Supplementary material for:

Biophysical cartography of the native and human-engineered antibody landscapes quantifies the plasticity of antibody developability. Bashour H, Smorodina E, Pariset M, Zhong J et al.

Corresponding authors:

Habib Bashour: habib.bashour@medisin.uio.no
Victor Greiff: victor.greiff@medisin.uio.no

This document contains:

- Supplementary Tables (1–4)
- Supplementary Figures (1–21)
- Supplementary Methods.
- Supplementary Notes (1-3).

Supplementary Tables

Supplementary Table 1 | **Amino acid categorization based on physicochemical properties as defined by [1].** This classification was used as a reference in the computation of several sequence-based DPs.

Category	Amino acids
Aromatic	F+H+W+Y
Tiny	A+C+G+S+T
Small	A+C+D+G+N+P+S+T+V
Aliphatic	A+I+L+V
Nonpolar	A+C+F+G+I+L+M+P+V+W+Y
Polar	D+E+H+K+N+Q+R+S+T
Basic	H+K+R
Acidic	D+E

Supplementary Table 2 | **Minimum weight dominating set (MWDS) DPs as identified by the ABC-EDA for the full set (170,473) of human IgG antibodies at a Pearson correlation coefficient threshold of 0.6.** Doublet parameters (a pair of DPs, see Methods) are mentioned within the same cell and only one (randomly selected) is carried forward to perform the PC analysis (Figure 5C, Figure 7B, Supplementary Figure 20A, Supplementary Figure 21A) and the predictability analysis in Figure 6. DPs with incomplete values (**bolded**) were excluded from these analyses as well.

Sequence DPs				
AbChain_small_content	AbChain_10_charge"	AbChain_hydrophobicity	AbChain_acidic_content	AbChain_aromatic_content
AbChain_length + AbChain_mw	AbChain_percentextcoef + AbChain_molextcoef	AbChain_cysbridges_molextcoef + AbChain_cysbridges_percenextcoef	AbChain_aliphatic_content + AbChain_aliphindex	AbChain_polar_content + AbChain_nonpolar_content
AbChain_hmom	AbChain_solubility	AbChain_instaindex	**AbChain_min_rank**	**AbChain_immunopeptide_regions_span**
AbChain_num_strong_binders				
Structure DPs				
AbStruc_plane_group_interactions	AbStruc_cation_pi_interactions	AbStruc_metsulphur_pi_interactions	AbStruc_sasa	AbStruc_vdw_clashes

AbStruc_pbonds	AbStruc_carbon_pi_interactions	AbStruc_unfolded_pI	AbStruc_loops	AbStruc_beta_bridges
AbStruc_beta_strands	AbStruc_alpha_helices	AbStruc_beta_bends	AbStruc_carbonyl_interactions	AbStruc_steric_clashes
AbStruc_covbonds	AbStruc_hbonds	AbStruc_ibonds	AbStruc_phi_angle	AbStruc_psi_angle
AbStruc_vdw_interactions	AbStruc_folding_energy	AbStruc_free_cys	AbStruc_cys_bridges	AbStruc_pcharge_hetrgen
AbStruc_ncharge_hetrgen	AbStruc_threeten_helices + AbStruc_beta_turns	AbStruc_aromatic_interactions + AbStruc_plane_plane_interactions		

Supplementary Table 3 | **Antibody pairs from AbDb chosen for the structural variance study described in** Supplementary Figure 11**B,C.**

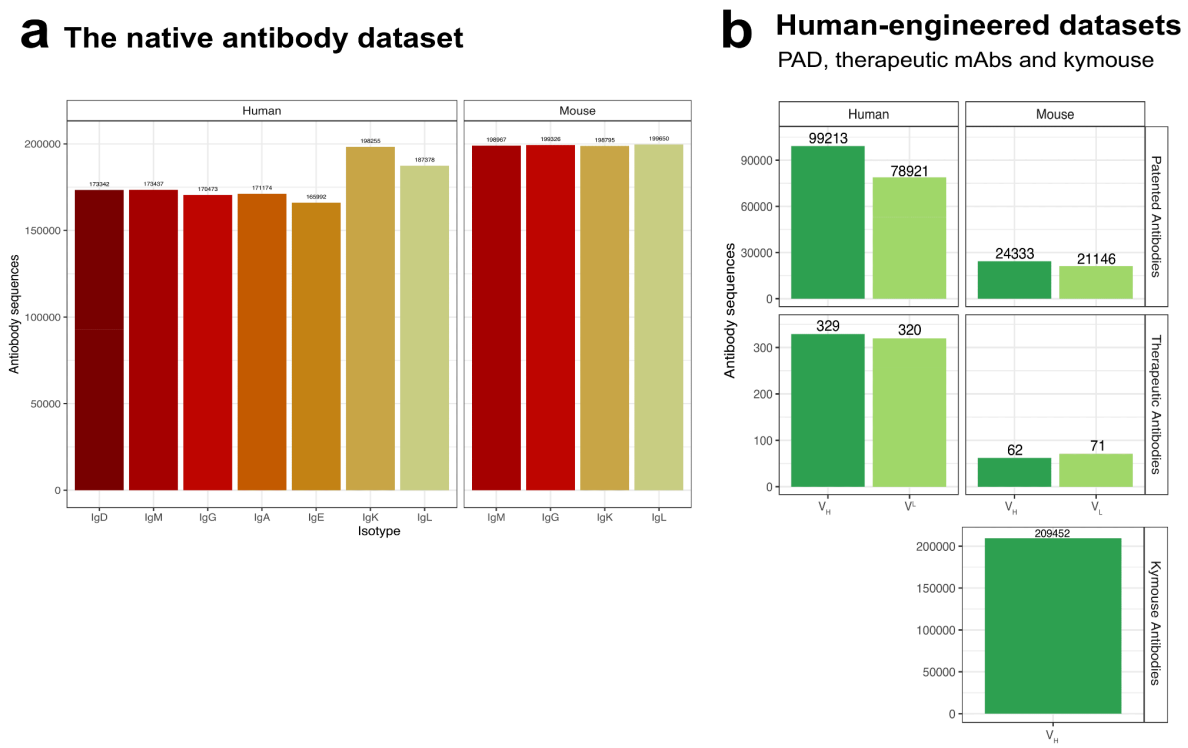
Pair PDB IDs (ID1–ID2)	**Sequence length**	**Position of single aa difference**	**Chain type**
2V7N–5GS1	118	85	V_H
2V7N–5GRV	118	85	V_H
5GRV–5GS1	118	85	V_H
1FVE–6MH2	110	54	V_L
3NAA–3NCJ	110	94	V_L
1DEE–1HEZ	110	93	V_L
5I19–6CNR	110	99	V_L
3X3G–5I1A	110	99	V_L
3HC0–3HC4	110	45	V_L
4HJG–5U5F	110	82	V_L

Supplementary Table 4 **| Crystal structures of mouse and human paired chain antibody Fv regions from AbDb chosen for MD simulations.**

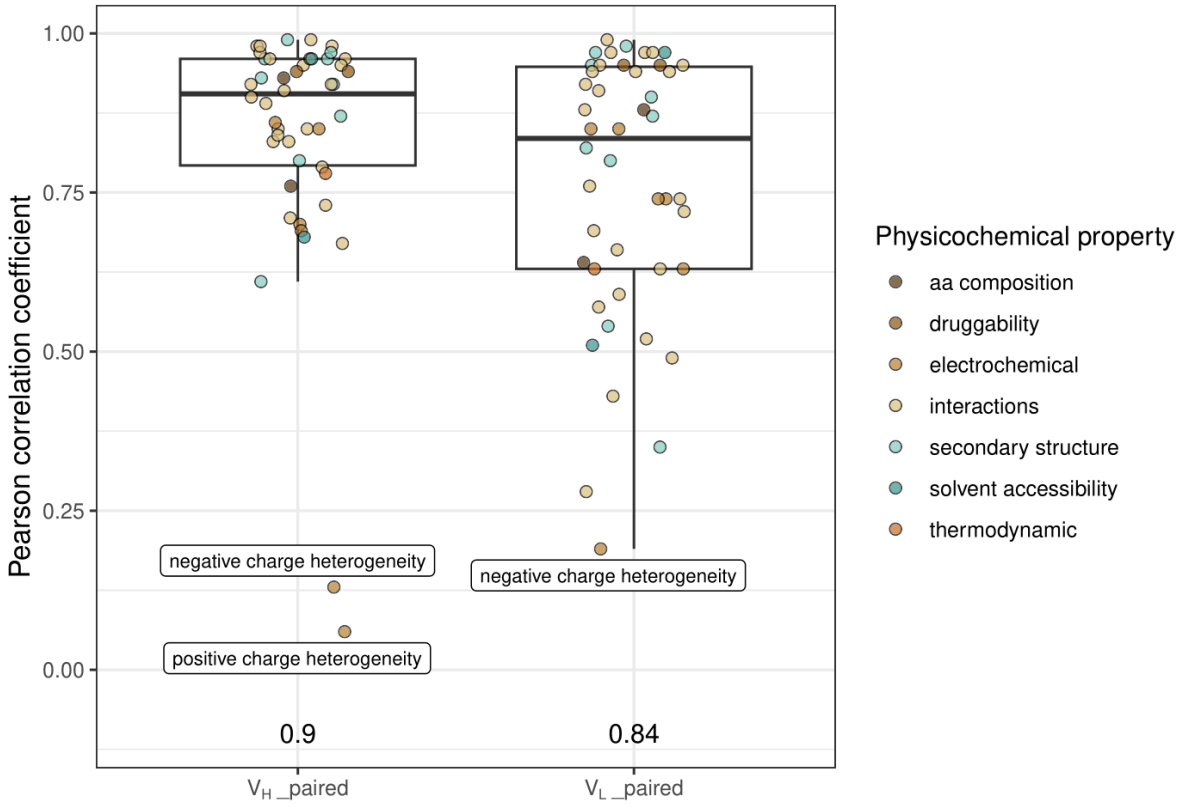
PDB ID	**Isotype**	**Organism**	**Resolution (Å)**	**Number of residues**	**Number of atoms**
1DLF	IgG2κ	Mouse	1.45	233	1763
1MQK	IgG1κ	Mouse	1.28	226	1764

4GXV	IgG1λ	Human	1.45	238	1803
5WCA	IgG1λ	Human	1.37	236	1777
6MEG	IgG1λ	Human	1.41	232	1747

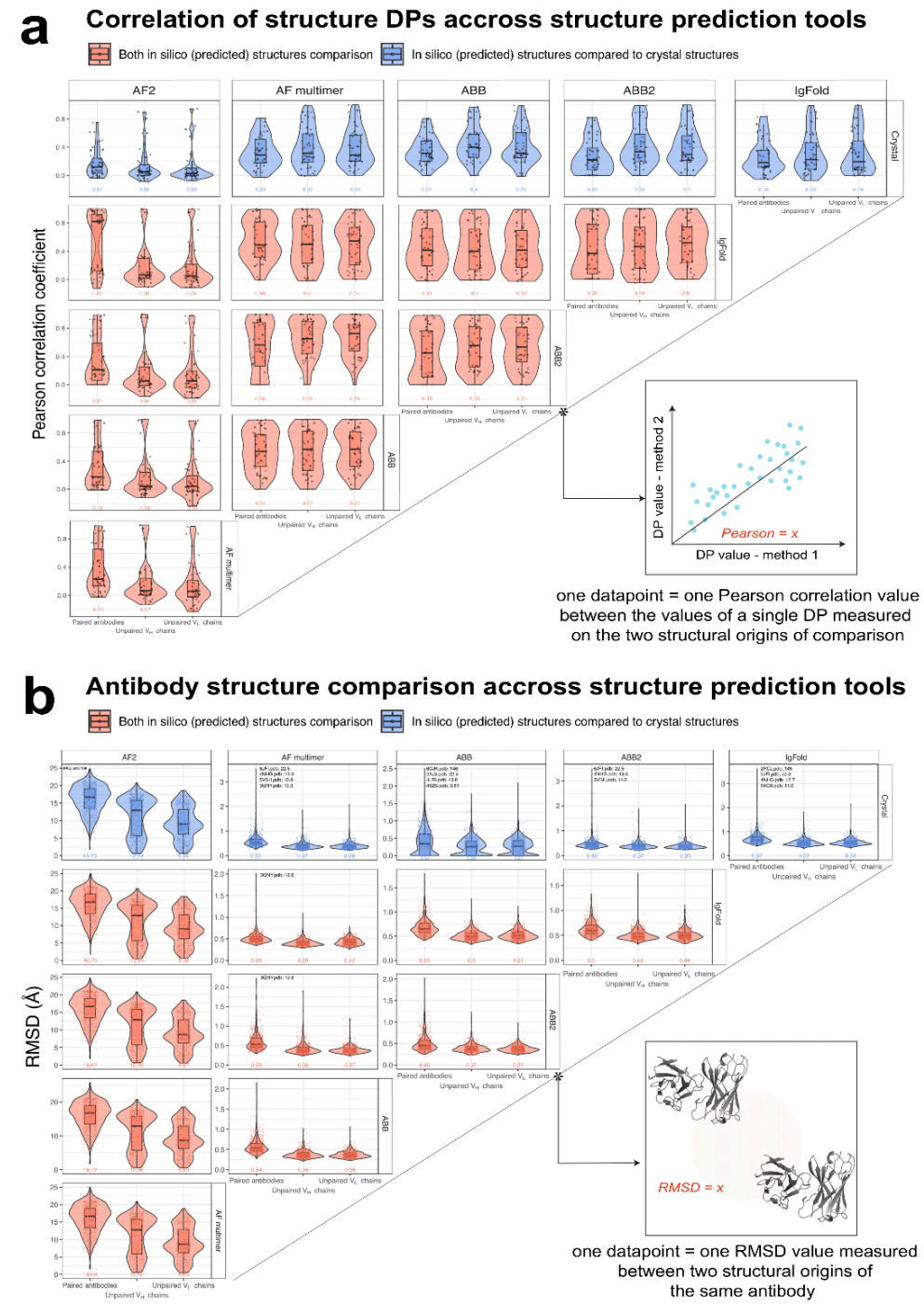
Supplementary Figures



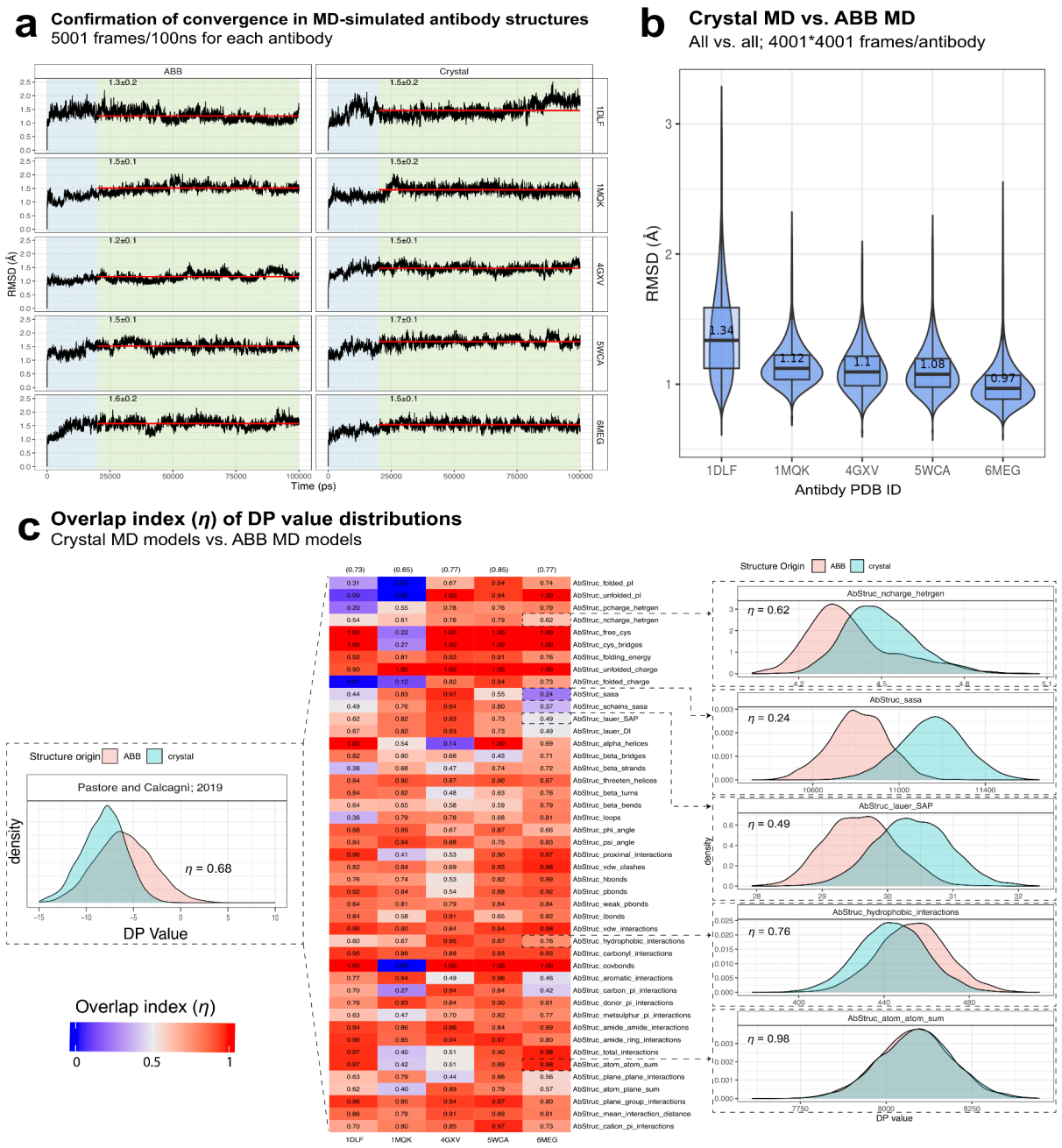
Supplementary Figure 1 | **Overview of the native and human-engineered antibody datasets. (A)** The native antibody datasets. We collected a dataset of 2,036,789 unpaired native antibody sequences (variable regions: Fv) from human and mouse repertoires. The majority of them were sourced from Observed Antibody Space (OAS) database [2–7]. We also included our own experimentally-generated sequences (298,698 antibodies) within the IgD, IgK and IgL human datasets to provide balanced antibody counts among isotypes (see Methods). **(B)** The human-engineered antibody datasets consist of (i) 223,613 patented antibody sequences obtained from the NaturalAntibody company under a non-commercial agreement (credit: Dr. Konrad Krawczyk). Isotype information was only available for the V_L sequences (IgK or IgL), (ii) 782 therapeutic antibody sequences from TheraSAbDAb [8]. V_H sequences were all for the IgG isotype and V_L sequences were from both IgK and IgL isotypes, (iii) 209,452 sequences from humanized mice (Kymouse) [9]. All antibodies in this dataset are V_H sequences of IgM isotype. Relates to Figure 1.



Supplementary Figure 2 | **Pearson correlation of 46 structure DPs measured on single-chain (V_H: "heavy", V_L: "light") compared to corresponding paired-chain structure DP values.** DP values were computed on the crystal structure dataset (859 antibodies – see Methods). The figure legend refers to the classification of DPs by physicochemical property (Supplementary Data 1). Numbers on the x-axis represent the median Pearson correlation coefficient. The positive and negative charge heterogeneity DPs show the lowest correlation across all structures (Pearson correlation <0.2, labeled data points). Relates to Figure 2.


a
Correlation of structure DPs accross structure prediction tools
Both in silico (predicted) structures comparison
In silico (predicted) structures compared to crystal structures
AF2
AF multimer
ABB
ABB2
IgFold
Crystal
Pearson correlation coefficient
DP value - method 2
Pearson = x
DP value - method 1
one datapoint = one Pearson correlation value
between the values of a single DP measured
on the two structural origins of comparison
b
Antibody structure comparison accross structure prediction tools
Both in silico (predicted) structures comparison
In silico (predicted) structures compared to crystal structures
RMSD (Å)
RMSD = x
one datapoint = one RMSD value measured
between two structural origins of
the same antibody
c
Alignment of (in silico) predicted structure
Antibody 1DLF.pdb, Grey: experimental structure, Non-grey: predicted structure
ABB
ABB2
AF2
AF multimer
IgFold
Heavy chain (VH)
Light chain (VL)

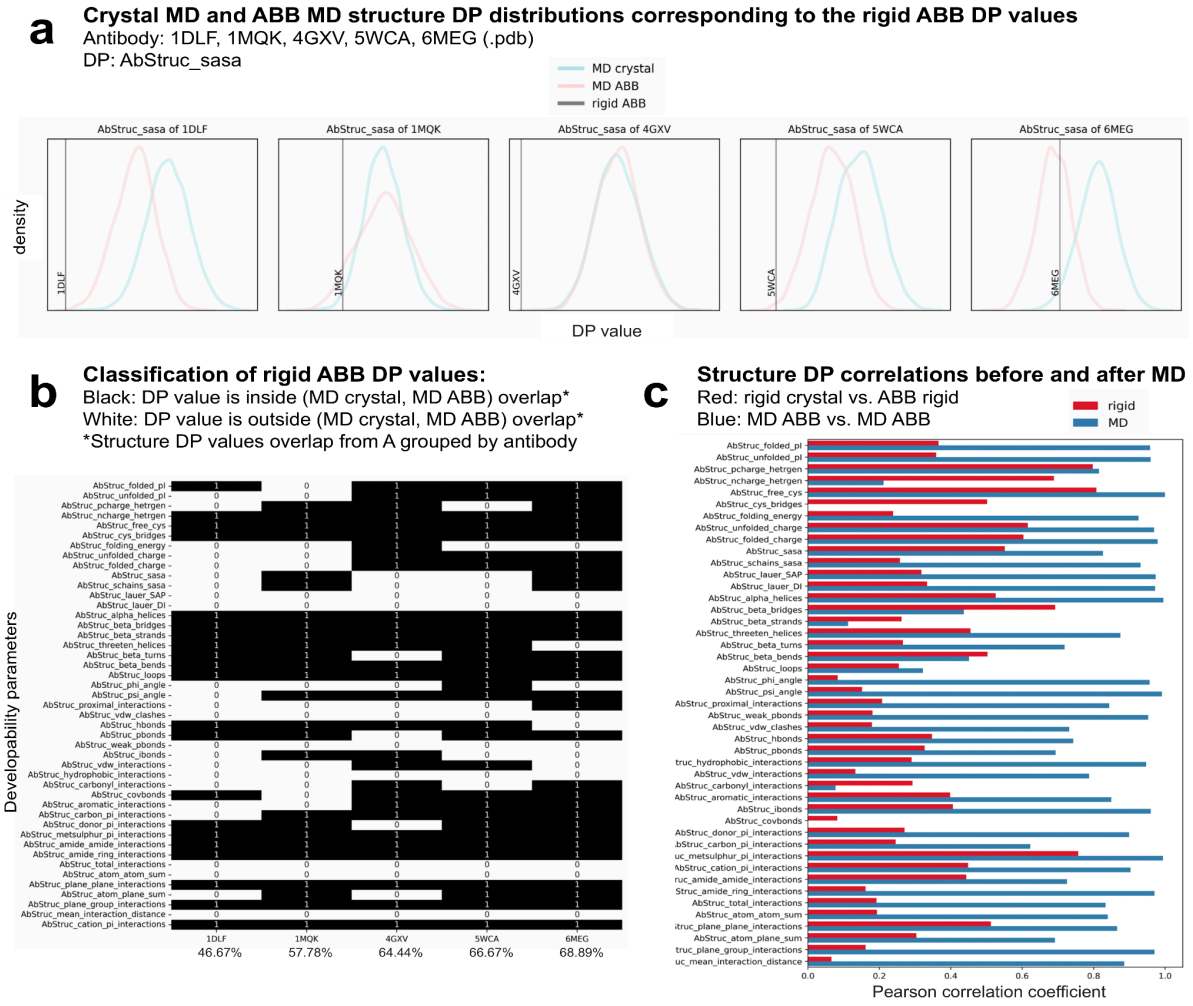
Supplementary Figure 3 | **ABodyBuilder-predicted antibody structures exhibit comparable conformational agreement and closer developability resemblance with reference antibody structures compared to several other state-of-the-art structure prediction tools. (A)** Pairwise Pearson correlation coefficients of structural DPs between antibody structure of experimental or in silico origin. The structures of antibodies (paired and unpaired) in the crystal structures dataset (859 structures – see Methods) were predicted using the in silico tools (AF2; AlphaFold2 [10], AF multimer; AlphaFold-multimer [11], ABB; ABodyBuilder [12], ABB2; ABodyBuilder2 [13], and IgFold [14]). Structure-based DP values were computed on all resulting structures (**46** structure-based DPs, Supplementary Data 1). Each data point in a boxplot refers to one Pearson correlation coefficient between the values of a single structural DP measured on the two structural origins of comparison (see inset graphic on the right) for all 859 antibody structures in the crystal datasets. Numbers on the x-axis represent the median of Pearson correlation values within each chain group (paired, unpaired V_H, unpaired V_L). **(B)** Pairwise root mean square deviation (RMSD) between two antibody structural origins. RMSD values were computed between experimental (crystal) and in silico predicted structures of antibodies from the same dataset (as in A). Each data point in a boxplot represents one RMSD value measured between two structural origins of the same antibody (see inset graphic on the right). Numbers on the x-axis represent the median RMSD within each chain group (paired, unpaired V_H, unpaired V_L). Values higher or equal to ten times the median RMSD within each chain group were considered outliers, and are listed above the corresponding chain group and antibody origin comparison pair (PDB ID; RMSD). **(C)** Example visualization of pairwise alignment between in silico predicted structures and the reference structure of a randomly selected antibody from the crystal structure dataset (PDB ID: 1DLF). Relates to Figure 2.



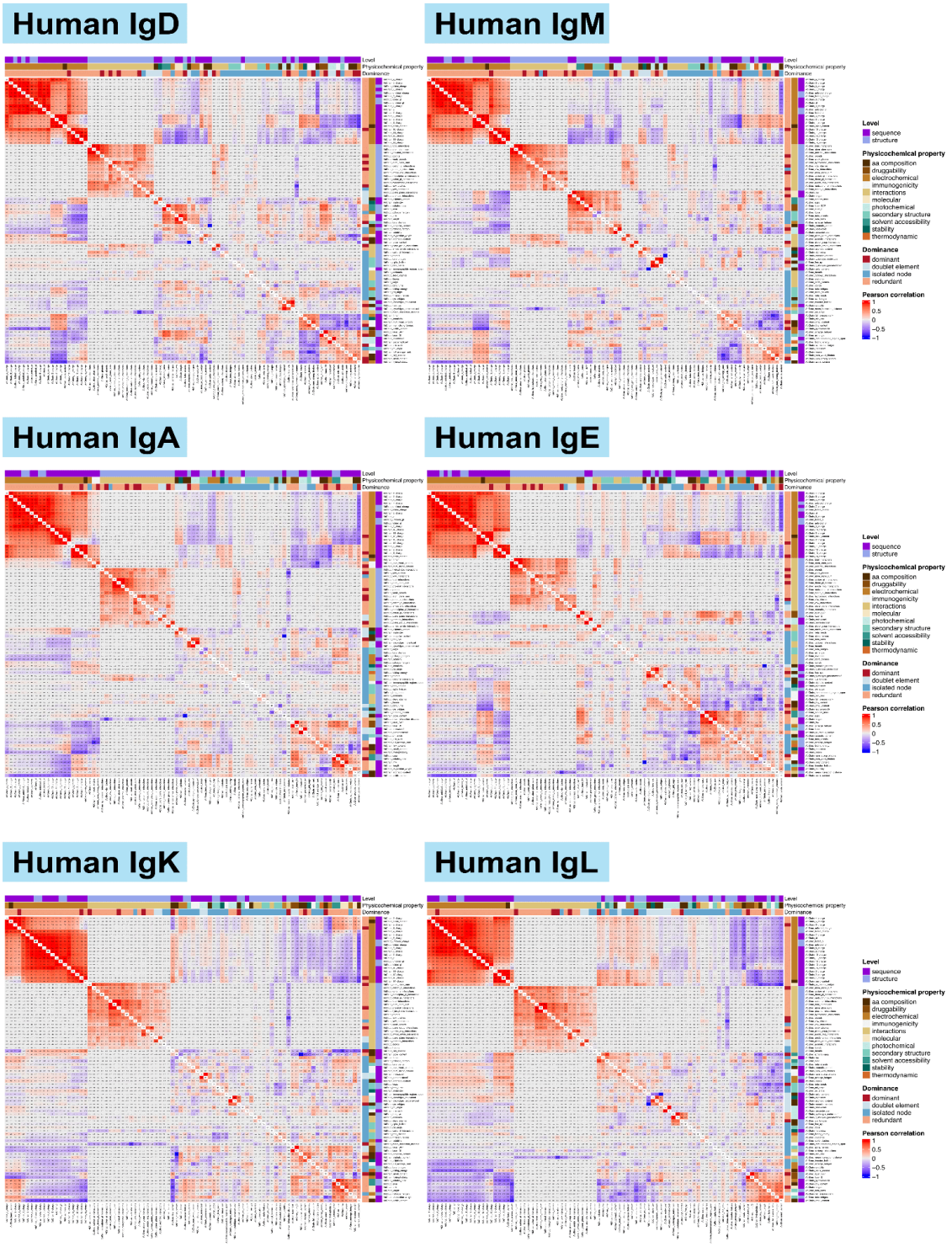
Supplementary Figure 4 | **ABodyBuilder(ABB)-predicted structures show minimal conformational variance and overlapping structural developability in molecular dynamic (MD) simulations in comparison to reference antibody crystal structures.**

Due to the computational cost associated with MD analysis, we selected five antibodies (PDB IDs: 1DLF, 1MQK, 4GXV, 5WCA, 6MEG) from the crystal structure dataset with the highest resolution, and performed the following analyses on their paired-chain structures (details in Methods section and Supplementary Table 4). **(A)** Confirmation of structural convergence. We simulated the ABB-predicted and the crystal structures of the five antibodies in MD for 100,000 ps (100 ns), resulting in a total of 5001 frames per antibody for each structural origin (one frame per 20 ps). The RMSD values for each antibody from the origin structure is shown on the y-axis and the simulation time is shown on the x-axis. The blue-shaded blocks represent pre-convergent frames (0–20,000 ps), and the green-shaded blocks highlight the convergent antibody frames (from 20 ns onwards). The red horizontal lines represent the mean RMSD for convergent frames. The text on the plots represents the mean RMSD ± SD values. Only convergent frames were included in downstream analyses (panels B, C of this figure). **(B)** Pairwise RMSD measurements of MD-simulated antibody structures comparing the conformations of ABB against the conformations of crystal (reference) structures (all vs. all; 4001*4001 datapoints/antibody). Numbers shown on the figure represent the median RMSD (Å). **(C)** Overlap index (η) between the distribution of DP values when measured on the MD-simulated crystal structures and the MD-simulated ABB-predicted structures. AbStruc_steric_clashes are equal to 0 as

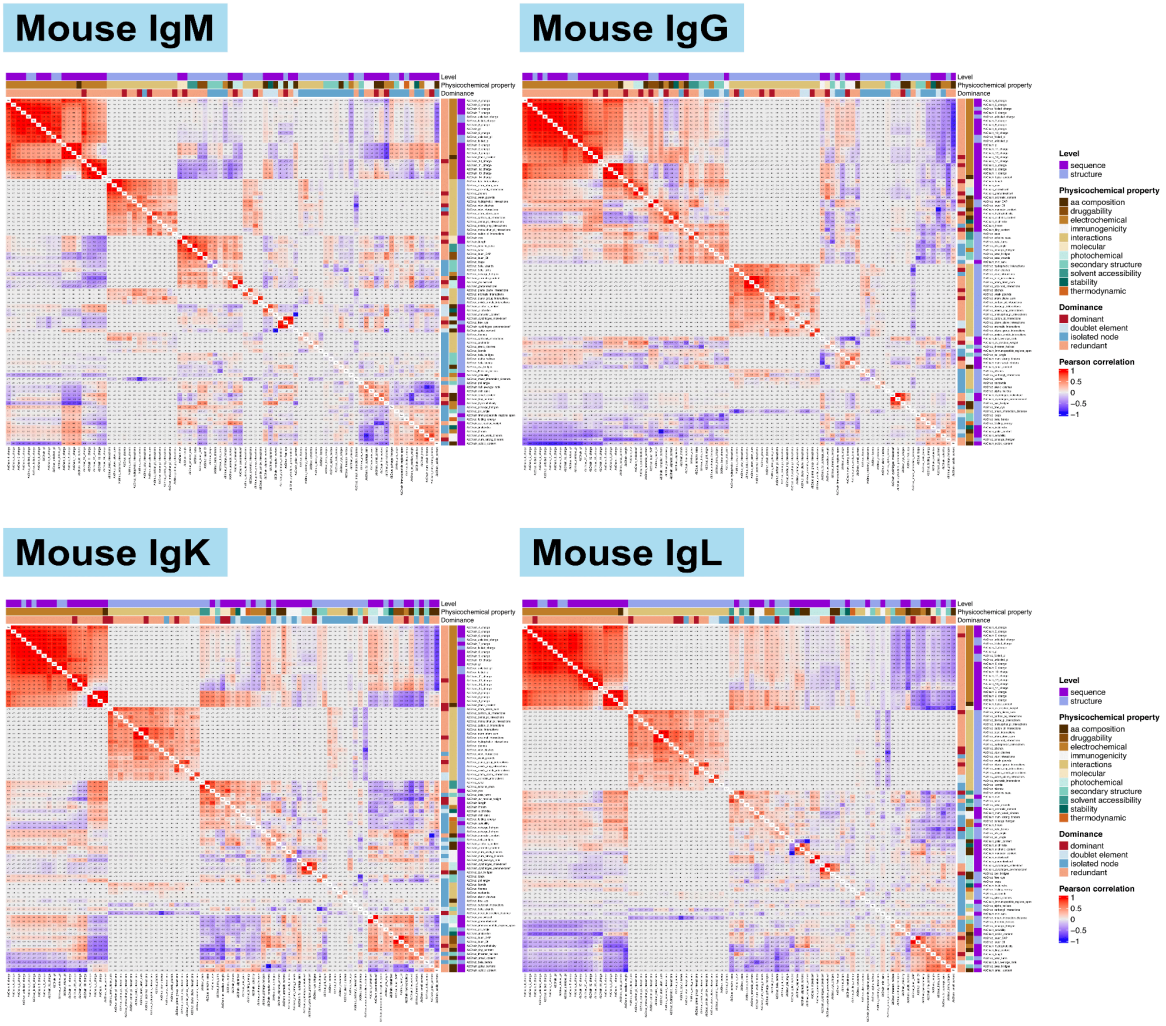
MD relaxed physical overlaps between atoms are not shown in the figure. Left panel: a graphical representation of the overlap index as defined by Pastore and Calcagnì [15]. η is a distribution similarity measure that can take values from 0–1 and it quantifies the overlapping area between two distribution curves. Middle panel: The overlap index measured for each antibody by DP. Each column represents an antibody and each row represents a structural DP. The color intensity in the heatmap and the numerical values within each cell reflect the overlap index. Numbers shown in brackets above each column represent the mean overlap index for each antibody. Right panel: examples of the overlap index and the value distribution of five DPs for the crystal (reference) and ABB-predicted MD-simulated convergent frames (4001 datapoints/origin) of the antibody with PDB ID 6MEG. Relates to Figure 2.



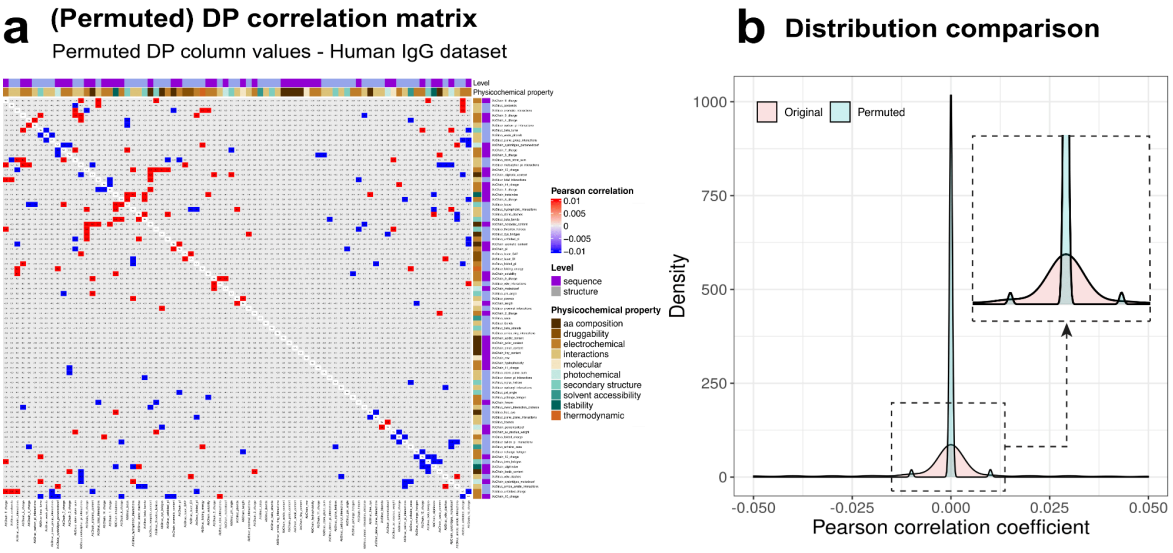
Supplementary Figure 5 | **ABodyBuilder (ABB) models reflect the real-world structures in terms of DP values in 60% of the case (on average), and molecular dynamics (MD) improves poor correlations between structure DPs caused by the used structure prediction tool. (A)** The positioning of DP values measured on the rigid ABB structures for each antibody (PDB IDs: 1DLF, 1MQK, 4GXV, 5WCA, 6MEG) in relation to the overlap between DP value distribution when measured on crystal MD and ABB MD ensembles for the structure-based solvent accessible surface area DP (AbStruc_sasa). The corresponding rigid DP values are shown as a black vertical line. Each subplot represents a separate antibody system. **(B)** Binary classification: 1 (black) indicates rigid ABB structure DP value is within the overlap of crystal MD and ABB MD structure DP distributions; 0 (white) – otherwise. Numbers shown at the bottom of each column represent the proportion of black cells for each antibody (PDB IDs: 1DLF, 1MQK, 4GXV, 5WCA, 6MEG). **(C)** Improvement in correlations of DP values across ABB and crystal data between the values from the rigid (red) and dynamic (blue) structures.



Supplementary Figure 6 | **Pairwise developability parameter Pearson correlation for the native human datasets.** Isotype-specific pairwise Pearson correlation coefficient for non-IgG human antibody datasets. Relates to Figure 2.



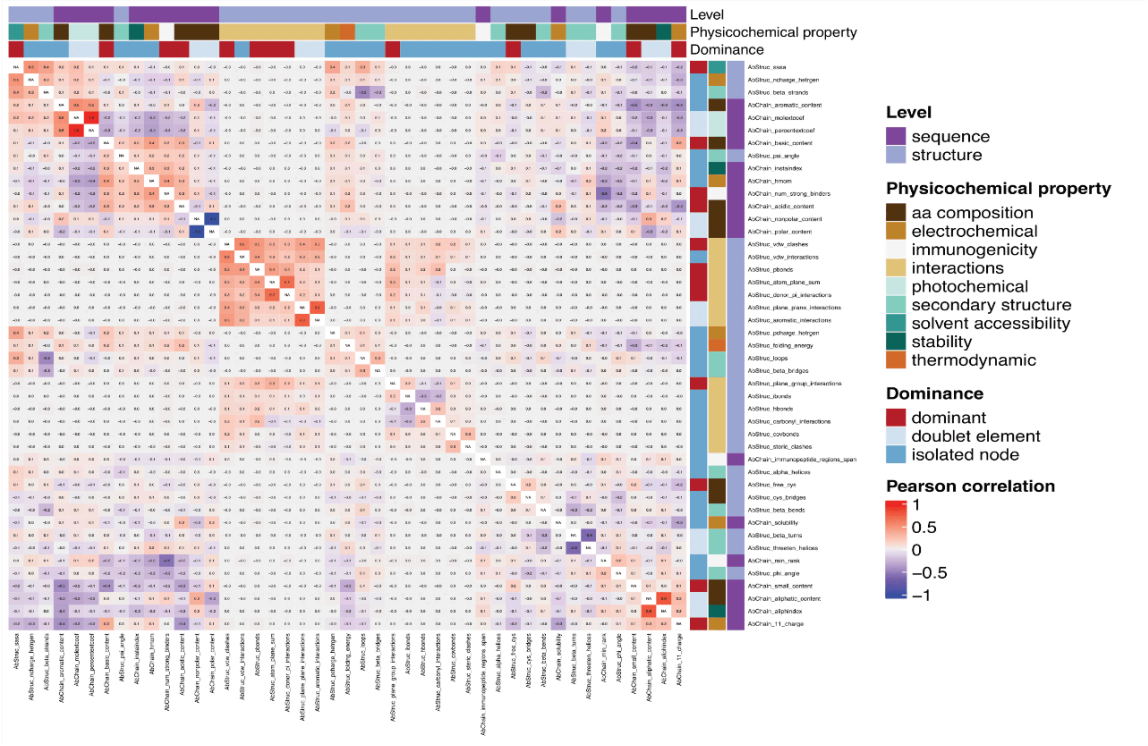
Supplementary Figure 7 | **Pairwise developability parameter Pearson correlation for the native murine datasets.** Isotype-specific pairwise Pearson correlation coefficient for murine antibody datasets. Relates to Figure 2.



Supplementary Figure 8 | **Comparison of pairwise Pearson correlation coefficients calculated on observed and permuted DP column values. (A)** Pairwise DP correlation matrix (equivalent to Figure 2A) where Pearson correlation coefficients were computed on permuted DP column values for the human IgG native dataset, i.e. every antibody from this datasets (170,473) was assigned to a permuted DP value within each DP column. **(B)** A comparison between the distribution of pairwise Pearson correlation coefficients computed on the original and the permuted DP column values for the same dataset (IgG human). A two-sample Kolmogorov-Smirnov test reported a significant difference between the two distributions (p-value < $2.2e^{-16}$). The permutation procedure was repeated 100 times, generating 100 permuted DP correlation matrices (n=2 biologically independent experiments repeated 100 times). The p-values of the Kolmogorov–Smirnov test were adjusted using the Benjamin-Hochberg method. All p-values reported statistical significance, thus the null hypothesis that the Pearson correlation coefficients computed on original and permuted DP column values are similar in their distributions was rejected. Relates to Figure 2.

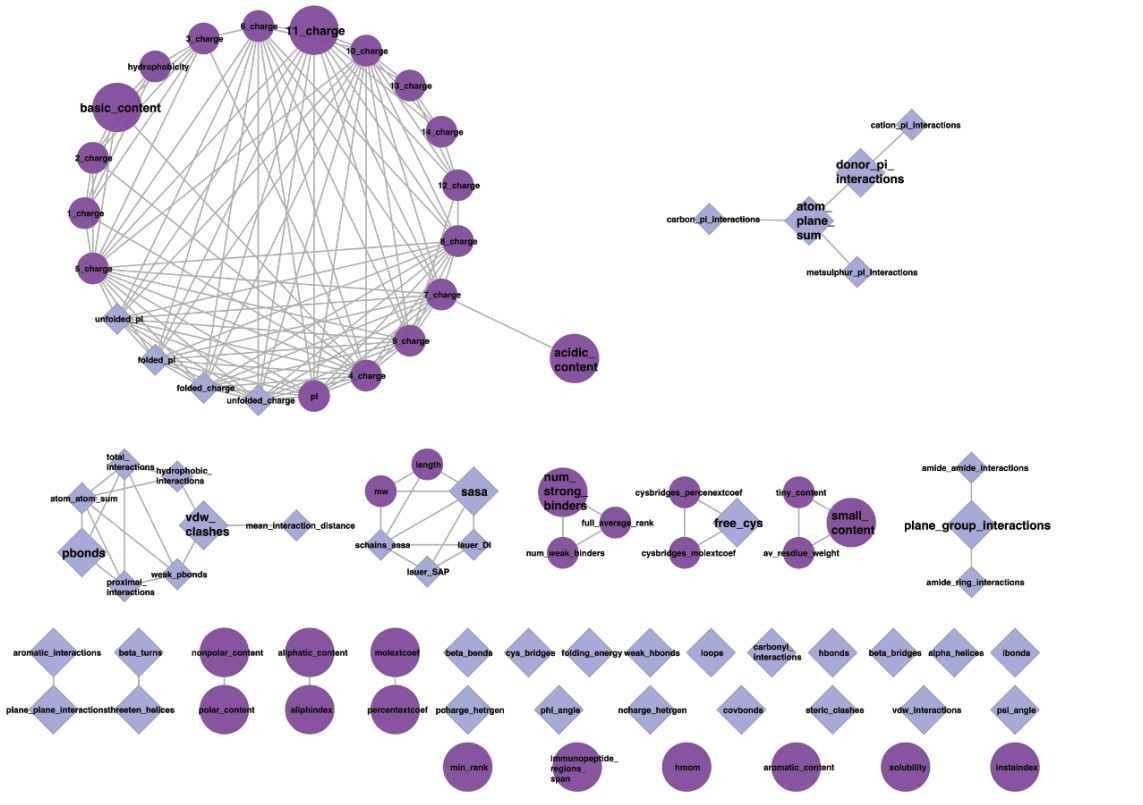
a Non-redundant DP correlation

Human IgG dataset, correlation threshold = 0.6

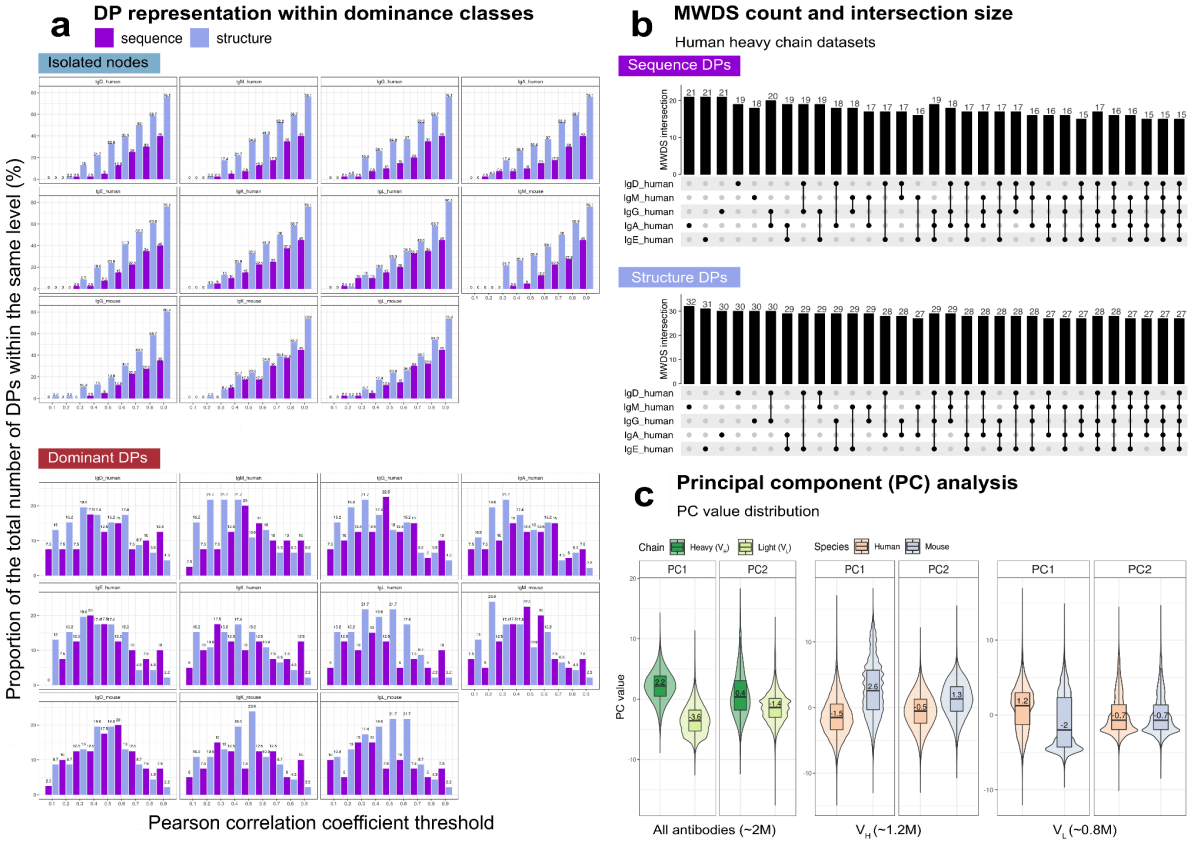


b DP correlation network

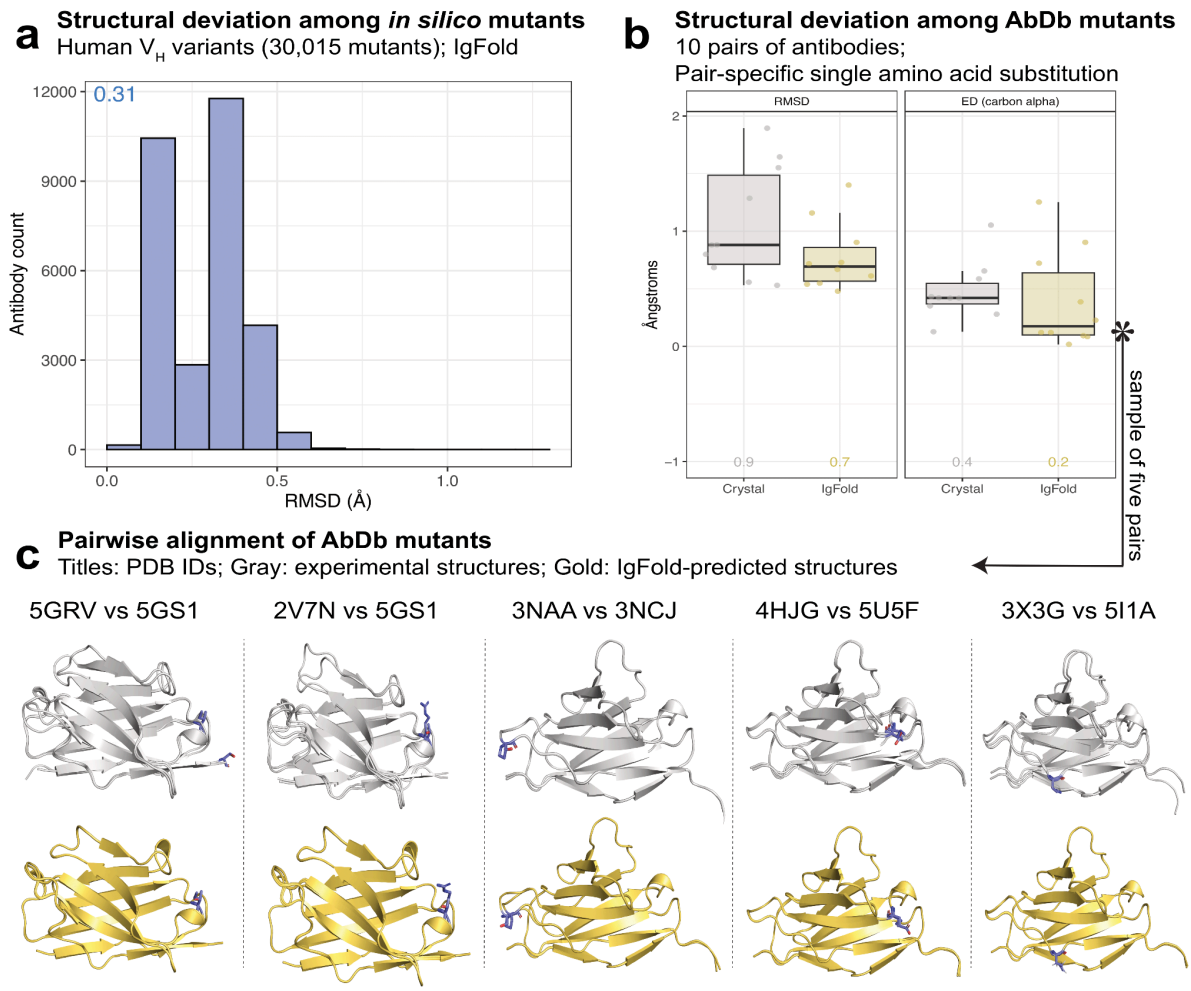
Human IgG dataset, correlation threshold = 0.6



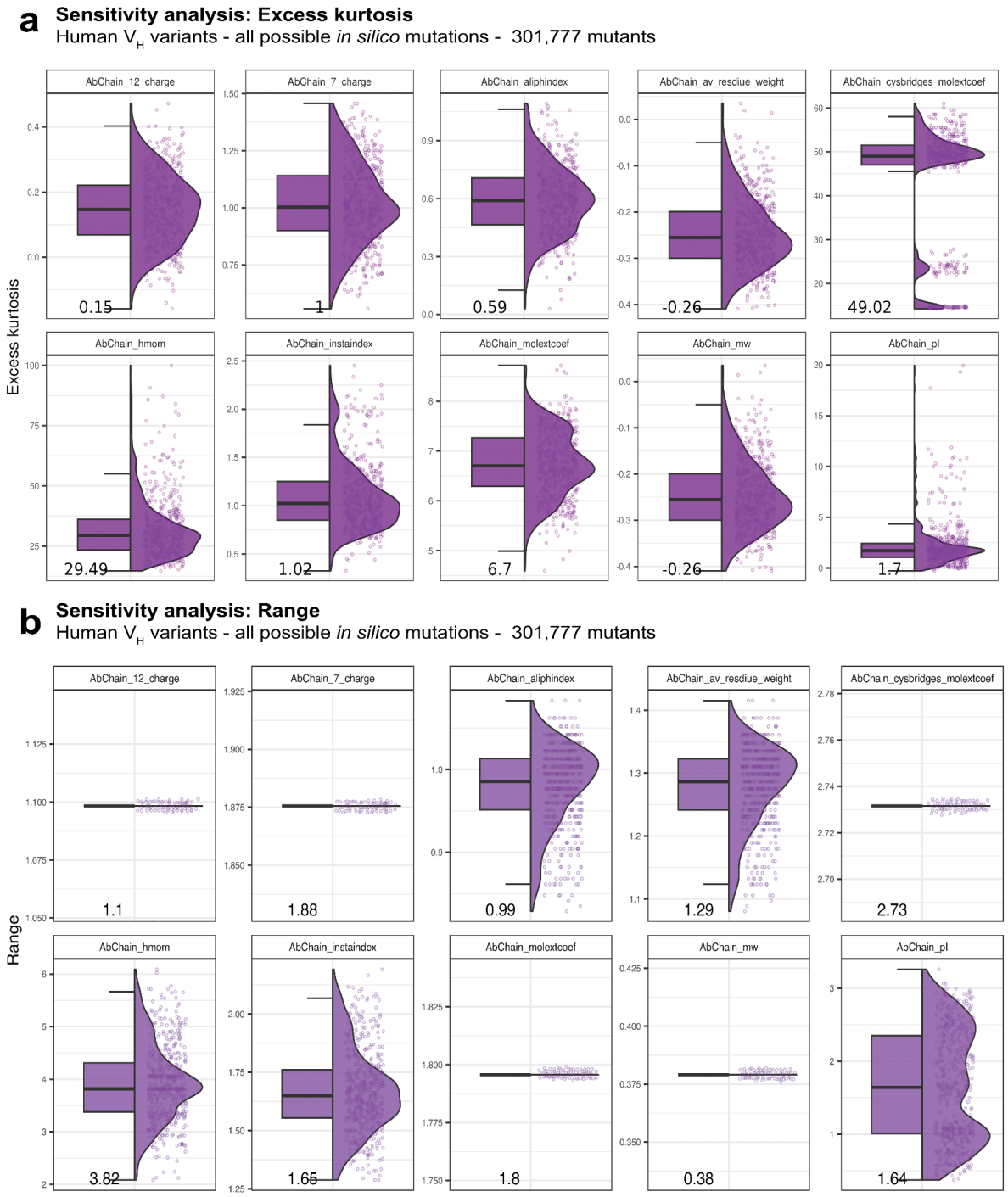
Supplementary Figure 9 | **The ABC-EDA algorithm facilitates the determination of non-redundant DPs for a specified Pearson correlation threshold. (A)** Pairwise correlation among non-redundant DPs as selected by the ABC-EDA algorithm for an absolute Pearson correlation threshold of 0.6, starting from the correlation heatmap presented in Figure 2B. Of note, doublet DPs were all retained in this analysis as no meaningful selection based only on correlation can be made by the algorithm **(B)** An undirected network graph of the pairwise parameter Pearson correlation data from (Figure 2B) with nodes representing DPs. An edge was drawn between a pair of nodes when the absolute pairwise Pearson correlation coefficient was > 0.6 (see Methods). Circular (dark orchid) nodes represent sequence DPs and square (cloudy blue) nodes represent structure ones. The node's size reflects its dominance state (large: dominant, small: redundant) as determined by the ABC-EDA output. Out of the 86 total parameters, 10 were classified as doublets, 23 as isolated nodes and 12 as dominant parameters for the human IgG repertoire. Structure-based developability parameters are majorly present as isolated nodes due to the lack of correlation with sequence parameters and among themselves (pairwise Pearson correlation mostly lower than 0.6). In contrast, sequence-based parameters are mainly present in larger subnetworks as they are more correlated. Relates to Figure 22.



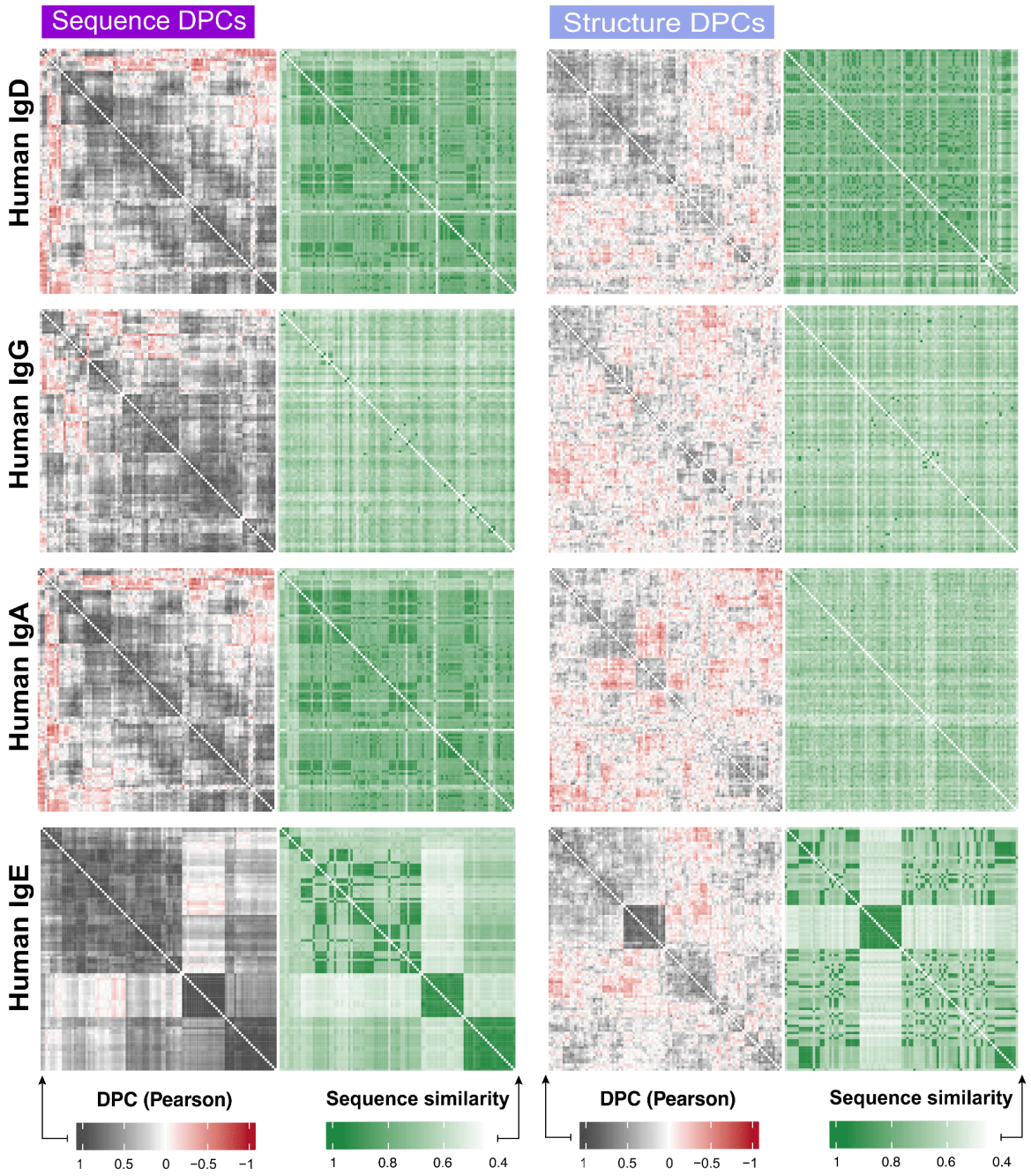
Supplementary Figure 10 | **Structure-based developability parameters exhibit lower pairwise association compared to sequence-based parameters. (A)** The proportion of DPs within the isolated nodes (top panel) and dominant (bottom panel) classes in relation to an ascending Pearson correlation cut-off (0.1–0.9, step size: 0.1) on sequence and structure levels. **(B)** MWDS intersection size for the human heavy chain datasets (IgD, IgM, IgG, IgA and IgE). The MWDS parameters were identified using the ABC-EDA algorithm (see Methods) at a threshold of absolute Pearson correlation of 0.6. The MWDS intersection study with the murine antibody datasets is shown in Figure 3A. **(C)** The value distribution of the first two principal components (PCs) when analyzing the entire native dataset (left panel ~2M antibodies), the heavy chains (middle panel, ~1.2M antibodies) and the light chains (right panel, ~0.8M antibodies). The numerical values annotated in the figure represent the median of the corresponding metric. Relates to Figure 2 and Figure 3.



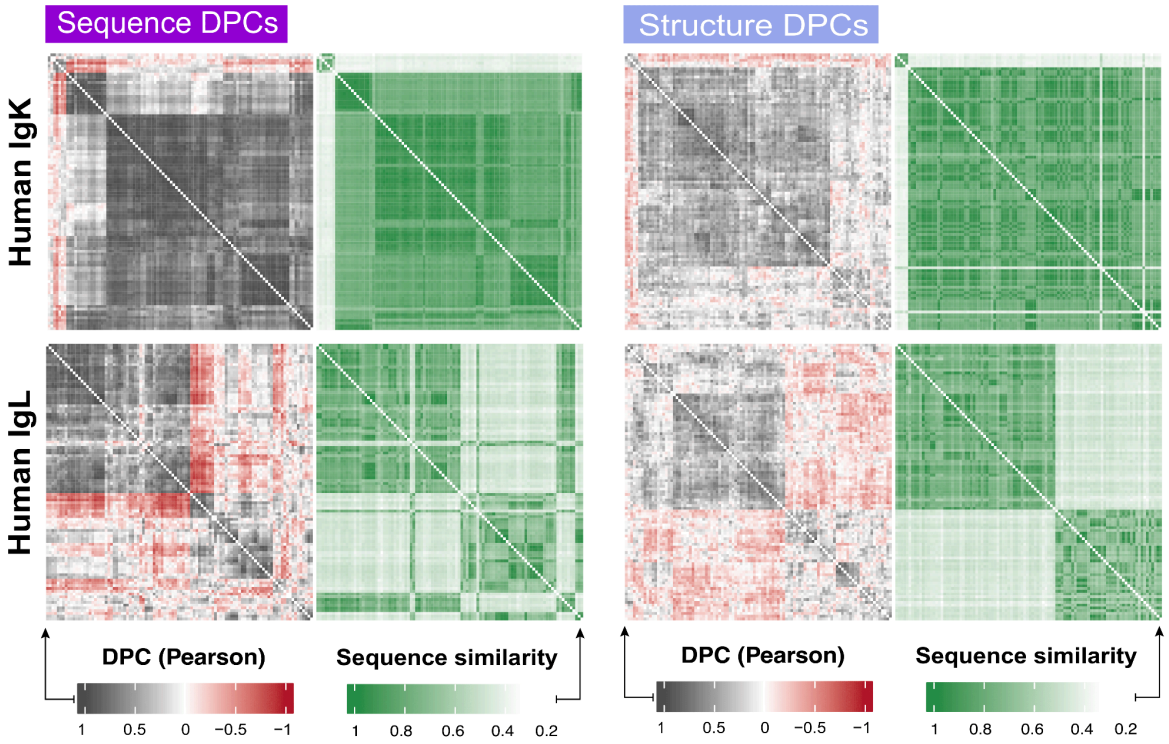
Supplementary Figure 11 | **Structural-DP variance induced by single amino acids is challenging to recapitulate. (A)** The distribution of RMSD values among the human V_H *in silico* mutants (30,015 antibodies). RMSD was measured between all CDR mutants and their corresponding WT (native) antibody after predicting their structures with IgFold (see Methods). The numerical value reported on the plot (0.31) represents the median RMSD. **(B)** The structural deviation among AbDb [16] antibody pairs measured with RMSD (global measure – left facet) and the Euclidean distance (ED) between carbon alpha atoms (focused measure – right panel). Of note, the carbon alpha of interest refers to the first carbon atom that bears the amino acid side chain in the exact location of sequence disparity between the antibodies of each pair (Supplementary Table 3). Both metrics (RMSD and $ED_{carbon\ alpha}$) were measured on crystal (experimental) structures and IgFold-predicted structures. The numerical values shown in the figure reflect the median of the corresponding metric. **(C)** Visualisation of the pairwise structural alignment of five antibody pairs (from B) shown for the crystal structures (top – gray) and IgFold-predicted structures (bottom – yellow). The side chains where the pair differ in a single mutation are drawn in all structures (represented in blue sticks). The PDB IDs of the pairs are shown on the top. Relates to Figure 4**.**



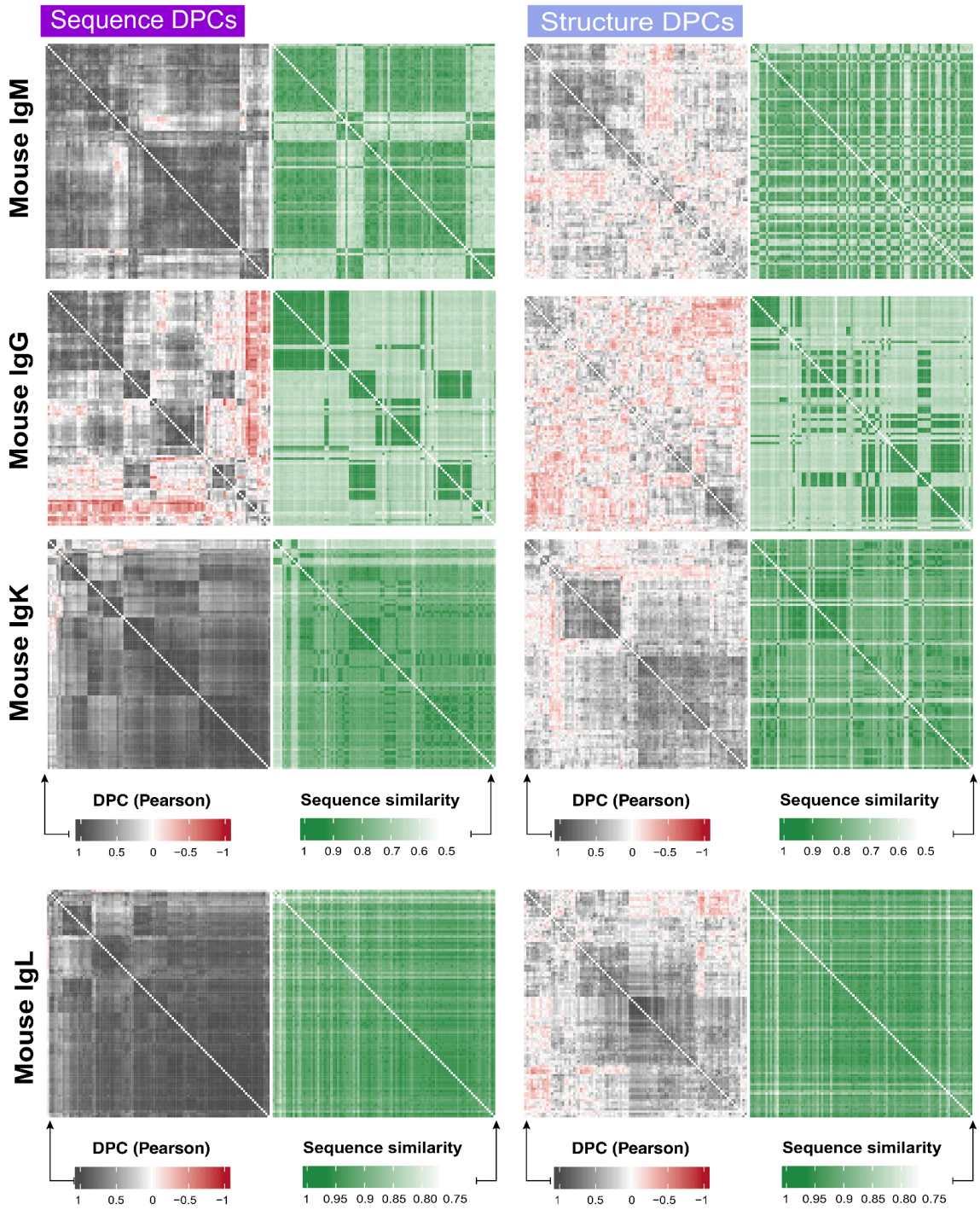
Supplementary Figure 12 | **Average and potential sensitivity of sequence-based developability parameters not shown in the main Figure 4.** DP values were computed for all possible single amino acid substituted variants of 500 sampled wildtype human heavy chain sequences (100 sequences sampled per isotype; 301,777 mutants in total). For each DP, the sensitivity was quantified by analyzing the variants' dispersion of DP values compared to their wildtype antibody. Average sensitivity **(A)** was measured by excess kurtosis (low kurtosis = high average sensitivity), while potential sensitivity **(B)** was measured by the range. Relates to Figure 4.



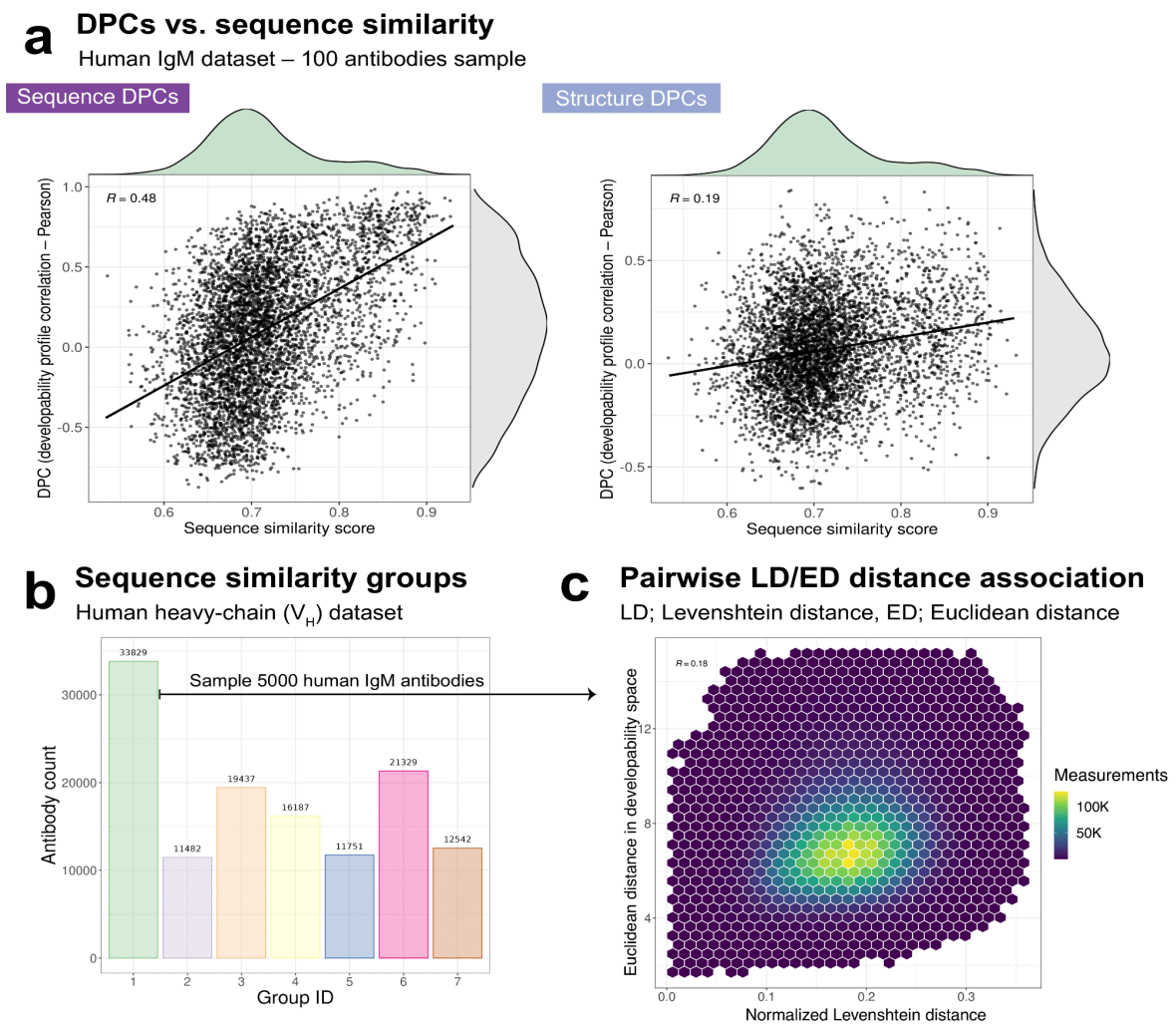
Supplementary Figure 13 | **The pairwise developability profile correlation (DPC) alongside the pairwise sequence similarity score for a random sample of 100 natural antibodies from the human heavy-chain datasets (IgD, IgG, IgA and IgE).** Each row and each column represent a single antibody sequence. Rows and columns in the DPC (left) panels were hierarchically clustered. In the sequence similarity (right) panels, rows and columns were ordered in the same order as the corresponding left panel for ease of comparison. Relates to Figure 5.



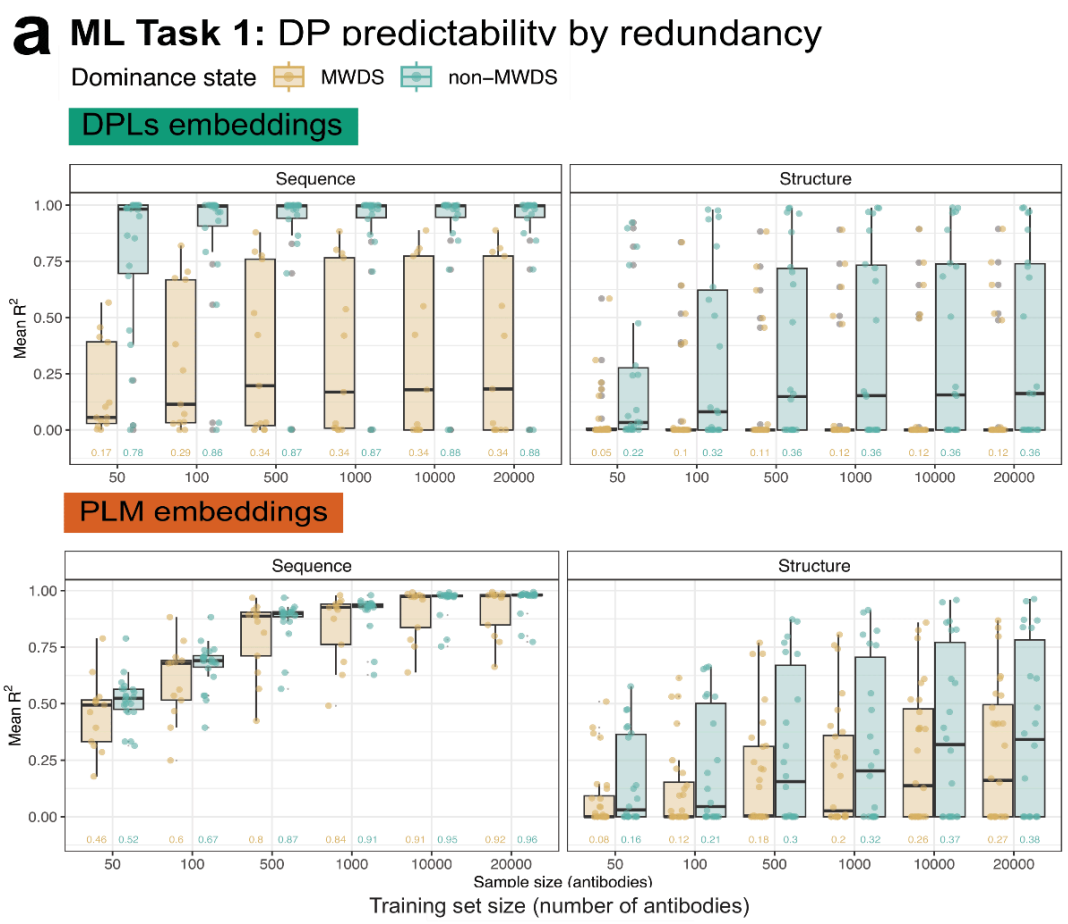
Supplementary Figure 14 | **The pairwise developability profile correlation (DPC) alongside the pairwise sequence similarity score for a random sample of 100 natural antibodies from the human light-chain datasets (IgK and IgL).** Each row and each column represent a single antibody sequence. Rows and columns in the DPC (left) panels were hierarchically clustered. In the sequence similarity (right) panels, rows and columns were ordered in the same order of the corresponding left panel for the ease of comparison. Relates to Figure 5.

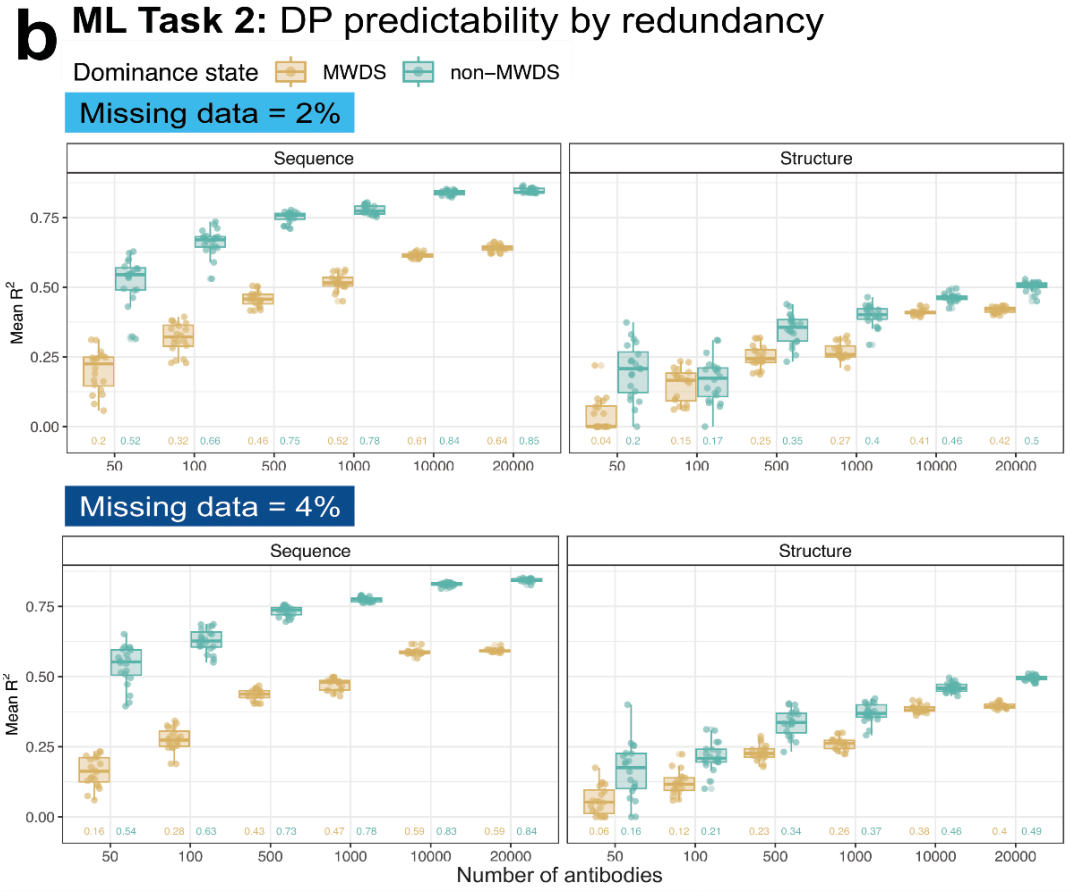


Supplementary Figure 15 | **The pairwise developability profile correlation (DPC) alongside the pairwise sequence similarity score for a random sample of 100 natural antibodies from the mouse datasets (IgM, IgG, IgK and IgL).** Each row and each column represent a single antibody sequence. Rows and columns in the DPC (left) panels were hierarchically clustered. In the sequence similarity (right) panels, rows and columns were ordered in the same order of the corresponding left panel for the ease of comparison. Relates to Figure 5.

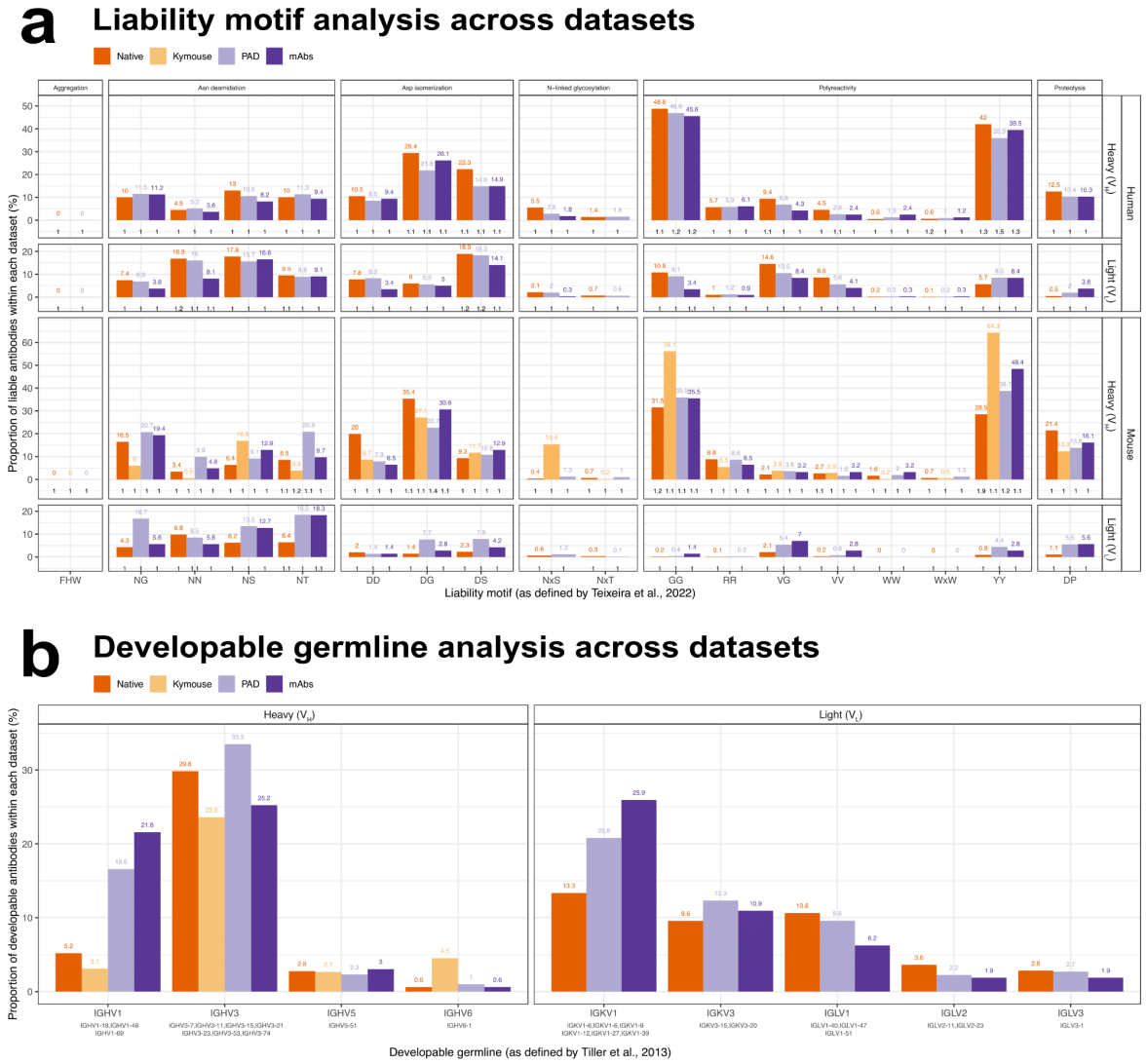


Supplementary Figure 16 | **Pairwise antibody sequence similarity does not correlate with pairwise antibody developability similarity. (A)** The distribution of the pairwise developability profile correlation (DPC) and pairwise sequence similarity values for a random sample of 100 natural human IgM antibodies that share the same IGHV gene family annotation. (shown in Figure 5A) This is shown on the sequence level (left panel) and the structure level (right panel). **(B)** The count of antibodies within each sequence similarity group (as shown in Figure 5C). Antibodies that belong to the same group share at least 75% sequence similarity (see Methods). **(C)** The relationship between pairwise (normalized) Levenshtein distance and pairwise Euclidean distance in the developability (*R*) space for 5000 human IgM antibodies that belong to sequence similarity group 1 and share the same IGHV gene family annotation. Out of the total (12,497,500) pairwise measurements, we removed 0.8% where the Euclidean distance was greater than 15 to avoid the effect of outliers on data visualization. Numerical values on the figures in **(A)** and **(C)** represent the Pearson correlation coefficient. Relates to Figure 5.


a
ML Task 1: DP predictability by redundancy
Dominance state
MWDS
non–MWDS
DPLs embeddings
Sequence
Structure
Mean R²
PLM embeddings
Sample size (antibodies)
Training set size (number of antibodies)

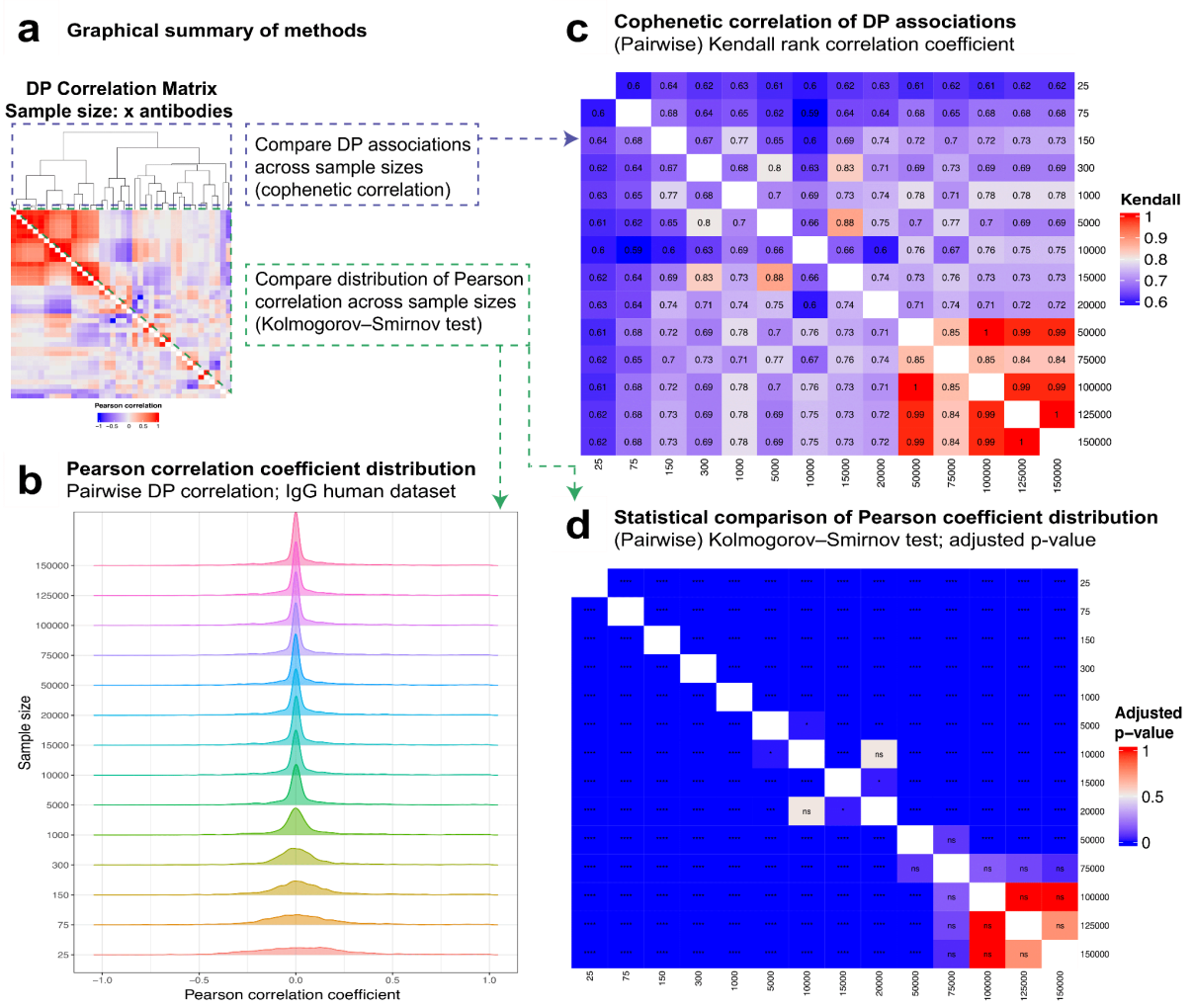

b
ML Task 2: DP predictability by redundancy
Dominance state
MWDS
non–MWDS
Missing data = 2%
Sequence
Structure
Mean R²
Missing data = 4%
Number of antibodies

Supplementary Figure 17 | **Developability parameters belonging to the MWDS are more challenging to predict than (redundant) non-MWDS ones. (A)** Evaluating the predictive power of single-DP-wise incomplete developability profiles embeddings (top) and PLM-based embedding (bottom) to predict the missing values of (non-redundant) MWDS DPs (yellow lines) and (redundant) non-MWDS DPs using multiple linear regression (MLR – ML Task 1). MWDS developability parameters were identified as previously explained in Figure 6B and listed in Supplementary Table 2 (13 sequence-based and 28 structure-based). To evaluate the predictability of non-MWDS parameters, we included an equal number of (randomly sampled) sequence-based DPs (13) and all the remaining structure DPs that do not belong to the MWDS. X-axis reflects the number of antibodies used for the model embedding (sample size). For each sample size, we repeated the prediction of missing DPs 20 times (20 independent subsamples). Y-axis represents the mean coefficient of determination (R^2) for sequence DPs (left facets) and structure DPs (right facets). **(B)** Evaluating the predictive power of cross-DP-wise incomplete developability profiles embeddings to predict the missing values of (non-redundant) MWDS DPs (yellow lines) and (redundant) non-MWDS DPs multivariate imputation by chained random forests algorithm (MICRF – ML Task 2). MWDS and non-MWDS DPs were identified similarly to (A). The y-axis represents the mean R^2 for sequence DPs (left facets) and structure DPs (right facets) when the proportion of the missing data is either 2% (top) or 4% (bottom). Numbers on the x-axis in both (A) and (B) represent the mean values of mean R^2. Relates to Figure 6.

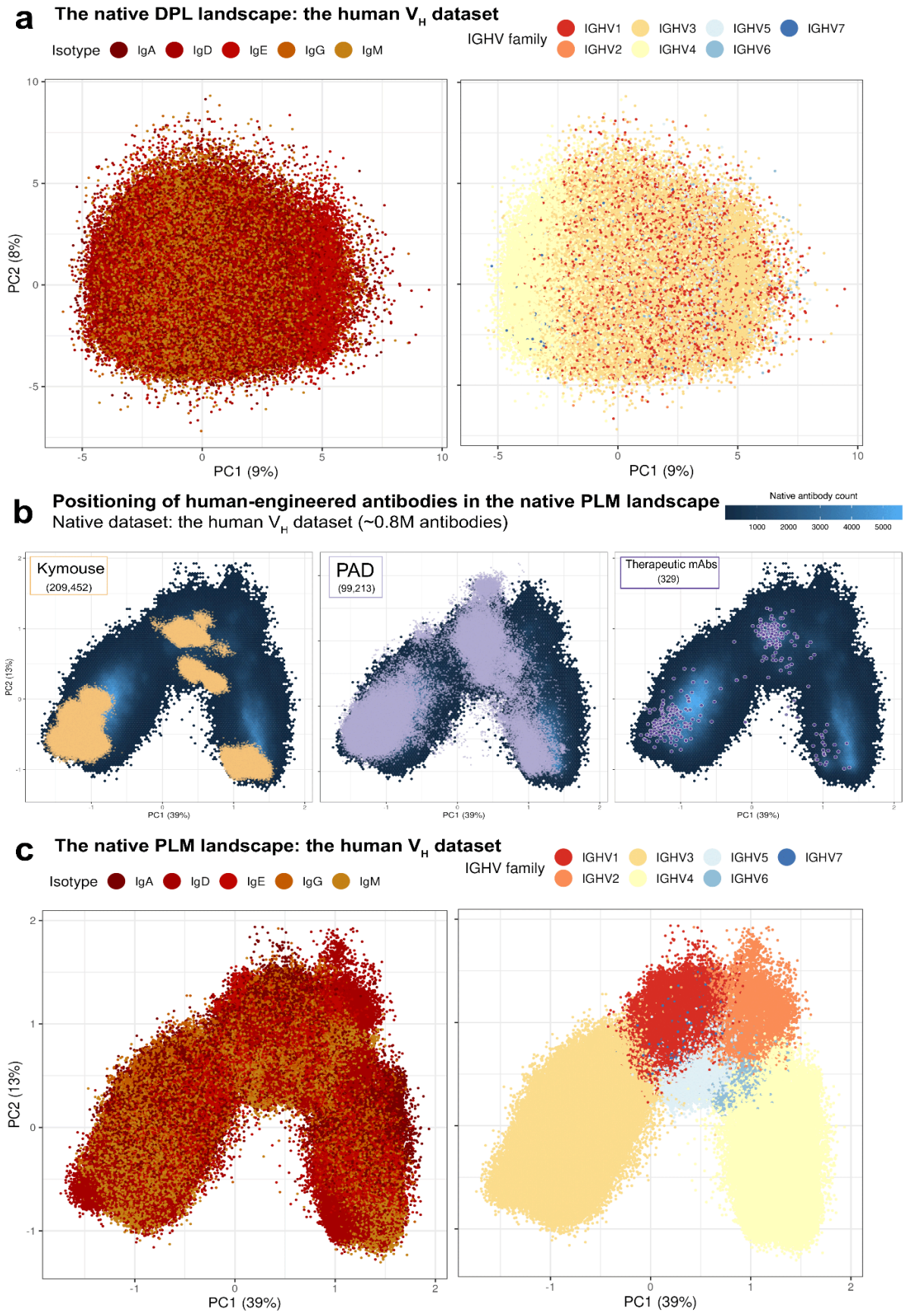


Supplementary Figure 18 | **Liability motifs at the native and human-engineered antibody datasets.** Proportional abundance of (**A**) liability motifs and **(B)** developable germlines across the native (human and murine) antibody datasets (Native – Supplementary Figure 1A), and the human-engineered datasets (humanized mouse antibodies: Kymouse, patented antibody dataset: PAD, and therapeutic monoclonal antibodies: mAbs). In (A), the x-axis refers to amino acid sequence liability motifs as identified by Teixeira and colleagues [17]. The figure is segmented into columns (sequence liability type) and

rows (antibody species and chain type). The height of the bars and the color-coded text (above them) represent the proportion of antibodies (%) that contain the respective liability motif at least once. The black text on the x-axis refers to the average number of occurrences of liability motifs within liable antibodies (minimum of 1). In (**B**), the x-axis represents the IGHV gene family followed by the developable germline genes (below) within each family, as reported by Tiller and colleagues [18]. The figure is segmented into columns (antibody chain type), and the height of the bars and color-coded text (above them) represent the proportion of antibodies (%) that were annotated with developable germline genes within the respective chain type. Relates to Figure 7.

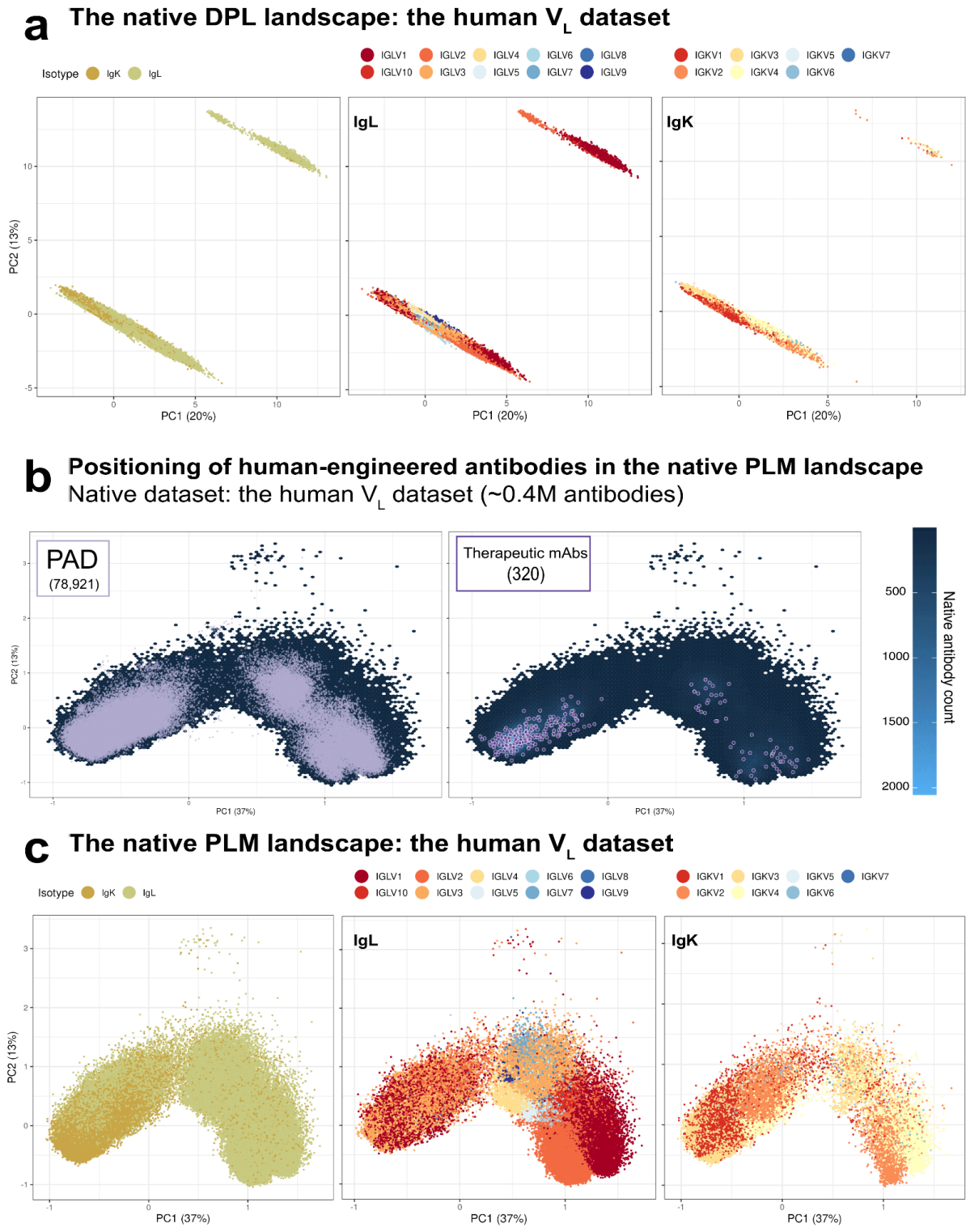


Supplementary Figure 19 | **A minimum of 50K antibodies is required to stabilize the pairwise associations among DPs.** **(A)** Graphical summary of methods followed in this analysis in which we compare DP correlation matrices as a function of sample size. Two pieces of information can be obtained from correlation matrices being (i) DP associations (clustering dendrograms – blue dashed lines) and (ii) Pearson correlation coefficients (green dashed lines). We compare the DP associations (in C) and the distribution of Pearson correlation coefficient values (in B and D) as extracted from pairwise DP correlation matrices for randomly sampled antibodies from the human IgG dataset with increasing antibody sequence sample size (25, 75, 150, 300, 1000, 5000, 10000, 15000, 20000, 50000, 75000, 100000, 125000, 150000 sequences). **(B)** Visual comparison of the distribution of Pearson correlation coefficient values as a factor of increasing antibody sequence sample size. **(C)** Pairwise Kendall correlation coefficient was computed among the hierarchical clustering dendrograms extracted from pairwise DP correlation matrices computed on. At least 50K antibodies are required to sustain the associations (dendrograms) among DPs (Kendall correlation coefficient ≥ 0.84). (**D**) Statistical comparison of the distribution of Pearson correlation coefficient values as a factor of increasing antibody sample size using the Kolmogorov-Smirnov test. At least 75K antibodies are required to preserve the distribution of Pearson correlation values among DPs (the distribution difference becomes statistically non-significant from 75K antibodies onwards). Each cell in the heatmap represents a comparison of two distributions; n=2 biologically independent samples. ns; not-significant, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ and **** $p < 0.0001$. Relates to Figure 7.



Supplementary Figure 20 | **The role of germline gene family annotation and isotype in the localization of V_H antibodies in the native developability and PLM-based sequence landscapes. (A)** The distribution of the native V_H human antibodies

(854,418 antibodies) in the developability landscape (DPL-based PCA analysis) shown by antibody isotype (left) and IGHV gene family (right). **(B)** The positioning of the human-aligned human-engineered V_H antibodies (Kymouse; 209,452, PAD; 99,213 and therapeutic mAbs; 329) in the PLM space of the native human V_H datasets (854,418 antibodies). The hexagonal bins (shown in the back layer) represent the count of native antibodies (scale shown on the top right of the panel), while the human-engineered antibodies are represented as data points. **(C)** The distribution of the native V_H human antibodies (854,418 antibodies) in the protein language model landscape (PLM-based PCA analysis) shown by antibody isotype (left) and IGHV gene family (right). Antibody clustering based on germline gene family annotation is observed due to the high sequence similarity (≈80%) among antibodies that belong to the same V-gene family, resulting in PLM-based clustering biases [19]. Relates to Figure 7.



Supplementary Figure 21 | **The role of germline gene family annotation and isotype in the localization of V_L antibodies in the native developability and PLM-based sequence landscapes. (A)** The distribution of the native V_L human antibodies (386,173 antibodies) in the developability landscape (DPL-based PCA analysis) shown by antibody isotype (left), IGLV gene family (middle) and IGKV gene family (right). **(B)** The positioning of the human-aligned human-engineered V_L antibodies (PAD; 78,921 and therapeutic mAbs; 320) in the PLM space of the native human V_L dataset (386,173 antibodies). The hexagonal bins (shown in the back layer) represent the count of native antibodies (scale shown on the right of the panel), while, the human-engineered antibodies are represented as data points. **(C)** The distribution of the native V_L human antibodies (386,173 antibodies) in the protein language model landscape (PLM-based PCA analysis) shown by antibody isotype, IGLV gene family (middle – for IgL antibodies) and IGKV gene family (right – for IgK antibodies). Antibody

clustering based on germline gene family annotation is observed due to the high sequence similarity (≈80%) among antibodies that belong to the same V-gene family, resulting in PLM-based clustering biases [19]. Relates to Figure 7.

Supplementary Notes

Supplementary Note 1: Verification of strong correlation between DP values measured on single and paired-chain antibody structures

We found that the values of structure-based DPs measured on a subset of 859 paired-chain antibodies (see Methods) strongly correlated with the corresponding DP values measured on each of the unpaired single chains (V_H and V_L) separately (median Pearson correlation 0.84–0.90, Supplementary Figure 2) thus supporting the notion that some DP work performed on single-chains data translates to paired-chain data.

Supplementary Note 2: Analysis of dependence of structure-based DPs on computational antibody structure prediction method

We set out to verify to what extent structure DPs are stable across computational antibody structure prediction methods. We performed these analyses both for rigid (Supplementary Figure 3) and dynamic antibody models (i.e., molecular dynamics, Supplementary Figure 4) starting from 859 reference (crystal) antibody structures obtained from AbDb (crystal structures dataset, see Methods) [16].

(1) Rigid model analysis:

The rigid structure prediction tools included in this analysis were ABodyBuilder [12], ABodyBuilder2 [13], IgFold [14], AlphaFold2 [10], and AlphaFold-multimer [11]. For this analysis, we measured the structural variance (RMSD – Supplementary Figure 3B) and the correlation of DP values (Pearson correlation – Supplementary Figure 3A) between crystal structures and their predicted models. We conducted this analysis on V_H-V_L paired structures as well as V_H and V_L (unpaired) structures separately (Supplementary Figure 3A,B).

ABB-predicted rigid structures harbored minimal atomic distances from their reference counterparts (median RMSD = 0.34, 0.2 6, 0.27 ordered for paired, V_H and V_L structures – Supplementary Figure 3B), stemming from high structural alignment with reference antibodies (example shown for the antibody 1DLF in Supplementary Figure 3C). Nevertheless, despite this seemingly high structural similarity, the correlation of the DP values measured on ABB structures with those measured on crystal structures was low (median Pearson correlation coefficient = 0.31, 0.4, 0.31 for paired, V_H and V_L respectively – Supplementary Figure 3A). Moreover, we reported similar low correlation coefficients between DP values measured on crystal structures and all predicted structures included in this study (Supplementary Figure 3A). These observations were reported despite ABB being a former software release in comparison to the remaining tools included in this study. These results demonstrate the dependence of structure-based antibody developability predictions on the choice of computational structure prediction method. A similar observation was made recently by Jain and colleagues [20].

Given the differences in DP values among the structure prediction tool, we asked in the next section whether these differences may be explained by comparing rigid structures from frames belonging to a conformational ensemble (dynamic continuum).

(2) Molecular dynamics analysis:

We investigated whether the rigid ABB-predicted structures represented a mere snapshot of the structural conformations of the corresponding crystal (reference) antibodies when examined with molecular dynamics simulations (MD). To this end, we included five paired-chain antibodies in this analysis, prioritizing those with the highest structural resolution within the crystal structure dataset (PDB IDs: 1DLF, 1MQK, 4GXV, 5WCA, 6MEG – Supplementary Table 4). The selection of sample size (five) and antibody structure type (paired-chain) was made considering (i) the computational cost associated with molecular dynamics simulations, and (ii) the strong correlation between single and paired-chain structure-based antibody developability as previously shown (Supplementary Figure 2). We simulated the crystal and ABB-predicted structures of these antibodies over 100 ns resulting in 5001 frames/antibody, and confirmed structural convergence for all antibodies (crystal and ABB-predicted) from 20 ns onwards (see Methods – Supplementary Figure 4A). Examining the pairwise structural variance (measured with RMSD) among convergent ABB frames (4001 frames/antibody) and convergent reference frames (4001 frames/antibody) revealed limited structural disparities (Supplementary Figure 4B). Indeed, ABB frames were minimally distant from their reference counterparts (median pairwise RMSD 1DLF;1.34Å, 1MQK;1.12Å, 4GXV;1.1Å, 5WCA;1.08Å, 6MEG;0.97Å – Supplementary Figure 4B). Importantly, the distribution of their structural DP values highly overlapped with those of the crystal (reference) equivalents with a mean overlap index (η) of 0.73, 0.65, 0.77, 0.85, 0.77 for 1DLF, 1MQK, 4GXV, 5WCA, 6MEG, respectively (Supplementary Figure 4C), suggesting that the difference in structure-based DPs as a function of 3D-structure origin (be it experimental or computational) may be a consequence of undersampling the dynamics of the antibody structure space. Future research in ML-based MD may enable repertoire-level MD as, currently, performing MD calculations beyond a handful of structures is prohibitively computationally expensive [21].

(3) Rigid model vs. Molecular dynamics analysis:

As current MD approaches are computationally expensive and time-consuming, scientists rely on rigid antibody structures as the primary models for computational developability studies. Hence, without an experimental structure determination, we set out to examine the applicability of rigid ABB models to possible real-world DP investigations. We hypothesized that when structure DP values calculated on rigid ABB structures are within the overlap between the conformational ensembles of crystal MD and ABB MD for a given antibody, then rigid ABB models can replicate to some extent the behavior of the real molecules and thus reflect the real values of DPs. ABB was used as an example of the rigid antibody structure prediction method since we applied this tool to model the native and human-engineered datasets (Supplementary Figure 1, see Methods).

First, we investigated the consistency of structure DP values calculated based on crystal MD, ABB MD, and ABB rigid. Computationally, the closest entity to real-world scenarios are MD simulations

of experimentally determined structures (crystal MD). Above, we showed a noticeable overlap between crystal MD and ABB MD distributions (Supplementary Figure 4A). Since crystal MD is a baseline of accuracy, when we compare crystal MD with ABB MD, the overlap in a given DP values when measured on both MDs would show where the most reliable ABB conformations are placed. Thus, crystal MD distributions are considered an "oracle" that highlights (1) the most solid parts of ABB MD distributions where the structural conformations and hence structure DPs are the closest to the real world (in the resolution of force field accuracy) and (2) the most reliable rigid DPs that can exist in dynamics as they also intersect with ABB MD. To this end, we investigated where the rigid ABB models are located in relation to the DP value distribution overlap between ABB MD and crystal MD Supplementary Figure 5A,B. If a rigid ABB structure DP value is within the overlap, the computational conclusions made with the DPs from the overlap are more likely to remain similar in the experimental verification. For example, the values of the solvent accessible surface area (AbStruc_sasa) measured on the rigid ABB structures were within the DP value distribution overlap for 1MQK and 6MEG, while for the remaining antibodies (1DLF, 4GXV, and 5WCA) the rigid values were outside the distribution overlap (Supplementary Figure 5A).

To follow up on this observation quantitatively, we performed a binary categorization (separation into two groups) if the rigid ABB DP value is in the overlap between crystal MD and ABB MD DP value distributions stratified by antibody (Supplementary Figure 5B). Investigating rigid ABB at the overlap, we found that on average, 60.89% of rigid ABB DP parameters are within the DP value overlap (1DLF: 46.67%, 1MQK: 57.78%, 4GXV: 64.44%, 5WCA: 66.67%, 6MEG: 68.89%). These findings suggest that ABB models, even without MD simulations, might have the potential for predicting structure DPs, as they capture aspects of the conformational landscape observed in experimentally determined antibodies. This implies a practical link between the models and real-world scenarios (Supplementary Figure 4, Supplementary Figure 5).

To further speculate on the poor correlation between DP values measured on rigid ABB and rigid crystal structures (Supplementary Figure 3A), we investigated a method to improve these correlations by employing DP values measured on the ensembles of structural conformations. We hypothesized that MD would reduce the difference caused by using rigid structures since it decreases the noise caused by the non-optimal amino acid side chain conformations. An increased correlation suggests the importance of dynamics and MD simulations as a valuable tool for in silico DP investigations.

To improve the correlation between structure DP across different sources of antibody structures (ABB models and crystal data), we computed the DP values for each converged conformation from MD simulations of five selected antibodies (1DLF, 1MQK, 4GXV, 5WCA, 6MEG). Next, we calculated the average of these values (one value per antibody conformation) for each structure DP parameter. We correlated averaged ABB MD with averaged crystal MD (blue in Supplementary Figure 5C), and compared that with the correlation of rigid ABB with rigid crustal (red bars in Supplementary Figure 5C). We observed a substantial improvement in DP value correlations across ABB and crystal data between the values from the rigid (red) and dynamic (blue) structures (rigid DP value range: 0.07–0.80, mean: 0.36; MD DP value range: 0.00–1.00, mean: 0.76) (Supplementary Figure 5C). We conclude that MD simulations reduced the intra-DP variation caused by the single rigid conformation from the conformational ensemble, highlighting the feasibility of using only one structure prediction tool with an additional conformational ensemble rather than many (rigid) structure prediction tools.

To conclude, we found that despite the poor correlation of DP values between rigid structures from different prediction tools, the computational DPs are real-world informative since many rigid DPs are

within the possible DP ranges defined by crystal MD. Hence, the models yielding the DP parameters are meaningful as they fall within the conformation ensemble produced by MD simulations (real-world possible antibody structures within the resolution of the MD force field error). Moreover, MD improves poor correlations between structure DP obtained from models predicted by different tools (Supplementary Figure 5). Moreover, although the rigid correlations were poor, the models yielding the DP parameters are meaningful as they fall within the possible ensemble of conformations produced by MD simulations (real-world possible antibody structures within the resolution of the MD force field error). Our findings are in line with recent MD-driven reports regarding structure-based DPs [22–24], where an improved correlation between DP values measured on experimental and predicted structures was reported only after averaging the DP values of the two conformational ensembles.

Supplementary Note 3: Feasibility analysis of inferring structures of close-distance antibody mutants

To potentially complement the sequence-based sensitivity analysis (Figure 4, Supplementary Figure 12) with a structure-based one, we set out to investigate whether computational antibody structure prediction methods may be suitable for sensitivity analysis where structure-based DPs of *single* amino acid substituted antibody mutants would be compared to that of their wildtype. To this end, we studied the predicted structures of all single-amino acid substituted CDR variants of 49 sampled heavy chain sequences (10 per isotype except 9 for IgD; 30,015 mutants in total – see Methods). Of note, we performed structure prediction for this analysis using the deep learning-based tool IgFold [14] because template search-based tools (e.g., ABB) tend to utilize an identical structural template for antibodies that are single-amino-acid-distant from their wildtype antibody due to the high sequence similarity [25,26]. We found that structural deviations induced by single amino acid substitutions in CDRs (as determined by RMSD) between wildtype and mutant were minimal (0.09–1.2 Å) with a median value of 0.31 Å (Supplementary Figure 11A). This range of RMSD values is within or below the current technical resolution of experimental protein crystallography (RMSD ≈ 2Å) [16,27].

Thus, we asked whether these findings obtained by computational structure prediction are comparable with the variation observed among antibody structures determined experimentally by crystallography. To this end, we compiled a dataset of ten antibody pairs from the AbDb database [16] where antibodies of the same pair are one-amino-acid different and are similar in sequence length (see Methods, Supplementary Table 3). Of note, the small size of this dataset is due to the limited availability of antibody crystal structure pairs that satisfy these criteria (same length and one amino acid difference in the loop).

We found that the structural variations within pairs were greater on average for experimental and IgFold predicted structures (median RMSD 0.9 and 0.7 Å, respectively – Supplementary Figure 11B) in comparison to what we reported for the *in silico*-generated native antibody mutants (Supplementary Figure 11A). To further investigate the difference between experimental and IgFold-predicted models, we visually inspected the structural alignment of five randomly-sampled antibody pairs from the same dataset (Supplementary Figure 11C). In the case of IgFold models (shown in gold, Supplementary Figure 11C), antibody pairs exhibited complete superposition, while crystal structures (shown in gray, Supplementary Figure 11C) displayed subtle global variations. Nevertheless, it is challenging to solely attribute such variations to the single amino acid difference within each pair as several factors could be the origin of variance including (i) the inherent limitations

and uncertainties associated with experimental variability to resolve each individual structure within the pair [28,29] and (ii) the flexibility of antibody structures, especially their CDR loops [30,31]. To limit the effect of external factors (as much as possible), we focused on the exact locus of amino acid difference by computing the Euclidean distance (ED) between carbon alpha atoms ($ED_{carbon\ alpha}$) within each pair (Supplementary Figure 11B). In both types of structures (crystal and IgFold-predicted), we reported negligible $ED_{carbon\ alpha}$ values (median of 0.4 (crystal) and 0.2 (IgFold) Å).

Based on our analysis, and considering the (current) shortfalls of antibody structure prediction tools and experimental variance, we hypothesize that the actual distance values among antibody structural variants, particularly at the mutation site, to be lower than experimentally-measured values and higher than computationally-predicted values. This indicates that the actual impact of single-amino-acid antibody structural variance might fall somewhere between that of the predicted computationally and that of measured experimentally. Nevertheless, as *in silico* structure-based developability prediction relies on 3D antibody structures, this analysis highlighted how current technical insufficiency hinders conducting a single-amino-acid sensitivity analysis on structure-based antibody developability parameters. For the above reasons, we limited our sensitivity analysis to sequence-based DPs (Figure 4, Supplementary Figure 12).

References

1. Osorio, D., Rondon-Villarreal, P. & Torres, R. Peptides: A Package for Data Mining of Antimicrobial Peptides. *The R Journal* vol. 7 4–14 Preprint at (2015).
2. Cui, A. *et al.* A Model of Somatic Hypermutation Targeting in Mice Based on High-Throughput Ig Sequencing Data. *J. Immunol.* **197**, 3566–3574 (2016).
3. Greiff, V. *et al.* Systems Analysis Reveals High Genetic and Antigen-Driven Predetermination of Antibody Repertoires throughout B Cell Development. *Cell Rep.* **19**, 1467–1478 (05/2017).
4. Gupta, N. T. *et al.* Hierarchical Clustering Can Identify B Cell Clones with High Confidence in Ig Repertoire Sequencing Data. *J.I.* **198**, 2489–2499 (2017).
5. Kovaltsuk, A. *et al.* Observed Antibody Space: A Resource for Data Mining Next-Generation Sequencing of Antibody Repertoires. *The Journal of Immunology* **201**, 2502–2509 (2018).
6. Ohm-Laursen, L. *et al.* Local Clonal Diversification and Dissemination of B Lymphocytes in the Human Bronchial Mucosa. *Front. Immunol.* **9**, 1976 (2018).
7. Thörnqvist, L. & Ohlin, M. Data on the nucleotide composition of the first codons encoding the complementary determining region 3 (CDR3) in immunoglobulin heavy chains. *Data Brief* **19**, 337–352 (2018).
8. Raybould, M. I. J. *et al.* Thera-SAbDab: the Therapeutic Structural Antibody Database. *Nucleic*

Acids Res. **48**, D383–D388 (2020).

9. Richardson, E. *et al.* Characterisation of the immune repertoire of a humanised transgenic mouse through immunophenotyping and high-throughput sequencing. *Elife* **12**, (2023).
10. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
11. Evans, R. *et al.* Protein complex prediction with AlphaFold-Multimer. *bioRxiv* 2021.10.04.463034 (2022) doi:10.1101/2021.10.04.463034.
12. Leem, J., Dunbar, J., Georges, G., Shi, J. & Deane, C. M. ABodyBuilder: Automated antibody structure prediction with data–driven accuracy estimation. *MAbs* **8**, 1259–1268 (2016).
13. Abanades, B. *et al.* ImmuneBuilder: Deep-Learning models for predicting the structures of immune proteins. *Commun Biol* **6**, 575 (2023).
14. Ruffolo, J. A., Chu, L.-S., Mahajan, S. P. & Gray, J. J. Fast, accurate antibody structure prediction from deep learning on massive set of natural antibodies. *Nat. Commun.* **14**, 2389 (2023).
15. Pastore, M. & Calcagnì, A. Measuring Distribution Similarities Between Samples: A Distribution-Free Overlapping Index. *Front. Psychol.* **10**, 1089 (2019).
16. Ferdous, S. & Martin, A. C. R. AbDb: antibody structure database-a database of PDB-derived antibody structures. *Database* **2018**, (2018).
17. Teixeira, A. A. R. *et al.* Simultaneous affinity maturation and developability enhancement using natural liability-free CDRs. *MAbs* **14**, 2115200 (2022).
18. Tiller, T. *et al.* A fully synthetic human Fab antibody library based on fixed VH/VL framework pairings with favorable biophysical properties. *MAbs* **5**, 445–470 (2013).
19. Tran, C., Khadkikar, S. & Porollo, A. Survey of Protein Sequence Embedding Models. *Int. J. Mol. Sci.* **24**, (2023).
20. Jain, T., Boland, T. & Vásquez, M. Identifying developability risks for clinical progression of antibodies using high-throughput in vitro and in silico approaches. *MAbs* **15**, 2200540 (2023).
21. Noé, F., Tkatchenko, A., Müller, K.-R. & Clementi, C. Machine Learning for Molecular Simulation. *Annu. Rev. Phys. Chem.* **71**, 361–390 (2020).

22. Park, E. & Izadi, S. Molecular Surface Descriptors to Predict Antibody Developability. *bioRxiv* 2023.07.18.549448 (2023) doi:10.1101/2023.07.18.549448.

23. Licari, G. *et al.* Embedding Dynamics in Intrinsic Physicochemical Profiles of Market-Stage Antibody-Based Biotherapeutics. *Mol. Pharm.* (2022) doi:10.1021/acs.molpharmaceut.2c00838.

24. Raybould, M. I. J., Turnbull, O. M., Suter, A., Guloglu, B. & Deane, C. M. Contextualising the developability risk of antibodies with lambda light chains using enhanced therapeutic antibody profiling. *Commun Biol* **7**, 62 (2024).

25. Yan, R., Xu, D., Yang, J., Walker, S. & Zhang, Y. A comparative assessment and analysis of 20 representative sequence alignment methods for protein structure prediction. *Sci. Rep.* **3**, 2619 (2013).

26. Kuhlman, B. & Bradley, P. Advances in protein structure prediction and design. *Nat. Rev. Mol. Cell Biol.* **20**, 681–697 (2019).

27. Blow, D. *Outline of Crystallography for Biologists*. (OUP Oxford, 2002).

28. Andrec, M. *et al.* A large data set comparison of protein structures determined by crystallography and NMR: statistical test for structural differences and the effect of crystal packing. *Proteins* **69**, 449–465 (2007).

29. Improta, R., Vitagliano, L. & Esposito, L. The determinants of bond angle variability in protein/peptide backbones: A comprehensive statistical/quantum mechanics analysis. *Proteins* **83**, 1973–1986 (2015).

30. Abanades, B., Georges, G., Bujotzek, A. & Deane, C. M. ABlooper: Fast accurate antibody CDR loop structure prediction with accuracy estimation. *Bioinformatics* (2022) doi:10.1093/bioinformatics/btac016.

31. Giovanoudi, E. & Rafailidis, D. Multi-Task Learning with Loop Specific Attention for CDR Structure Prediction. *arXiv [cs.LG]* (2023).