square fluctuation (RMSF) of individual residues in Atezolizumab with and without the A54S mutation, calculated from the MD trajectories. **(E)** Energy-minimized structure of Atezolizumab in white, with regions in green indicating residues that experienced relatively large fluctuations in the MD simulations, as indicated in **(D)** by square brackets. **(F)** Zoomed-in view of region-of-interest **(A)** from panels **(D, E)**, showing the favorable hydrophobic binding pocket of A54 in Atezolizumab (top) and the stability of surrounding loop regions (bottom; 15 overlaid snapshots, equally spaced across the entire trajectory). **(G)** MD simulation snapshot of Atezolizumab with (blue) and without (green) the germline-reverted mutation A54S, highlighting differences in local secondary structure (top) and relatively large fluctuations in surrounding loops (bottom). A black dashed line indicates a hydrogen bond.

mAbs listed above with this S54A mutation. To test for increased global stability of the mAbs, we utilized the online protein stability prediction algorithm PremPS (52) to determine the impact of reverting the alanine at position 54 in each of the four mAbs back to its identity of a serine in the respective germline sequences (Figure 7B). The A54S reversion mutation was predicted by PremPS to decrease mAb stability (calculated as $\Delta\Delta G$) for all four mAbs, indicating the forward mutation (S54A) in the mature FDA-approved sequences was predicted to be stabilizing. In addition, we see on average a 1.9-fold reduction in peptide: MHC-II affinity for peptides from FDA-approved mAbs with this mutation. Thus, *in silico* methods predict that S54A reduces the potential for CD4+ T cell immunogenicity and increases global stability in these four mAbs.

To gain a more mechanistic picture of how the S54A mutation mediates stability in the mature FDA-approved mAbs, we performed molecular dynamics (MD) simulations of Atezolizumab both in its mature form (e.g., with A54) and with the germline-reverted mutation (e.g., with S54). Production simulations were performed for 0.28 s in the NPT ensemble, using the GROMACS-2021 (30, 31) simulation engine and the AMBER99SB-ILDN (33) and TIP3P (32) force fields to describe the mAbs and solvent, respectively (see *Methods*). Figure 7C shows the root mean square fluctuation (RMSD) of the mature and mutant forms of Atezolizumab from the respective energy-minimized crystal structures as a function of simulation time. The results show the RMSD of the mature mAb was consistently lower than that of the mutant mAb, indicating increased stability of the mature mAb due to the S54A mutation, in line with both our mutation scores and PremPS results.

Increased stability due to the S54A mutation is further supported by the higher root mean square fluctuation (RMSF) of residues in the mutant *versus* mature mAb during the simulations (Figure 7D). In particular, we observe a distinct region in the V_H where residues in the mutant mAb experienced increased flexibility (decreased stability) during the simulations, namely IMGT-aligned V_H residues 26-72. This region corresponds to a β -sheet on which residue 54 is situated (Figure 7E), indicating the observed changes in RMSF are directly attributable to this mutation. In the mature mAb simulation, we observe A54 to be optimally situated within a hydrophobic pocket formed by the inward-facing side chains of surrounding residues (Figure 7F, top). Favorable binding interactions between the methyl group on A54 with these hydrophobic side chains during the simulation results in relatively small fluctuations of nearby loops (Figure 7F, bottom), which includes the CDRH2 loop that is often in direct contact with antigens. Conversely, substituting a serine into this hydrophobic pocket leads to S54 having an unsaturated hydrogen bond. The polar side chain of S54 is observed to rotate freely during the simulation, forming a hydrogen bond only transiently with a nearby residue (Figure 7G, top) and causing increased fluctuations in the nearby loops (Figure 7G, bottom). The decreased motion of the CDRH2 region due to the S54A mutation is consistent with previous studies showing Ab rigidification may selectively occur during affinity maturation for certain GL families upon binding antigen (53). Future computational studies could test this directly

through MD simulations of FDA-approved mAbs in complex with their cognate antigens. Figure S11 demonstrates this mechanism of stability due to the S54A FR mutation is the same for V_H3-66 -based mAbs as well, underscoring the ‘universal’ nature of this mutation.

We observed a second mutation (Y103F) with even higher prevalence among human repertoire antibodies. Figure 8A shows the predicted scores for the mutation across residues with similar hydrophobic side chains and/or aromatic structures (F, H, M, W and Y) for all of the 25 GL V_H genes studied. Our results show that this mutation broadly occurs across many of the GL V_H genes, with high scores ranging between 18 and 22. In addition, it should be noted that 103F is present in several mouse germlines (V_H1 , V_H9 , V_H11 , V_H13) (26), providing evidence that this mutation should be structurally tolerated particularly for mouse-derived antibodies. Our collected data gives a frequency of 12% for the Y103F mutation, which is consistent with the ~15% frequency found in a prior antibody evolution study (19).

To determine the impact of the Y103F mutation on mAb stability, we analyzed the single FDA-approved mAb (Daratumumab) containing this mutation, as well as three other mAbs identified through an IMGT BlastSearch to have this mutation. The three non-FDA-approved mAbs included the anti-influenza FluA-20 Fab from the V_H4-61 GL family, the anti-HIV 17b Fab from the V_H1-69 GL family, and the anti-malarial 092096 Fab from the V_H1-24 GL. As with the S54A mutation, we first utilized PremPS to assess the change in mAb stability upon reversion of the phenylalanine at position 103 in the mature sequences of these mAbs to its identity of a tyrosine in all of the respective germlines. Figure 8B shows that this mutation, in contrast with the S54A mutation, is predicted to be destabilizing to all four of the analyzed mAbs. This prompted us to explore this mutation further: why would an antibody repeatably select a destabilizing mutation during affinity maturation, particularly one far away from the binding site?

To address this question, we performed MD simulations of Daratumumab both with and without the Y103F mutation. Production simulations were carried out as described earlier for the simulations with Atezolizumab and Omalizumab (see also *Methods*). Over the converged, second half of the simulation, we observe the RMSD of the mature antibody to be consistently higher than that of the mutant mAb (Figure 8C), suggesting this mutation is indeed destabilizing. This result aligns with our PremPS results but still contradicts our mutation scores.

Decreased mAb stability due to the Y103F mutation is also evidenced by the decreased RMSF values of residues in the mutant *versus* mature mAb, averaged over the converged portion of the simulations (Figure 8D). The largest RMSF differences are observed for IMGT-aligned V_H residues 19-80. This region corresponds to FR2, CDR2 and FR3 and directly contacts the Y103F mutation, which is situated between the V_H and V_L domains of the antibody Fab (Figure 8E). In the simulation of the mutant mAb, the tyrosine is observed to form consistent hydrogen bonds with neighboring residues, whereas the phenylalanine in the simulation with the mature mAb is unable to do so (Figures 8F, G, top). Consistent hydrogen bonds formed by Y103 essentially lock the tyrosine

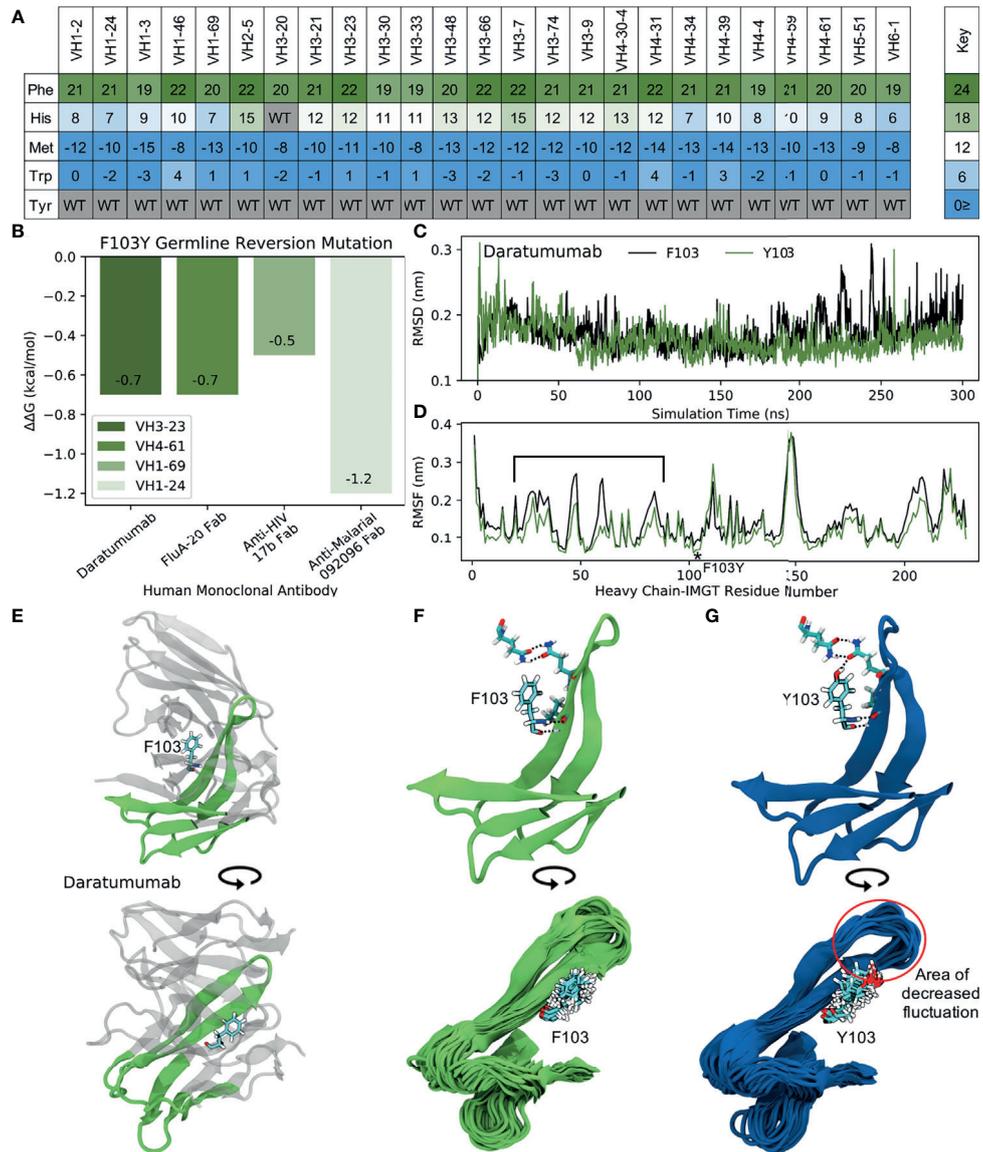


FIGURE 8 | MD simulations indicate the ‘universal’ FR mutation Y103F decreases mAb stability. **(A)** Heatmap of FR scores for the mutation of residue 103 to different amino acids for each of the 25 germline genes. **(B)** Change in stability as predicted by PremPS upon introducing the F103Y germline-reverted mutation into the sequence of Daratumumab and three other human mAbs found to have the F103Y mutation. **(C)** Root mean square deviation (RMSD) of Daratumumab with and without the germline-reverted mutation F103Y as a function of MD simulation time. **(D)** Root mean square fluctuation (RMSF) of individual residues in Daratumumab with and without the F103Y mutation. **(E)** Energy-minimized structure of Daratumumab in white with highly fluctuating regions in green, as indicated in **(D)** by square brackets. **(F)** Zoomed-in view of region-of-interest from panels **(D, E)**, highlighting the reduced hydrogen bonding capabilities of F103 in Daratumumab (top), leading to increased fluctuations in surrounding loop regions (bottom). **(G)** MD simulation snapshot of Daratumumab with the F103Y mutation, highlighting the strong hydrogen bonding capacity of Y103 (top), leading to reduced fluctuations in surrounding loops (bottom).

into place, leading to relatively small fluctuations in surrounding loops during the simulation (**Figure 8G**, bottom). Conversely, large fluctuations in neighboring loops are observed in the simulation with the mature mAb (**Figure 8F**, bottom).

Overall, our simulations support the PremPS result that Y103F destabilizes the structure of Daratumumab, which indicates that the effects of this particular mutation may be much more context specific. However, due to the mutation’s

unique location at the intersection between the V_H and V_L domains of the antibody, we hypothesize that Y103F may facilitate a favorable conformational transition of the antibody upon antigen binding. This could explain the high prevalence with which this mutation is observed in baseline human antibody repertoires, but would require data on the light chain sequences of the antibodies to make more firm conclusions. Regardless, the negative impact of the Y103F mutation on the stability of

Daratumumab suggests it may be important to consider a wide variety of mutations - whether frequently observed or not - when analyzing or engineering new therapeutic antibodies.

DISCUSSION

In this work, we generated PSSMs using repertoire sequences from 25 GL genes. We used these scores to assess framework mutations of FDA-approved therapeutic antibodies and to determine the degree of similarity with human antibody sequences. We found that mutations in FDA-approved antibodies are also common in human antibody repertoires.

To understand why high frequency repertoire mutations are observed, we analyzed various high-scoring mutations for stability effects, as we expect thermostability to be a top priority to select for in both baseline human repertoire antibodies as well as antibody therapeutics. Due to scant evidence in literature, we were unable to make firm conclusions about this claim. A comprehensive study of stability would be necessary to resolve this issue by directly testing many antibody variants *in vitro*. The potential for CD4+ T cell immunogenicity is another critical factor that should be selected for in baseline repertoire and human-engineered antibodies. Our work shows that increasing PSSM scores correlate with decreasing peptide:MHC-II binding affinity, signifying removal of CD4+ T cell (MHC-II) epitopes and potentially decreasing immunogenicity, a critical quality in antibody therapeutics. In future therapeutic mAb development, we speculate that repertoire-based PSSMs could be utilized as early warning signals that a sequence may have developability flags or induce an immune response.

The high scoring FR mutation S54A was evaluated *in silico* and by simulations, where we found clear evidence supporting the claim that S54A confers benefits to both stability and reduced MHC-II peptide epitope content. While we anticipate the factors investigated (thermostability and MHC-II peptide epitope content) to be highly correlated with relative repertoire abundance, there are clearly many other factors to discover. This is made clear by the fact that some mutations that we would expect to be detrimental to stability, like Y103F, seem to have high scores. This must be because they were selected for other benefits not considered in this study. We expect that future research surrounding human antibody repertoires will continue to illuminate the poorly understood selection criteria that are implemented in affinity maturation with regards to framework mutations.

Our work highlights several important findings of baseline human antibody repertoires and the dynamics of affinity maturation that produce them. We found that framework mutations are remarkably consistent across individuals, suggesting that affinity maturation is selecting mutations in a consistent manner. This finding indicates that it may be well-suited for computational modeling, allowing for accurate prediction of the types and frequencies of framework mutations produced in affinity maturation. In addition, our results independently confirm the notion that mutation profiles are GL-dependent. We showed that the correlations between mutation scores of germline family members are much higher than those

across families. This analysis denotes that mutational frequencies are highly dependent on sequence, likely due to AID's intrinsic mutational biases. We look forward to future research aimed at determining the degree to which baseline human antibody repertoire mutations are dictated by the precursor local sequence environment.

Previous work on humanization of mouse Abs showed the importance of FR Vernier positions that scaffold and impact affinity of specific CDRs. We do find that such Vernier positions close to CDR1 and CDR2 have increased mutational frequency relative to other FR positions (**Figure S2**); future work could identify the covariation of specific FR and CDR mutations, and we expect that similar refinements may identify FR mutations that are context-specific for individual CDRs or, in the case of Y103F, different light chains.

With the current high costs of and increasing demand for mAb therapeutics, methods that can infer sequence to function relationships are very valuable. We speculate that incorporating high scoring mutations into mAb sequences will further enhance drug-like properties for existing therapies and could significantly reduce development costs for future mAbs.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

Designed experiments: BP, SU, ER, KS, MG, BD, and TW. Performed experiments: BP, SU, and MG. Performed simulations: ER and KS. Wrote paper with contributions from all co-authors: BP, SU, ER, KS and TW. All authors contributed to the article and approved the submitted version.

FUNDING

Research reported in this publication was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under Award Number R01AI141452 to TW and BD.

ACKNOWLEDGMENTS

The authors wish to acknowledge members of the Whitehead lab for constructive comments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2021.728694/full#supplementary-material>

REFERENCES

- The Business Research Company. *Monoclonal Antibodies (MAbS) Global Market Report 2021: COVID 19 Impact And Recovery To 2030* (2021). Available at: [https://www.thebusinessresearchcompany.com/report/monoclonal-antibodies-mabs-global-market-report#:~:text=The%20global%20monoclonal%20antibodies%20\(MABs\),\(CAGR\)%20of%207.1%25](https://www.thebusinessresearchcompany.com/report/monoclonal-antibodies-mabs-global-market-report#:~:text=The%20global%20monoclonal%20antibodies%20(MABs),(CAGR)%20of%207.1%25) (Accessed May 27, 2021).
- Wolchok JD, Kluger H, Callahan MK, Postow MA, Rizvi NA, Lesokhin AM, et al. Nivolumab Plus Ipilimumab in Advanced Melanoma. *N Engl J Med* (2013) 369:122–33. doi: 10.1056/nejmoa1302369
- Mease PJ, Ory P, Sharp JT, Ritchlin CT, Van Den Bosch F, Wellborne F, et al. Adalimumab for Long-Term Treatment of Psoriatic Arthritis: 2-Year Data From the Adalimumab Effectiveness in Psoriatic Arthritis Trial (ADEPT). *Ann Rheum Dis* (2009) 68:702–9. doi: 10.1136/ard.2008.092767
- Baum A, Fulton BO, Wloga E, Copin R, Pascal KE, Russo V, et al. Antibody Cocktail to SARS-CoV-2 Spike Protein Prevents Rapid Mutational Escape Seen With Individual Antibodies. *Science* (80-) (2020) 369:1014–8. doi: 10.1126/science.abd0831
- Jain T, Sun T, Durand S, Hall A, Houston NR, Nett JH, et al. Biophysical Properties of the Clinical-Stage Antibody Landscape. *Proc Natl Acad Sci USA* (2017) 114:944–9. doi: 10.1073/pnas.1616408114
- Persson H, Kirik U, Thörnqvist L, Greiff L, Levander F, Ohlin M. *In Vitro* Evolution of Antibodies Inspired by *In Vivo* Evolution. *Front Immunol* (2018) 9:1391. doi: 10.3389/fimmu.2018.01391
- Faber MS, Whitehead TA. Data-Driven Engineering of Protein Therapeutics. *Curr Opin Biotechnol* (2019) 60:104–10. doi: 10.1016/j.copbio.2019.01.015
- Lehmann M, Loch C, Middendorf A, Studer D, Lassen SF, Pasamontes L, et al. The Consensus Concept for Thermostability Engineering of Proteins: Further Proof of Concept. *Protein Eng* (2002) 15:403–11. doi: 10.1093/protein/15.5.403
- Amin N, Liu AD, Ramer S, Aehle W, Meijer D, Metin M, et al. Construction of Stabilized Proteins by Combinatorial Consensus Mutagenesis. *Protein Eng Des Sel* (2004) 17:787–93. doi: 10.1093/protein/gzh091
- Aerts D, Verhaeghe T, Joosten HJ, Vriend G, Soetaert W, Desmet T. Consensus Engineering of Sucrose Phosphorylase: The Outcome Reflects the Sequence Input. *Biotechnol Bioeng* (2013) 110:2563–72. doi: 10.1002/bit.24940
- Kunz P, Flock T, Soler N, Zaiss M, Vincke C, Sterckx Y, et al. Exploiting Sequence and Stability Information for Directing Nanobody Stability Engineering. *Biochim Biophys Acta - Gen Subj* (2017) 1861:2196–205. doi: 10.1016/j.bbagen.2017.06.014
- Yao H, Cai H, Li D. Thermostabilization of Membrane Proteins by Consensus Mutation: A Case Study for a Fungal $\Delta 8-7$ Sterol Isomerase. *J Mol Biol* (2020) 432:5162–83. doi: 10.1016/j.jmb.2020.02.015
- Apostoei AJ, Trabalka JR. *Review, Synthesis, and Application of Information on the Human Lymphatic System to Radiation Dosimetry for Chronic Lymphocytic Leukemia*. Oak Ridge, Tn. SENES Oak Ridge, Inc. Center for Risk Analysis (2010).
- Steipe B, Schiller B, Pluückthun A, Steinbacher S. Sequence Statistics Reliably Predict Stabilizing Mutations in a Protein Domain. *J Mol Biol* (1994) 240:188–92. doi: 10.1006/jmbi.1994.1434
- Kabat EA, Wu TT, Perry HM, Gottesman KS, Foeller C. *Sequences of Proteins of Immunological Interest*. 5th ed. Bethesda, MD: U.S. Dept. of Health and Human Services (1991).
- Sheng Z, Schramm CA, Kong R, Mullikin JC, Mascola JR, Kwong PD, et al. Gene-Specific Substitution Profiles Describe the Types and Frequencies of Amino Acid Changes During Antibody Somatic Hypermutation. *Front Immunol* (2017) 8:537. doi: 10.3389/fimmu.2017.00537
- Bonsignori M, Zhou T, Sheng Z, Chen L, Gao F, Joyce MG, et al. Maturation Pathway From Germline to Broad HIV-1 Neutralizer of a CD4-Mimic Antibody. *Cell* (2016) 165:449–63. doi: 10.1016/j.cell.2016.02.022
- Ellebedy AH, Jackson KJL, Kissick HT, Nakaya HI, Davis CW, Roskin KM, et al. Defining Antigen-Specific Plasmablast and Memory B Cell Subsets in Human Blood After Viral Infection or Vaccination. *Nat Immunol* (2016) 17:1226–34. doi: 10.1038/ni.3533
- Kirik U, Persson H, Levander F, Greiff L, Ohlin M. Antibody Heavy Chain Variable Domains of Different Germline Gene Origins Diversify Through Different Paths. *Front Immunol* (2017) 8:1433. doi: 10.3389/fimmu.2017.01433
- Briney B, Inderbitzin A, Joyce C, Burton DR. Commonality Despite Exceptional Diversity in the Baseline Human Antibody Repertoire. *Nature* (2019) 566:393–7. doi: 10.1038/s41586-019-0879-y
- Gutiérrez-González M, Fahad AS, Ardito M, Nanaware P, Lu L, Normandin E, et al. Human Antibody Immune Responses Are Personalized by Selective Removal of MHC-II Peptide Epitopes (2021). Available at: <https://www.biorxiv.org/content/10.1101/2021.01.15.426750v1> (Accessed April 27, 2021).
- Vaisman-Mentesh A, Gutierrez-Gonzalez M, DeKosky BJ, Wine Y. The Molecular Mechanisms That Underlie the Immune Biology of Anti-Drug Antibody Formation Following Treatment With Monoclonal Antibodies. *Front Immunol* (2020) 11:1951. doi: 10.3389/fimmu.2020.01951
- Cassotta A, Mikol V, Bertrand T, Pouzieux S, Le Parc J, Ferrari P, et al. A Single T Cell Epitope Drives the Neutralizing Anti-Drug Antibody Response to Natalizumab in Multiple Sclerosis Patients. *Nat Med* (2019) 25:1402–7. doi: 10.1038/s41591-019-0568-2
- Owen JA, Punt J, Stranford SA, Jones PP, Kuby J. *Adaptive Immunity: Effector Responses*. Kuby Immunology. New York, NY: W.H. Freeman and Company. (2013) p. 385–414.
- Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, et al. DrugBank 5.0: A Major Update to the DrugBank Database for 2018. *Nucleic Acids Res* (2018) 46:D1074–82. doi: 10.1093/nar/gkx1037
- Lefranc MP. IMGT, the International ImmunoGeneTics Database®. *Nucleic Acids Res* (2003) 31:307–10. doi: 10.1093/nar/gkg085
- Katoh K, Rozewicki J, Yamada KD. MAFFT Online Service: Multiple Sequence Alignment, Interactive Sequence Choice and Visualization. *Brief Bioinform* (2018) 20:1160–6. doi: 10.1093/bib/bbx108
- Dayhoff MO, Barker WC, Hunt LT. Establishing Homologies in Protein Sequences. *Methods Enzymol* (1983) 91:524–45. doi: 10.1016/S0076-6879(83)91049-2
- Reynisson B, Barra C, Kaabinejad S, Hildebrand WH, Peters B, Peters B, et al. Improved Prediction of MHC II Antigen Presentation Through Integration and Motif Deconvolution of Mass Spectrometry MHC Eluted Ligand Data. *J Proteome Res* (2020) 19:2304–15. doi: 10.1021/acs.jproteome.9b00874
- Lindahl, Abraham, Hess, Spoel van der. *GROMACS 2021 Source Code*. Zenodo (2021). doi: 10.5281/ZENODO.4457626.
- Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. Gromacs: High Performance Molecular Simulations Through Multi-Level Parallelism From Laptops to Supercomputers. *SoftwareX* (2015) 1–2:19–25. doi: 10.1016/j.softx.2015.06.001
- Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of Simple Potential Functions for Simulating Liquid Water. *J Chem Phys* (1983) 79:926–35. doi: 10.1063/1.445869
- Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, et al. Improved Side-Chain Torsion Potentials for the Amber Ff99sb Protein Force Field. *Proteins Struct Funct Bioinforma* (2010) 78:1950–8. doi: 10.1002/prot.22711
- Lee HT, Lee JY, Lim H, Lee SH, Moon YJ, Pyo HJ, et al. Molecular Mechanism of PD-1/PD-L1 Blockade via Anti-PD-L1 Antibodies Atezolizumab and Durvalumab. *Sci Rep* (2017) 7:5532. doi: 10.1038/s41598-017-06002-8
- Madej T, Lanczycki CJ, Zhang D, Thiessen PA, Geer RC, Marchler-Bauer A, et al. MMDb and VAST+: Tracking Structural Similarities Between Macromolecular Complexes. *Nucleic Acids Res* (2014) 42:D297–303. doi: 10.1093/nar/gkt1208
- Jensen RK, Plum M, Tjerrild L, Jakob T, Spillner E, Andersen GR. Structure of the Omalizumab Fab. *Acta Crystallogr Sect F* (2015) 71:419–26. doi: 10.1107/S2053230X15004100
- Humphrey W, Dalke A, Schulten K. VMD: Visual Molecular Dynamics. *J Mol Graph* (1996) 14:33–8. doi: 10.1016/0263-7855(96)00018-5
- Bussi G, Donadio D, Parrinello M. Canonical Sampling Through Velocity Rescaling. *J Chem Phys* (2007) 126:14101. doi: 10.1063/1.2408420
- Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular Dynamics With Coupling to an External Bath. *J Chem Phys* (1984) 81:3684–90. doi: 10.1063/1.448118
- Parrinello M, Rahman A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J Appl Phys* (1981) 52:7182–90. doi: 10.1063/1.328693
- Darden T, York D, Pedersen L. Particle Mesh Ewald: An N-log(N) Method for Ewald Sums in Large Systems. *J Chem Phys* (1993) 98:10089–92. doi: 10.1063/1.464397

42. Pilzecker B, Jacobs H. Mutating for Good: DNA Damage Responses During Somatic Hypermutation. *Front Immunol* (2019) 10:438. doi: 10.3389/fimmu.2019.00438
43. Victora GD, Nussenzweig MC. Germinal Centers. *Annu Rev Immunol* (2012) 30:429–57. doi: 10.1146/annurev-immunol-020711-075032
44. Julian MC, Li L, Garde S, Wilen R, Tessier PM. Efficient Affinity Maturation of Antibody Variable Domains Requires Co-Selection of Compensatory Mutations to Maintain Thermodynamic Stability. *Sci Rep* (2017) 7:1–14. doi: 10.1038/srep45259
45. Tiller KE, Tessier PM. Advances in Antibody Design. *Annu Rev BioMed Eng* (2015) 17:191–216. doi: 10.1146/annurev-bioeng-071114-040733
46. Franklin MC, Carey KD, Vajdos FF, Leahy DJ, de Vos AM, Sliwkowski MX. Insights Into ErbB Signaling From the Structure of the ErbB2-Pertuzumab Complex. *Cancer Cell* (2004) 5:317–28. doi: 10.1016/S1535-6108(04)00083-2
47. Schramm CA, Douek DC. Beyond Hot Spots: Biases in Antibody Somatic Hypermutation and Implications for Vaccine Design. *Front Immunol* (2018) 9:1876. doi: 10.3389/fimmu.2018.01876
48. Venkataramani S, Ernst R, Derebe MG, Wright R, Kopenhaver J, Jacobs SA, et al. In Pursuit of Stability Enhancement of a Prostate Cancer Targeting Antibody Derived From a Transgenic Animal Platform. *Sci Rep* (2020) 10:1–9. doi: 10.1038/s41598-020-66636-z
49. Wang S, Liu M, Zeng D, Qiu W, Ma P, Yu Y, et al. Increasing Stability of Antibody via Antibody Engineering: Stability Engineering on an Anti-hVEGF. *Proteins Struct Funct Bioinforma* (2014) 82:2620–30. doi: 10.1002/prot.24626
50. Miller BR, Demarest SJ, Lugovskoy A, Huang F, Wu X, Snyder WB, et al. Stability Engineering of Scfvs for the Development of Bispecific and Multivalent Antibodies. *Protein Eng Des Sel* (2010) 23:549–57. doi: 10.1093/protein/gzq028
51. Ewert S, Honegger A, Plückthun A. Structure-Based Improvement of the Biophysical Properties of Immunoglobulin VH Domains With a Generalizable Approach. *Biochemistry* (2003) 42:1517–28. doi: 10.1021/bi026448p
52. Chen Y, Lu H, Zhang N, Zhu Z, Wang S, Li M. PremPS: Predicting the Impact of Missense Mutations on Protein Stability. *PLoS Comput Biol* (2020) 16. doi: 10.1371/journal.pcbi.1008543
53. Ovchinnikov V, Louveau JE, Barton JP, Karplus M, Chakraborty AK. Role of Framework Mutations and Antibody Flexibility in the Evolution of Broadly Neutralizing Antibodies. *eLife* (2018) 7:e33038. doi: 10.7554/eLife.33038.001

Author Disclaimer: The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Petersen, Ulmer, Rhodes, Gutierrez-Gonzalez, Dekosky, Sprenger and Whitehead. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.