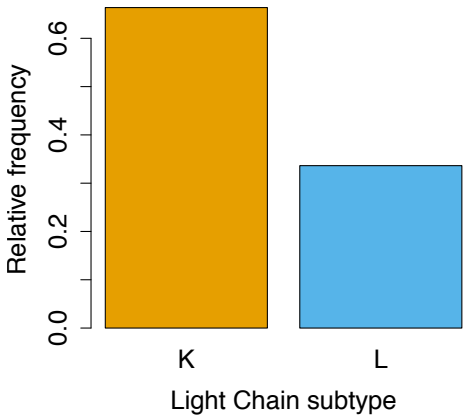
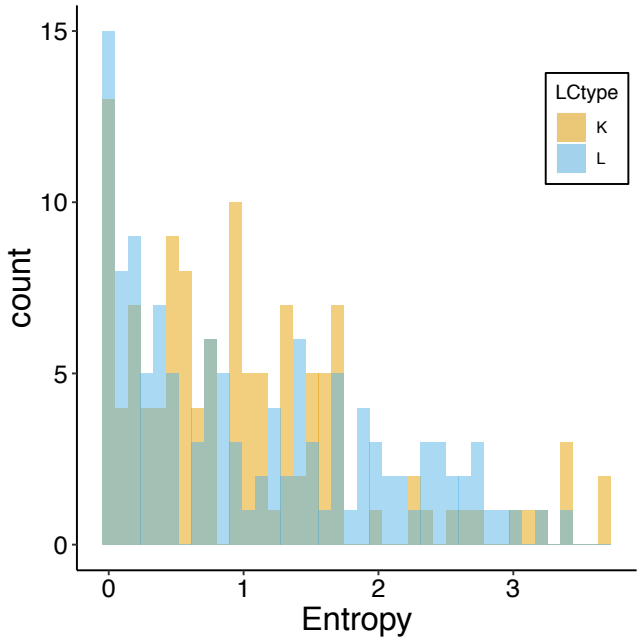
Supplementary Material

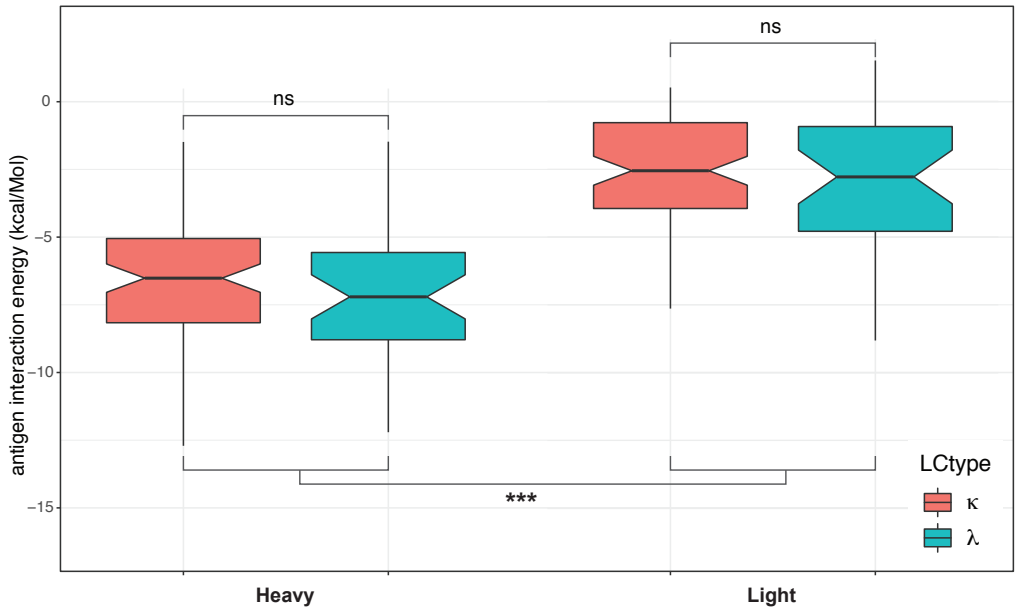
Supplementary Figures



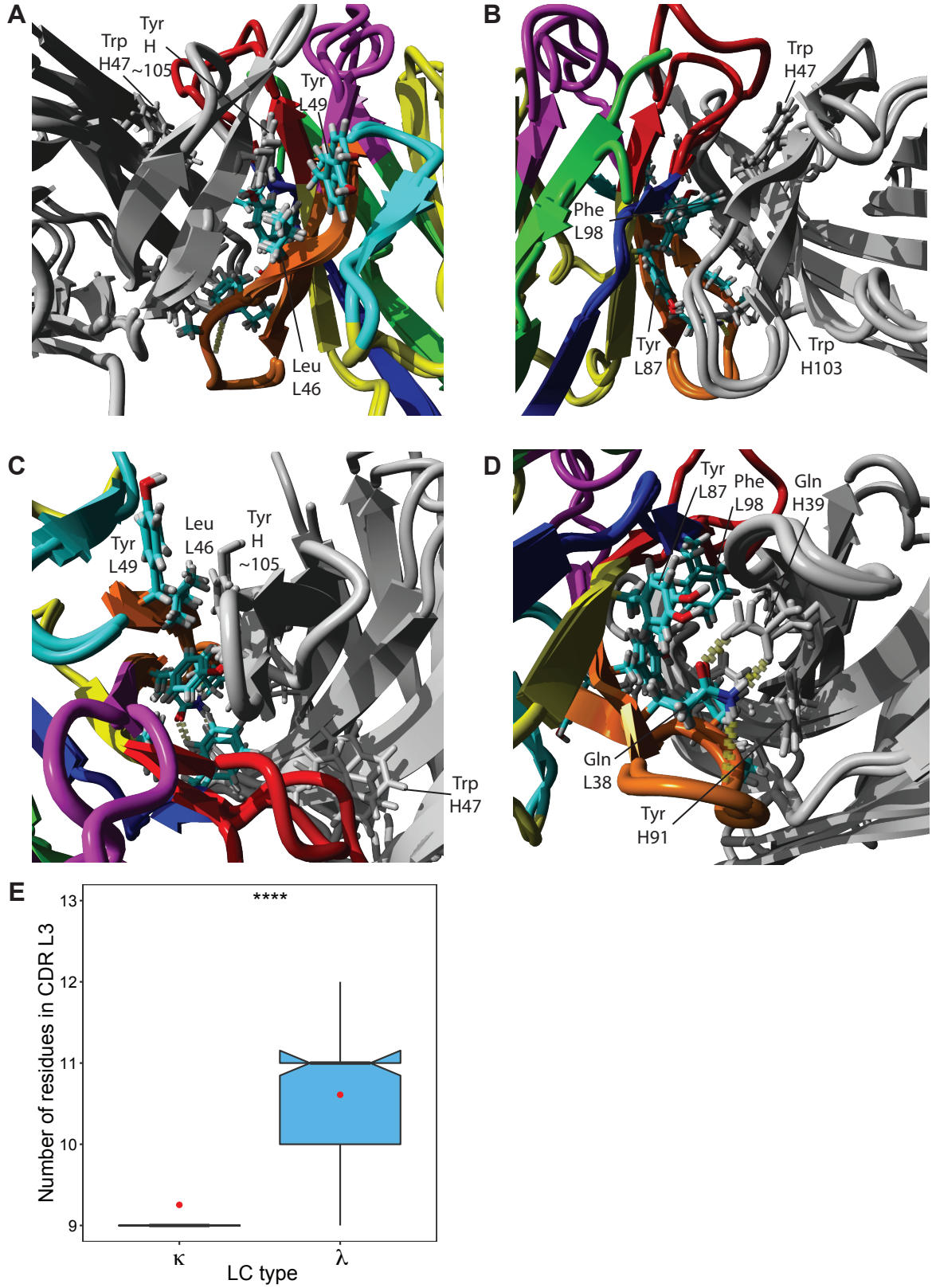
Supplementary Figure 1 – Relative frequency of κ (orange) and λ (blue) in the dataset analyzed.



Supplementary Figure 2 – The histogram displays the distribution of Shannon Entropy over the $V_L\lambda$ and $V_L\kappa$ multiple sequence alignments, calculated per sequence position. The Shannon entropy gives a measure of how the sequence diversity of each position. The Shannon Entropy was calculated per sequence position in the multiple sequence alignments of both, $V_L\kappa$ (orange) and $V_L\lambda$ (blue). The histogram was made by splitting the entropy over 40 bins.

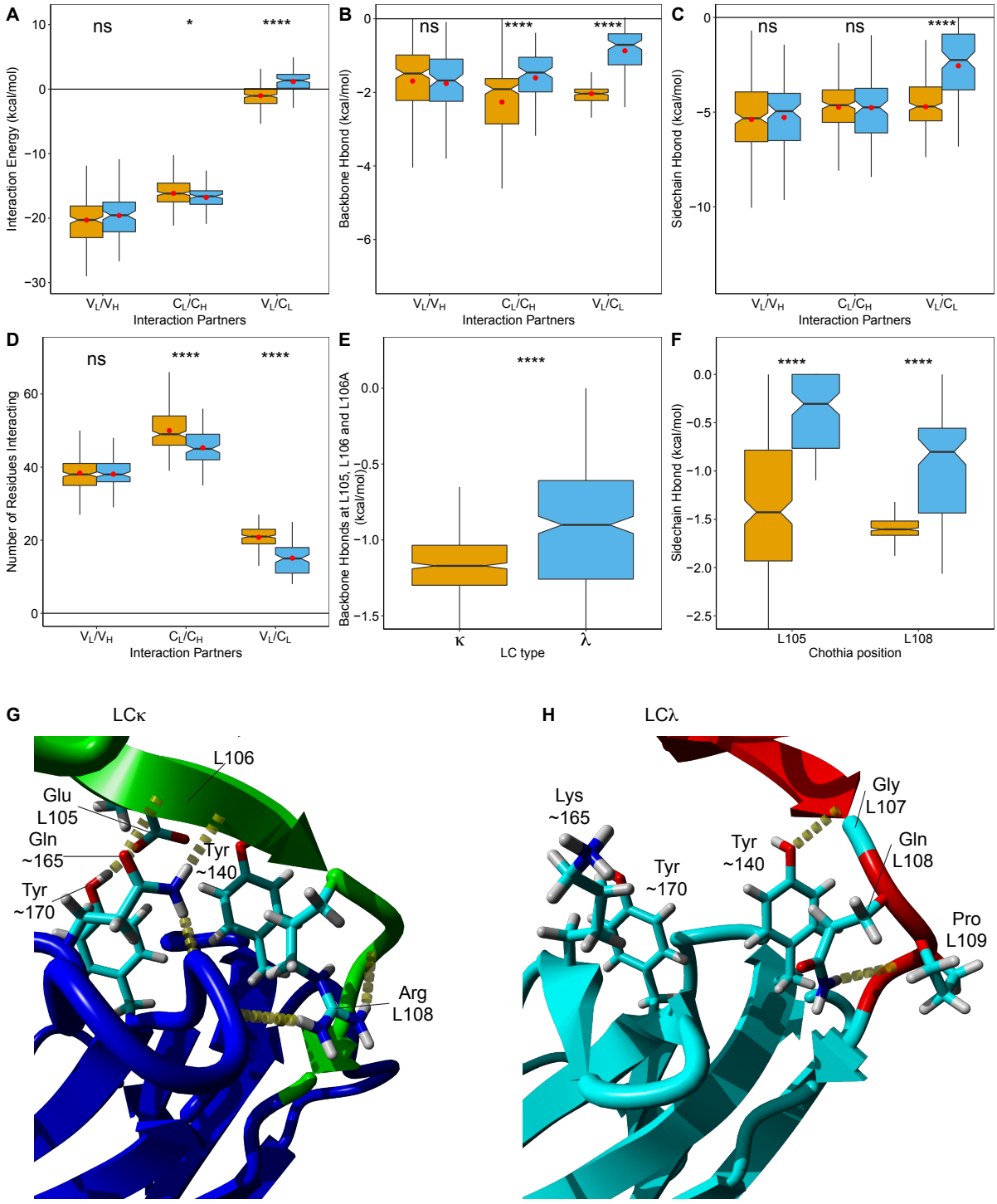


Supplementary Figure 3 – The individual interaction energy with the antigen was calculated separately for both isotypes (λ and κ) of the HC and LC. The particular energy contributions were computed by FoldX.



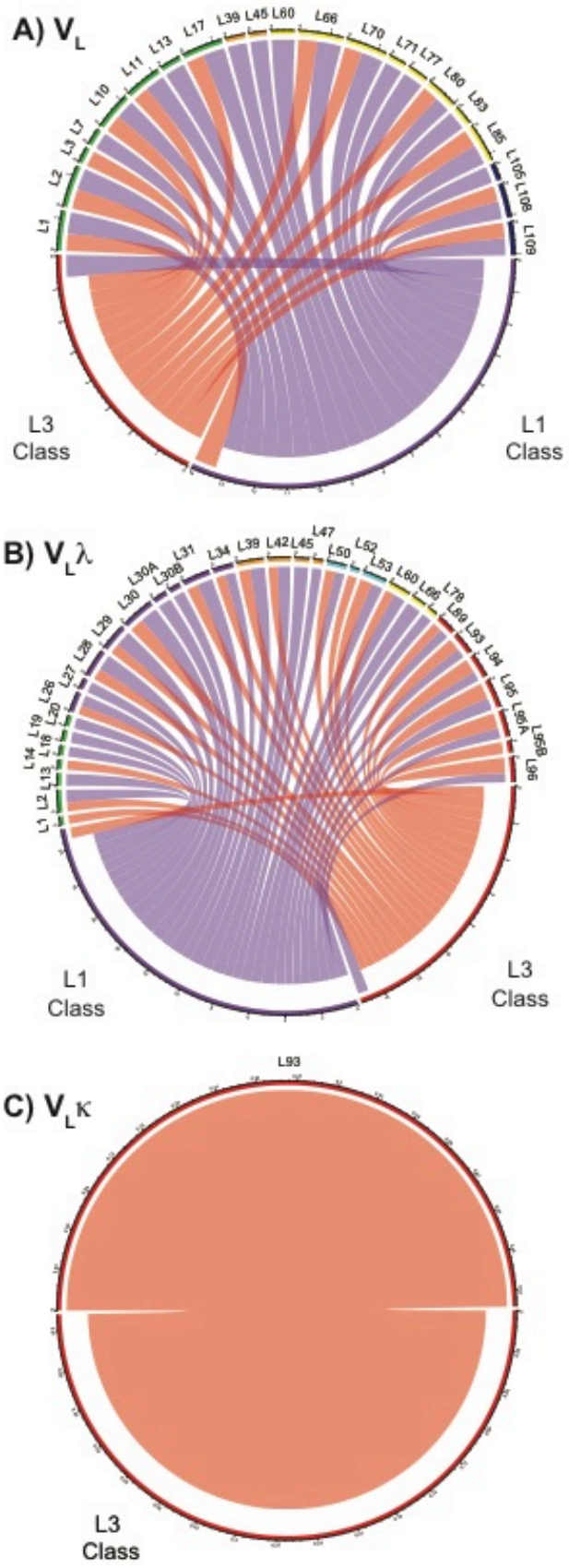
Supplementary Figure 4 – Conserved network of interacting amino acids within the V_L-V_H interface. (A-D) Representative PDB structures of $V_L\lambda$ (3ujj [1]) and $V_L\kappa$ (5ifa [2]) illustrate their shared structural environment that forms the V_L-V_H interface. The CDRs (L1, magenta; L2,

cyan; L3, red) and LFRs (LFR1, green; LFR2, orange; LFR3, yellow; LFR4, blue) of the V_L are color-coded and the V_H domain is shown in gray. Front **(A)** and back **(B)** view on CDRs that are oriented to the top. Top **(C)** and bottom **(D)** view at the V_L-V_H interface. **(E)** The length of CDR L3 differs significantly between the isotypes, since the loop of κ (orange) consistently comprises nine residues and that of λ (blue) exhibits a varying length of eleven residues on average.



Supplementary Figure 5 – Statistical analysis of the thermodynamic energies of inter-domain interactions within Fabs comprising κ or λ. (A-F) Boxplots incorporate the dataset's entire value distribution (whiskers), 25th (lower limit of the box) and 75th percentile (upper limit), mean (red dot) as well as median (central line) with a 95% confidence interval (notches). Statistical significances were obtained by a Wilcoxon rank sum test and are indicated by ns (P > 0.05), * (P ≤ 0.05), ** (P ≤ 0.01), *** (P ≤ 0.001) and **** (P ≤ 0.0001). Comparisons are for both V_Lκ (orange) and V_Lλ (blue). **(A)** Total interaction energy between Fab domains. The interaction between V_L and C_L is significantly more thermodynamically stable in κ than in λ. **(B)** The contribution of backbone H-bonding to the inter-domain interaction energy. The interaction of V_L-C_L and C_L-C_H is significantly more stabilized by

backbone H-bonds in κ if compared to λ. **(C)** The contribution of sidechain H-bonding to the interaction energy of distinct Fab domains. The sidechain H-bonding within the V_L-C_L interface is significantly more pronounced in κ. **(D)** Average number of residues that interact within the domain-domain interface. In κ, significantly more residues are involved in the V_L-C_L and C_L-C_H interaction. **(E)** The residual contribution of backbone H-bonding within the hinge region between V_L-C_L (L105, L106 and L106A). These residues consistently form more backbone H-bonds in κ if compared to λ. **(F)** The residual contribution of sidechain H-bonding within the hinge region. Both, L105 and L108, consistently form more sidechain H-bonds in κ than in λ. **(G, H)** Representative structures of LCκ (**G**, PDBid: 5ifa [2]) and LCλ (**H**, PDBid: 3ujj [1]) illustrate that the V_Lκ (green) - C_Lκ (blue) interface is considerably more hydrogen bonded (yellow bars) than that of V_Lλ (red) - C_Lλ (cyan).



Supplementary Figure 6 – Chord diagrams showing the mutual information between canonical structures and framework residues. (A, B, C) Mutual information between Deane's canonical structures and framework residues of V_L (**A**), $V_L\lambda$ (**B**) and $V_L\kappa$ (**C**) were calculated, and values > 0.5 are shown. The colour of the chord codes for the CDR (salmon: CDR L3, purple: CDR L1), and the width of the chords correlates with the extent of mutual information. The colouring scheme of Chothia positions resembles that of **Figure 1, 2 and 3**.

Supplementary Tables

Supplementary Table 1 – The occurrence of distinct Chothia positions of CDR L3 in κ (221 structures) and λ (112 structures).

Chothia #	κ occurrence	λ occurrence
L89	221	112
L90	221	112
L91	221	112
L92	211	110
L93	210	109
L94	209	109
L95	195	109
L95A	35	94
L95B	15	63
L95C	1	20
L95D	0	2
L96	221	112
L97	221	112

Supplementary Table 2 – Variables used in classification by RandomForest.

Variable	Definition	Unit
total.energy	The predicted overall stability	kcal/mol
Backbone.Hbond	The contribution of backbone H-bonds	kcal/mol
Sidechain.Hbond	The contribution of sidechain H-bonds	kcal/mol
Van.der.Waals	The contribution of VanderWaals forces	kcal/mol
Electrostatics	The contribution of electrostatic interactions	kcal/mol
Solvation.Polar	The penalty for burying polar residues	kcal/mol
Solvation.Hydrophobic	The contribution of hydrophobic groups	kcal/mol
Van.der.Waals.clashes	The penalty for VanderWaals' clashes (interresidue)	kcal/mol
entropy.sidechain	The entropy cost of fixing the sidechain	kcal/mol
entropy.mainchain	The entropy cost of fixing the mainchain/backbone	kcal/mol
cis_bond	The penalty for having a cis peptide bond	kcal/mol
torsional.clash	The penalty for VanderWaals' torsional clashes (intraresidue)	kcal/mol
backbone.clash	The penalty for VanderWaals' backbone-backbone clashes	kcal/mol
helix.dipole	The contribution of the helix dipole (electrostatic)	kcal/mol
disulfide	The contribution of disulfide bonds	kcal/mol
electrostatic.kon	Th electrostatic interaction between molecules in the precomplex	kcal/mol
energy.Ionisation	The contribution of ionisation energy	kcal/mol
sidechain.burial	Burial of the sidechain	fraction
mainchain.burial	Burial of the backbone	fraction
sidechain.Occ	Occupancy of the sidechain	fraction
mainchain.Occ	Occupancy of the backbone	fraction
GAP	Amount of gaps in structure	Count
FRcount	Amount of residues in frame region	Count
CDRcount	Amount of residues in CDRs	Count
VARcount	Amount of residues in variable region	Count
CONcount	Amount of residues in constant region	Count
rescount	Amount of residues	Count
Netcharge	Net charge of light chain	Count
CDRcountL1	Amount of residues in CDR L1	Count
CDRcountL2	Amount of residues in CDR L2	Count
CDRcountL3	Amount of residues in CDR L3	Count
CDRcountH1	Amount of residues in CDR H1	Count
CDRcountH2	Amount of residues in CDR H2	Count
CDRcountH3	Amount of residues in CDR H3	Count
CDRcountLFR1	Amount of residues in frame region LFR1	Count
CDRcountLFR2	Amount of residues in frame region LFR2	Count
CDRcountLFR3	Amount of residues in frame region LFR3	Count
CDRcountLFR4	Amount of residues in frame region LFR4	Count
CDRcountHFR1	Amount of residues in frame region HFR1	Count
CDRcountHFR2	Amount of residues in frame region HFR2	Count
CDRcountHFR3	Amount of residues in frame region HFR3	Count

CDRcountHFR4	Amount of residues in frame region HFR4	Count
Charge	Amount of charged residues	Count
MASS	Mass of light chain	kDa
arom	Amount of aromatic residues	Count

Supplementary References

[1] Gorny MK, Sampson J, Li H, Jiang X, Totrov M, Wang X-H, et al. Human anti-V3 HIV-1 monoclonal antibodies encoded by the VH5-51/VL lambda genes define a conserved antigenic structure. PLoS ONE2011. p. e27780.
[2] Jardine JG, Kulp DW, Havenar-Daughton C, Sarkar A, Briney B, Sok D, et al. HIV-1 broadly neutralizing antibody precursor B cells revealed by germline-targeting immunogen. Science2016. p. 1458-63.