Supplemental Methods and Figures

1. Rosetta Commands

For the benchmarking set, full commands and CDR Instruction File contents are given below. Note that some options for some experimental groups are the default and do not need to be specified, but are given in the file explicitly anyway. To prepare the structures for design, they are relaxed with the Rosetta force field using the FastRelax protocol with these commands:

Command

```
relax.mpi.linuxgccrelease -l PDBLIST.txt -nstruct 10
@pareto_optimal_flags.txt
```

*Contents of **pareto_optimal_flags.txt** file*

```
-no_optH false
-flip_HNQ
-use_input_sc
-constrain_relax_to_start_coords
-relax:ramp_constraints false
-relax:coord_constrain_sidechains
-ignore_unrecognized_res
-ignore_zero_occupancy false
-pdb_comments
-ex1
-ex2
-out:pdb_gz
-other_pose_to_scorefile
-scorefile_format json
-jd2:delete_old_poses
```

The benchmark antibodies are designed with the following commands (with variations for each protocol):

Command:

```
Antibody_designer.mpi.linuxgccrelease @common_flags.txt
@experimental_flags.txt —l PDBLIST.txt —cdr_instructions
instruction_file.txt
```

*Contents of **common_flags.txt** file:*

```
-output_ab_scheme AHo_Scheme
-design_protocol even_cluster_mc
-paper_ab_db True
-nstruct 100
-outer_cycle_rounds 100
-random_start True
-add_graft_log_to_pdb
-ignore_zero_occupancy false
-ignore_unrecognized_res
-pdb_comments
-ex1
-ex2
-use_input_sc
-out:pdb_gz
-other_pose_to_scorefile
-scorefile_format json
```

```
      -flip_HNQ
      -delete_old_poses
```

*Contents of **experimental_flags.txt** file:*

```
#Lambda/Min
      -light_chain lambda
      -remove_antigen False

#Lambda/Min/No antigen
      -light_chain lambda
      -remove_antigen True

#Lambda/Relax
      -light_chain lambda
      -remove_antigen False

#Lambda/Relax/No antigen
      -light_chain lambda
      -remove_antigen True

#Kappa/Min
      -light_chain kappa
      -remove_antigen False

#Kappa/Min/No antigen
      -light_chain kappa
      -remove_antigen True

#Kappa/Relax
      -light_chain kappa
      -remove_antigen False

#Kappa/Relax/No antigen
      -light_chain kappa
      -remove_antigen True
```

CDR Instruction File contents:

```
#RAbD Defaults:

      L1 MinProtocol Min_Neighbors L2 L3
      L2 MinProtocol Min_Neighbors L1
      L3 MinProtocol Min_Neighbors L1 H3
      H1 MinProtocol Min_Neighbors H2 H3
      H2 MinProtocol Min_Neighbors H1
      H3 MinProtocol Min_Neighbors L1 L3
      ALL MinProtocol MinType min

#Min
      L1 GraftDesign Allow
      L2 GraftDesign Allow
      L3 GraftDesign Allow
      H1 GraftDesign Allow
      H2 GraftDesign Allow
      ALL SeqDesign Allow
```

```
#Relax
      L1 GraftDesign Allow
      L2 GraftDesign Allow
      L3 GraftDesign Allow
      H1 GraftDesign Allow
      H2 GraftDesign Allow
      ALL SeqDesign Allow
      ALL MinProtocol MinType relax
```

For the redesign of the 2j88 antibody, the Rosetta command was:

antibody_designer.mpi.linuxgccrelease @common_flags.txt @experimental_flags.txt –cdr_instructions cdr_instructions.txt –s 2j88_pareto_optimal.pdb –nstruct 1000

*Contents of **common_flags.txt** file:*

```
-output_ab_scheme AHo_Scheme
-add_graft_log_to_pdb
-ignore_zero_occupancy false
-ignore_unrecognized_res
-pdb_comments
-ex1
-ex2
-use_input_sc
-out:pdb_gz
-other_pose_to_scorefile
-scorefile_format json
-flip_HNQ
-delete_old_poses
```

*Contents of **experimental_flags.txt**: Docking Off*

```
-outer_cycle_rounds 100
-s input_pdbs/pareto_2j88_renum_0002.pdb.gz
-run_relax
-random_start
-design_protocol even_cluster_mc
-light_chain kappa
```

*Contents of **experimental_flags.txt**: Docking On*

```
-outer_cycle_rounds 100
-run_relax
-random_start
-design_protocol even_cluster_mc
-light_chain kappa
-use_epitope_constraints
-do_dock
-inner_cycle_rounds 2
```

Contents of **cdr_instructions.txt**; H2 Design

```
H2 ALLOW
ALL GraftDesign mintype relax
H2 MinProtocol Min_Neighbors H1 H3
H2 CDRSet Cluster_Cutoffs 10
ALL CDRSET EXCLUDE PDBIDs 2J88
```

Contents of **cdr_instructions.txt;** L1 Design

```
L1 ALLOW
ALL GraftDesign mintype relax
L1 GraftDesign min_neighbors L3 L4
L1 CDRSet Cluster_Cutoffs 10
ALL CDRSET EXCLUDE PDBIDs 2J88
```

Contents of **cdr_instructions.txt**; L1/L4 Design

```
L1 ALLOW
ALL GraftDesign mintype relax
L1 GraftDesign min_neighbors L3 L4
L1 CDRSet Cluster_Cutoffs 10
ALL CDRSET EXCLUDE PDBIDs 2J88
L4 SeqDesign ALLOW
```

Note that many options, such as which CDRs to design, can alternatively be set via simple command-line options. For the redesign of the 2j88 antibody, the Rosetta command was:

General Command:

```
antibody_designer.mpi.linuxgccrelease @common_flags.txt –
cdr_instructions cdr_instructions.txt –s 4JAN_pareto_optimal.pdb
–nstruct 250
```

*Contents of **common_flags.txt** file:*

```
-outer_cycle_rounds 200
-run_relax
-random_start
-design_protocol even_cluster_mc
-light_chain lambda
-add_graft_log_to_pdb
-ignore_zero_occupancy false
-ignore_unrecognized_res
-pdb_comments
-ex1
-ex2
-use_input_sc
-out:pdb_gz
-other_pose_to_scorefile
-scorefile_format json
-flip_HNQ
-delete_old_poses
```

*Contents of **cdr_instructions.txt**; H2 Design, Profiled*

```
H2 ALLOW
H2 MinProtocol Min_Neighbors H1 H3
H2 CDRSet Cluster_Cutoffs 10
ALL CDRSet Exclude PDBIDs 4JAN
```

*Contents of **cdr_instructions.txt**; H2 Design, All*

```
H2 ALLOW
H2 MinProtocol Min_Neighbors H1 H3
ALL CDRSet Exclude PDBIDs 4JAN
```

*Contents of **cdr_instructions.txt**; L1/L3 Design, Profiled*

```
L1 ALLOW
L3 ALLOW
L3 MinProtocol Min_Neighbors L1 H3
```

```
L1 MinProtocol MIN_NEIGHBORS L3 H3
L1 CDRSet Cluster_Cutoffs 10
L3 CDRSet Cluster_Cutoffs 10
ALL CDRSet Exclude PDBIDs 4JAN
```

*Contents of **cdr_instructions.txt**; L1/L3 Design, All*

```
L1 ALLOW
L3 ALLOW
L3 MinProtocol Min_Neighbors L1 H3
L1 MinProtocol MIN_NEIGHBORS L3 H3
ALL CDRSet Exclude PDBIDs 4JAN
```

The AntibodyFeatures reporters (described below) were used in the analysis of the antibodies:

```
rosetta_scripts.macosxrelease –l DECOYS.txt @common_flags –
parser:protocol features_script.xml
```

RosettaScript used to obtain feature databases:

```
<ROSETTASCRIPTS>
   <MOVERS>
      <ReportToDB name='features' database_name='my_db'>
         <feature name='CDRClusterFeatures' numbering_scheme='AHO_Scheme'/>
         <feature name='AntibodyFeatures' numbering_scheme='AHO_Scheme'
cdr_definition='North' interface='LH_A' pack_separated='1' pack_together='1'/>
         <feature name='ResidueFeatures'/>
         <feature name='PdbDataFeatures'/>
         <feature name='ResidueTypesFeatures'/>
         <feature name='ScoreTypeFeatures'/>
         <feature name='StructureScoresFeatures'/>
         <feature name='ResidueSecondaryStructureFeatures'/>
         <feature name='ResidueScoresFeatures'/>
         <feature name='HBondFeatures'/>
         <feature name='SaltBridgeFeatures'/>
         <feature name='ResidueBurialFeatures'/>
      </ReportToDB>
   </MOVERS>
   <PROTOCOLS>
      <Add mover_name='features'/>
   </PROTOCOLS>
</ROSETTASCRIPTS>
```

Once each database was created, extra output in each decoy PDB file (added through the option *–add_graft_log_to_pdb),* was used to calculate recoveries and risk ratios through the use of Jade modules:
https://github.com/SchiefLab/Jade/blob/master/src/RAbD_BM/AnalyzeRecovery.py.

2. Antibody Feature Analysis

Three FeatureReporters were developed as a part of the Rosetta Features Reporter framework[1-3] to aid in the modeling and design of antibody structures. Each of these can be used through the RosettaScript framework on a list of structures. The physical attributes reported are output to a relational database, such as SQLITE3, across multiple tables for further analysis. These databases can easily be converted into CSV files or read by available packages such in R and Python.

The *CDRClusterFeatures* reporter identifies all North/Dunbrack CDR clusters in an antibody to the closest cluster centroid using the same metric described in **PyIgClassify** as well as

information pertaining to the dihedral distances of the CDRs[4]. It is the primary FeatureReporter used in benchmarking length and cluster recovery**.** The database tables output by the *CDRClusterFeatures* are detailed in **Table S6.**

The *InterfaceFeatures Reporter,* detailed in Table S7, analyzes protein-protein and protein-ligand interfaces, outputting a number of different tables and physical data. Much of the analysis is done through an updated version of the Rosetta InterfaceAnalyzer[5,6] where interface differences (such as an estimate of the interface **Δ**G – the enthalpic component of the full binding free energy) are calculated by physically separating the interface components (such as antibody from antigen) and optimizing interface residue side chains - both in the complexed and separated conformations. An interface distance of 6 Å is used as the default interface distance.

Separate tables are output for the overall complex, the individual proteins in the complex, and the interface residues. The main data output by this reporter are the estimated binding energy (**Δ**G) of the complex in Rosetta Energy Units (REU), the change in solvent accessible surface area upon binding (**Δ**SASA) using the Le Grand SASA calculation method[7], the Lawrence and Colman shape complementary of the interface (sc_value)[8], the packing quality (packstat)[9], and the number of unsaturated hydrogen bonds in the complex[6].

Alternative SASA radius sets were added to Rosetta, with the standard, now-defunct radii changing from the default to 'legacy'. We implemented a variety of radius sets found in the literature and used in various structural modeling programs in which they either implicitly or explicitly include hydrogen atoms. Once a particular radius set is used, the SASA machinery will change its consideration of implicitly or explicitly including hydrogen atoms during the calculation depending on the set.

The atomic radius set with implicit hydrogens is the one used by the program Naccess, a popular program used for the calculation of SASA[10]. This set was derived by Chothia in his seminal 1976 paper[11], while explicit hydrogen radius sets include the legacy radii, the Rosetta Lennard-Jones (LJ) radii (which are mostly the same as the LJ radii from the CHARMM molecular dynamics program[12]), and the radii used by the program reduce (a program for the placement of hydrogens onto molecular models and crystal structures)[13], originating from physical data obtained from Bondi, 1964[14] and Gavezzotti, 1983[15]. The reduce radius set is now the default in Rosetta.

The *AntibodyFeatures* Reporter is a type of *InterfaceFeatures* Reporter specific for antibody and antibody-antigen interfaces, while outputting a number of *additional* metrics for antibodies and CDRs. Some of the main metrics include CDR, antibody, and paratope charge, ΔG and ΔSASA, H3 kink statistics, number of contacts, and packing angle statistics[16]. The packing angle is a measure of the relative orientation between the light and heavy antibody chains. It uses four conserved residues of each chain in the framework beta-sheets at the VL and VH interface and principal component analysis to define four centroid points and a dihedral angle for which to quantify the orientation[17].

A full list of the metrics and tables output by the *AntibodyFeatures* Reporter can be found in Table S8. All tables output by the *InterfaceFeatures* Reporter are output by the *AntibodyFeatures* Reporter for specific antibody interfaces specified where A is the antigen: LH-A, L-H, L-A, H-A.

3. Outlier Control

In our original clustering of the antibody CDR structures, an affinity propagation clustering technique was used on a carefully curated dataset of high-resolution structures and few outliers[18]. In order to match new CDR structures with a proper cluster from that original clustering, we use the dihedral angle metric originally used for the affinity propagation, but

measure it against the centroid (representative structure) of all clusters of the same length. The cluster with the lowest dihedral distance is assigned as the cluster for that structure[4].

While this is useful to assign CDRs of known length to a particular cluster, many structures become outliers of the particular cluster and would have formed their own cluster if clustering was repeated. In order to optimize our CDR profiles, constraints, and all other aspects of the design program that rely on a clear cluster without outliers for the up-to-date antibody database that can be downloaded from PyIgClassify, we needed to quantitatively define what would be considered an outlier.

We used both the dihedral distance metric and RMSD of all backbone atoms to help define an outlier. In order to visualize the breadth of each cluster, we generated PyMol sessions of each of the clusters using python and PyRosetta[19] by aligning the CDRs to their cluster center either using all backbone heavy atoms or by aligning only the stem region (three framework residues on either side of the CDR loop).

We also generated plots of dihedral distance versus RMSD and length versus RMSD for both alignment types and for each length and cluster, where high RMSD can be seen even with lower dihedral distance, especially when only the stem was aligned.

We then used these plots and the PyMol visualizations for each CDR cluster to define two outlier definitions – one conservative and one liberal (used as the default). We calculate the RMSD for these definitions through the full CDR alignments as the stem alignment can result in very high RMSD for low dihedral angle distances, attributable to hinge-like motions in the CDR:

(Conservative):

```
if (DihDis >= 40° OR RMSD >= 1.5 Å) then CDR is Outlier
```

(Liberal):

```
if (DihDis >= 40° AND RMSD >= 1.5 Å) then CDR is Outlier
```

Outlier control is handled as an option in the Antibody Design framework, where each set of data used by the framework is first generated with and without outliers and using both the liberal and conservative definition of an outlier. For smaller clusters or H3 (which only cluster well at shorter lengths), outliers may be useful in the design search and an option will switch all aspects of the framework to include outliers for sequence and dihedral constraint statistics as well as graft sets. By default, outliers are left out, but used for H3 since it does not cluster well.

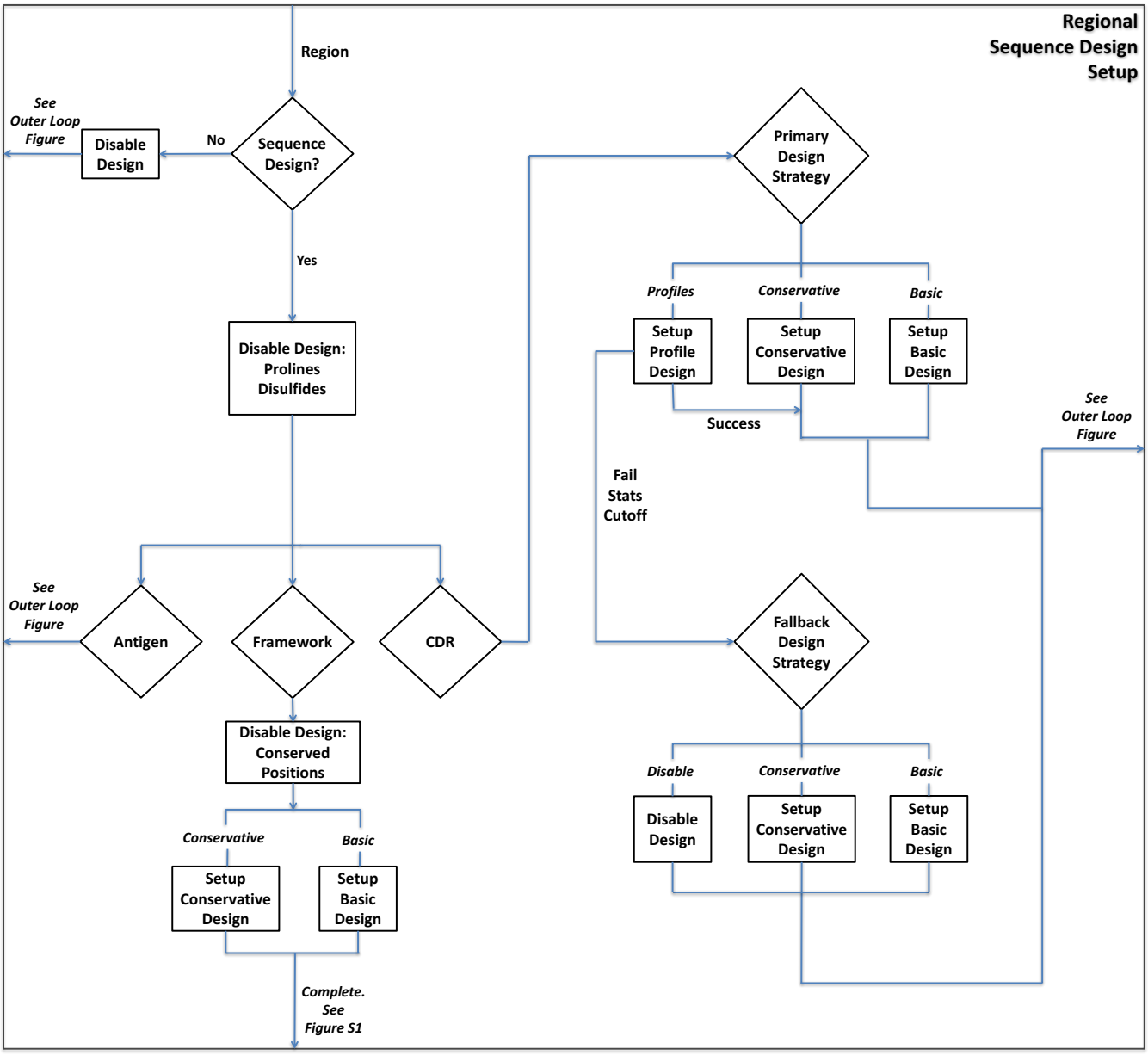


Figure S1: For each region (Antigen, Framework, Each CDR), we first determine if that region is set to design via the program options and CDR Instructions File. If the region is set to design, we disable design for prolines and disulfide-bonded Cysteine residues within the region by default. If it is an antigen region, we use basic design and exit the setup. If it is the framework region, we disable completely conserved positions, such as the tryptophan immediately after the H3 loop. Currently, we either do conservative or basic design for the framework region if the framework is to be designed. By default, the framework region is held fixed in sequence space. Finally, when designing a CDR region, we set up the Primary Sequence Design strategy that is set by the CDR Instructions File. If the Primary Sequence Design strategy is to use the CDR cluster-based profiles and there is scarce data, we use the set Fallback Design Strategy. After the CDR Sequence Design strategy is done, we have completed the setup for sequence design.

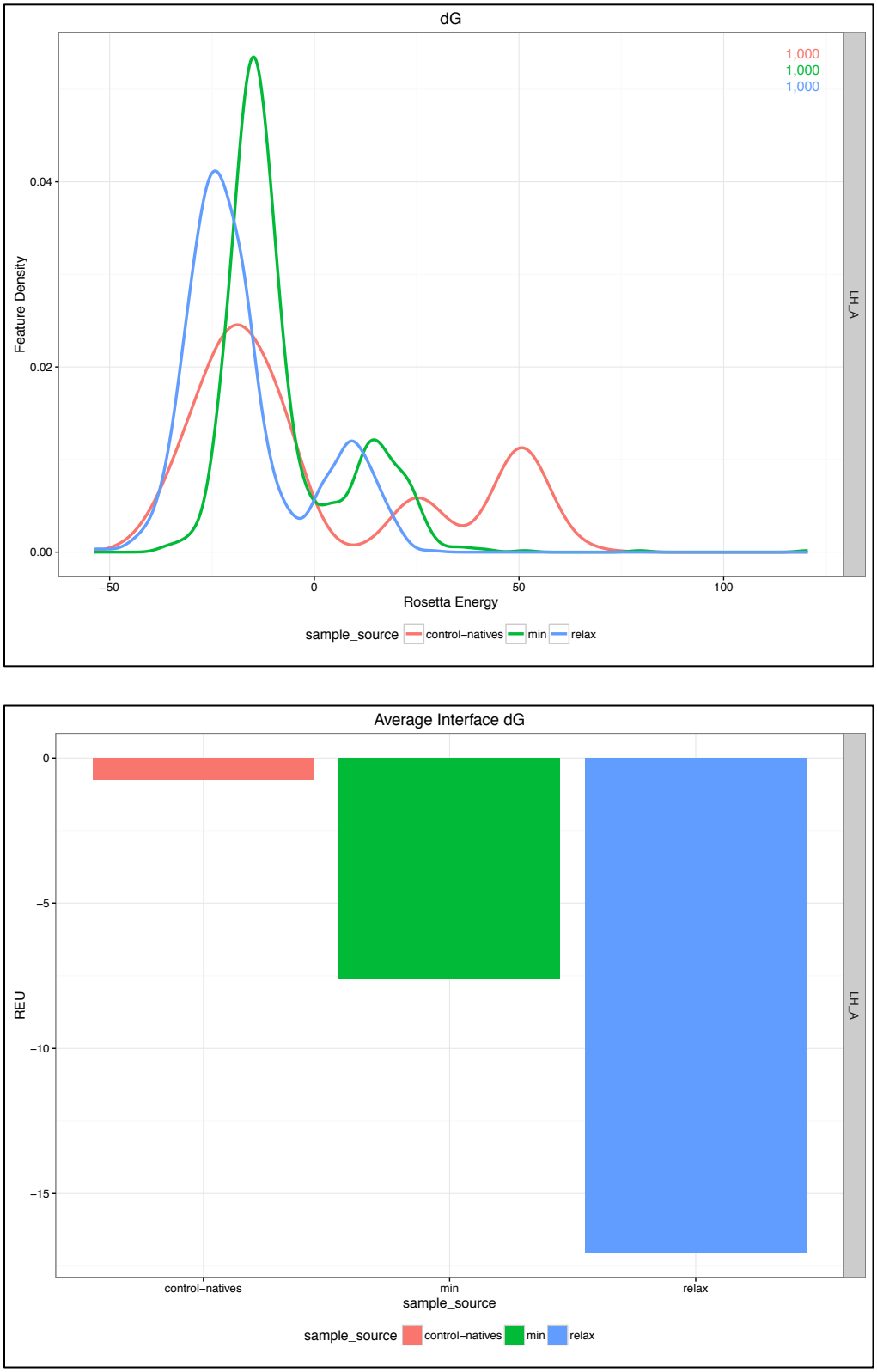


Figure S2: Kernel Density Estimates (KDE) and averages of the Rosetta Interface Energy (dG) of the decoys output for the min and relax benchmark.

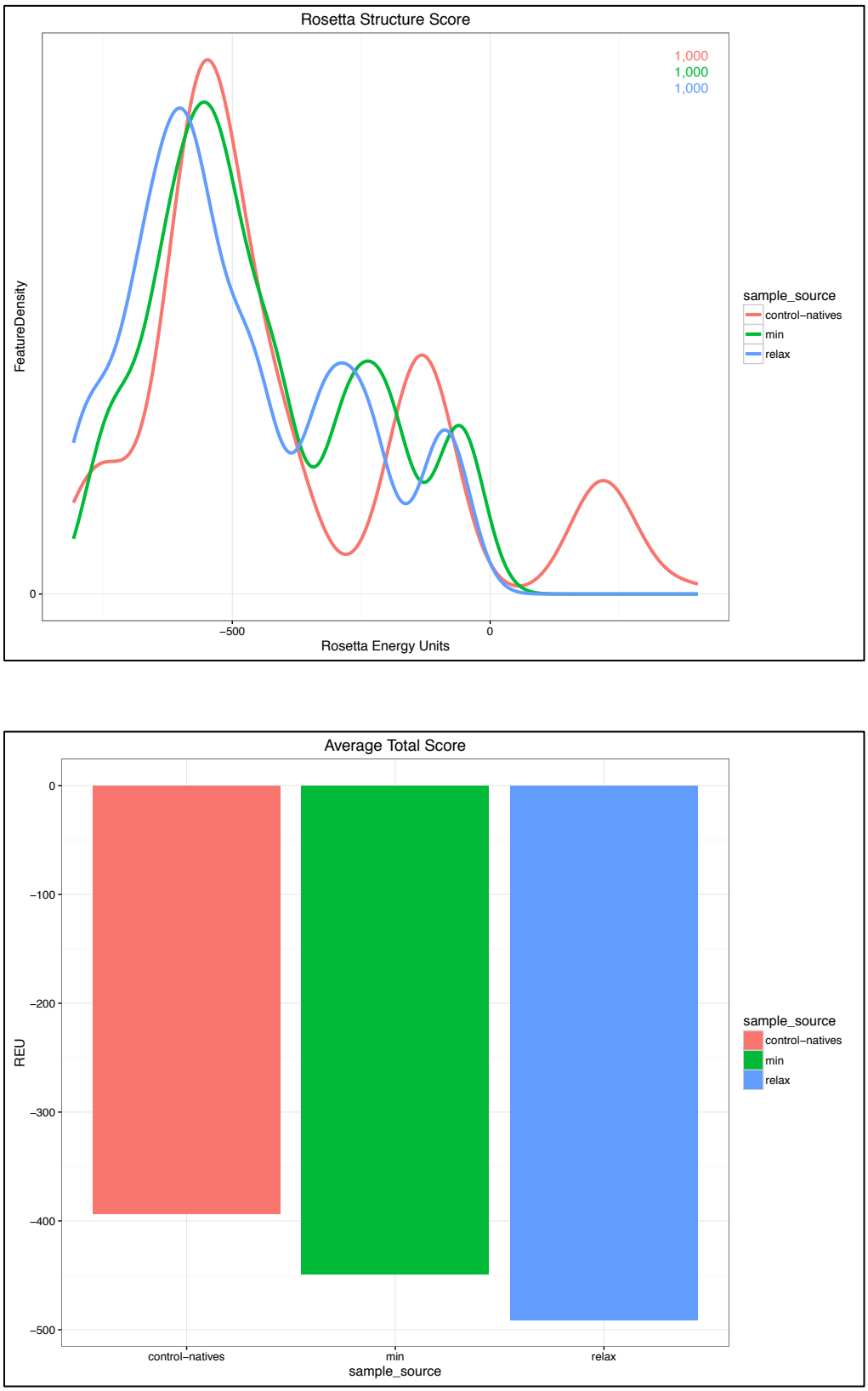


Figure S3 – Kernel Density Estimates and averages of total Rosetta energy (REU) of the min vs relax benchmark decoy set using the Rosetta Talaris2014 Energy function.

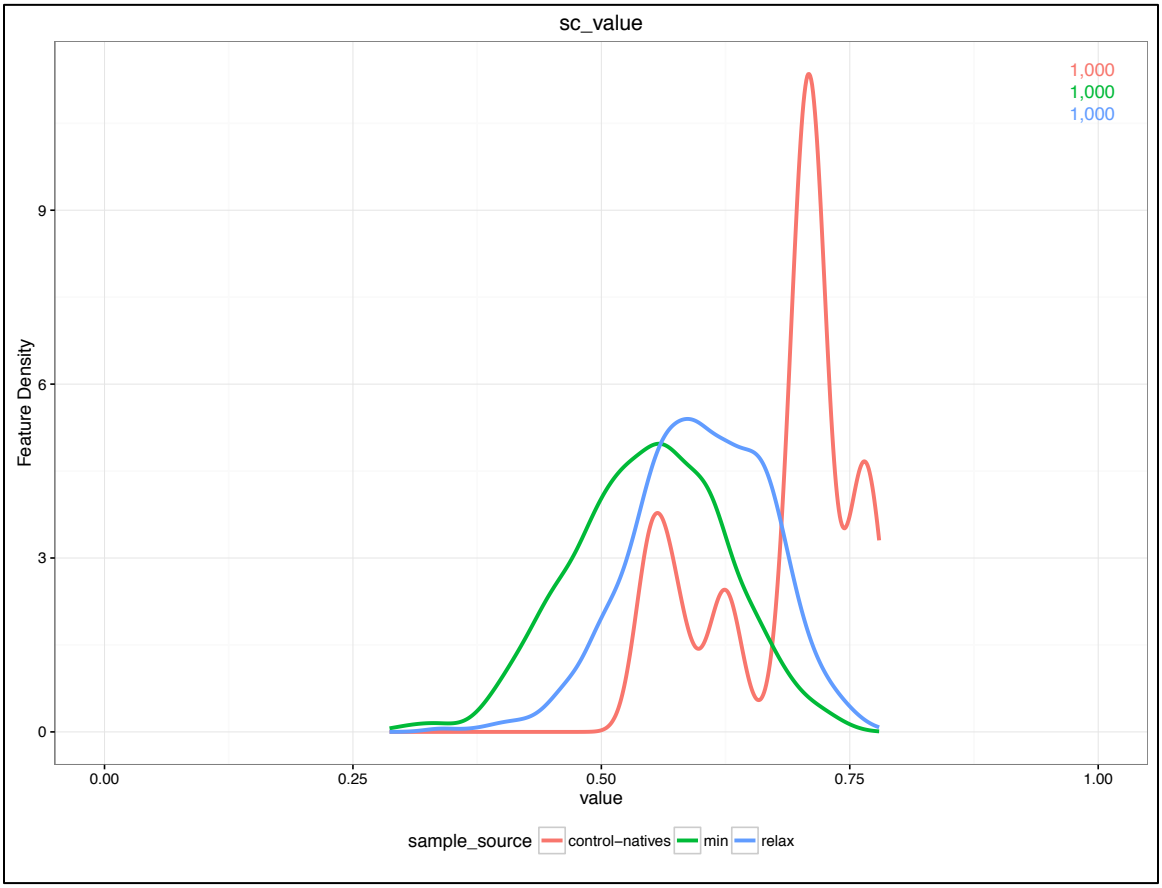


Figure S4: Kernel Density Estimates of the Lawrence and Colman Shape Complementarity value (*sc_value*) of the min and relax benchmark decoys using the Rosetta *AntibodyFeatures* reporter.

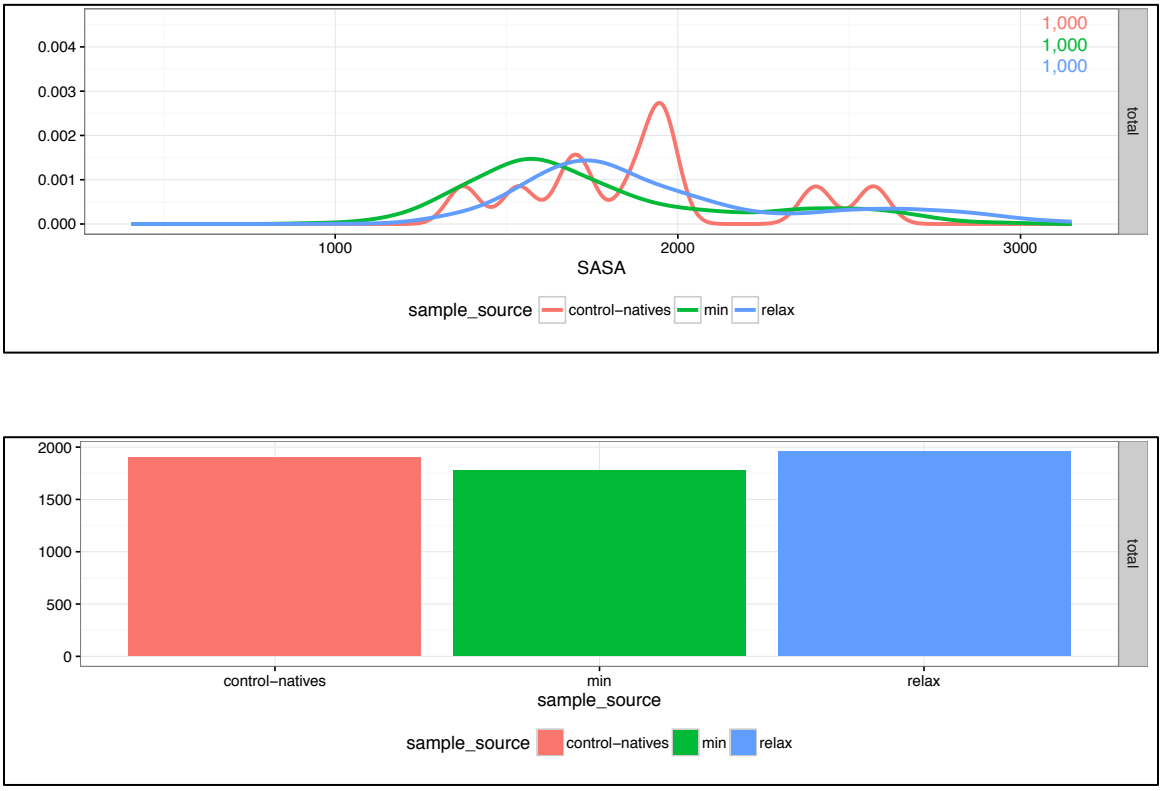


Figure S5: Kernel Density Estimates and averages of the buried Solvent Accessible Surface Area (**Δ**SASA) of the min vs. relax benchmark decoy set at the antibody/antigen interface.

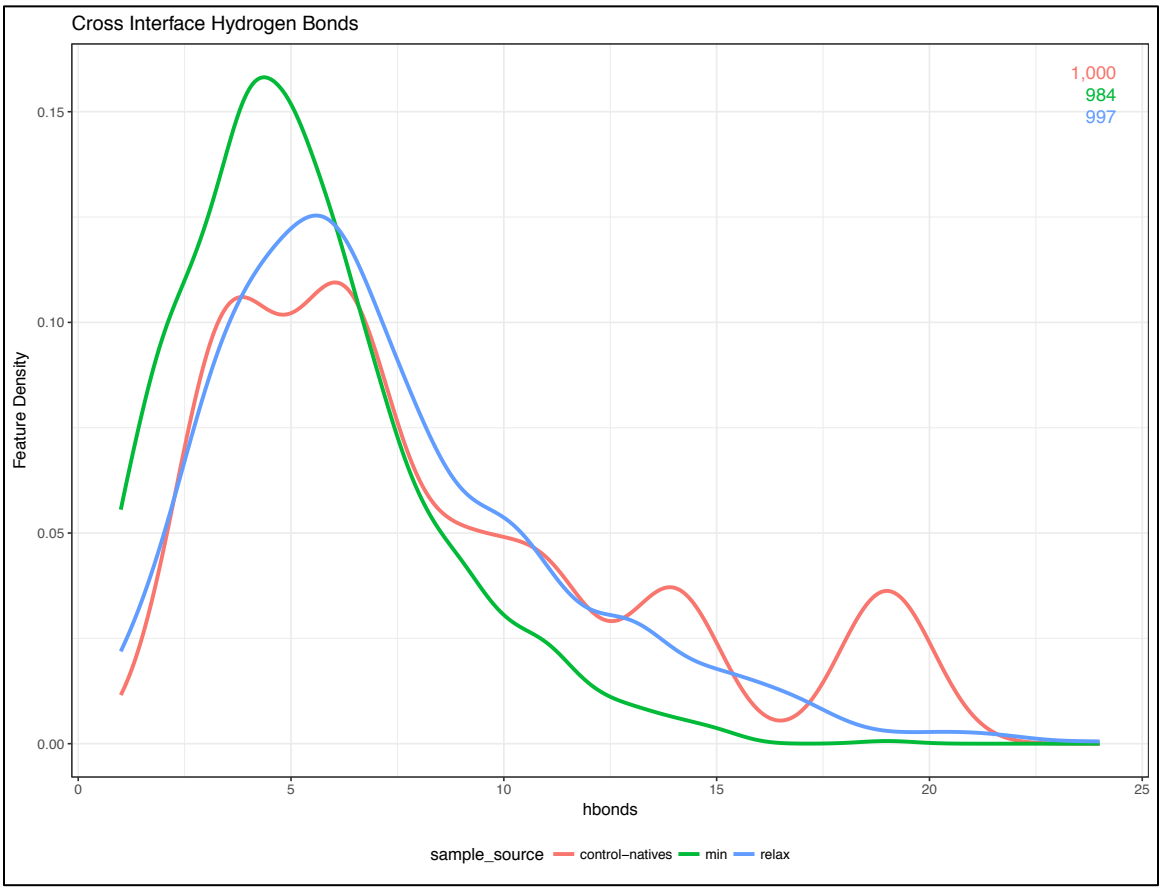


Figure S6: Kernel Density Estimates of the cross-interface hydrogen bonds of the min and relax benchmarks compared to pareto-optimal relaxed natives reported by the *HBondFeatures* reporter.

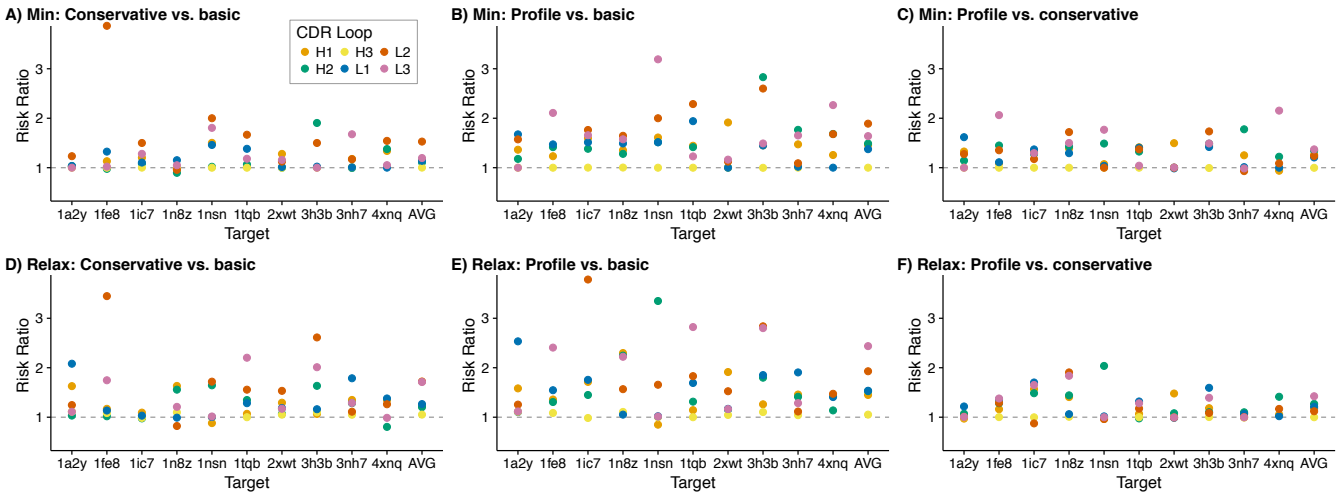


Figure S7: Sequence antigen risk ratios of all CDR residues. Values greater than 1.0 indicate that the native residue types were present in the design simulations more often than they were present in the design simulations with the alternative sequence design strategy

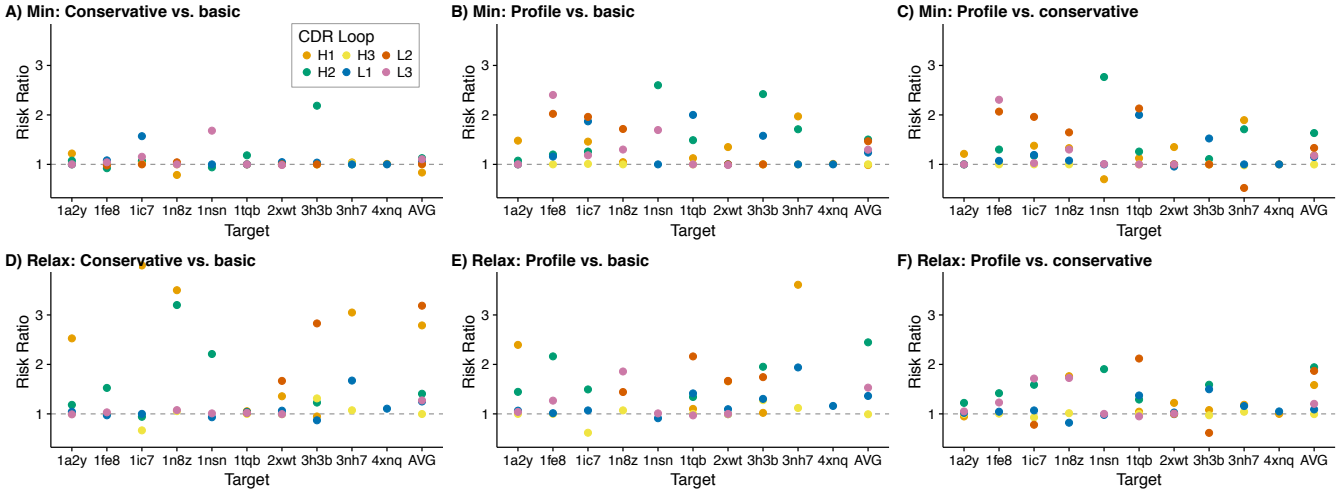


Figure S8: Sequence antigen risk ratios for residues of each antibody in contact with the antigen. Values greater than 1.0 indicate that the native residue types were present in the design simulations more often than they were present in the design simulations with the alternative sequence design strategy

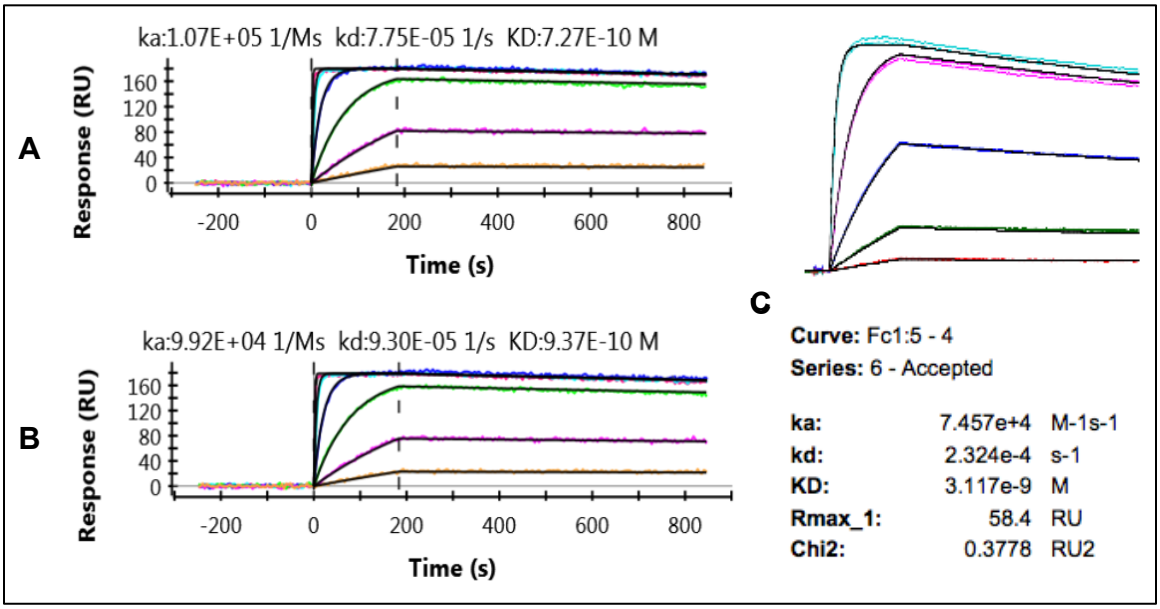


Figure S9: Kinetic sensorgrams of design L1_10 to Bee Hyaluronidase. **(A)** *XPR Repeat 1*; **(B)** *XPR Repeat 2*; **(C)** *Biacore 4000.*

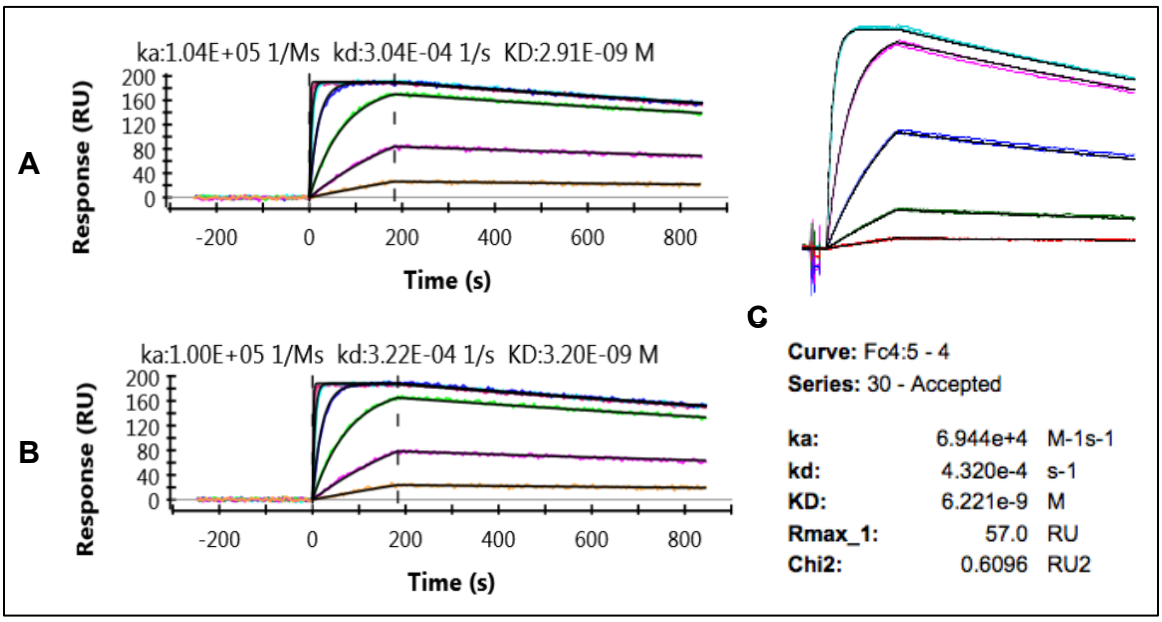


Figure S10: Kinetic sensorgrams of design L1_5 to Bee Hyaluronidase. **(A)** *XPR Repeat 1*; **(B)** *XPR Repeat 2*; **(C)** *Biacore 4000.*

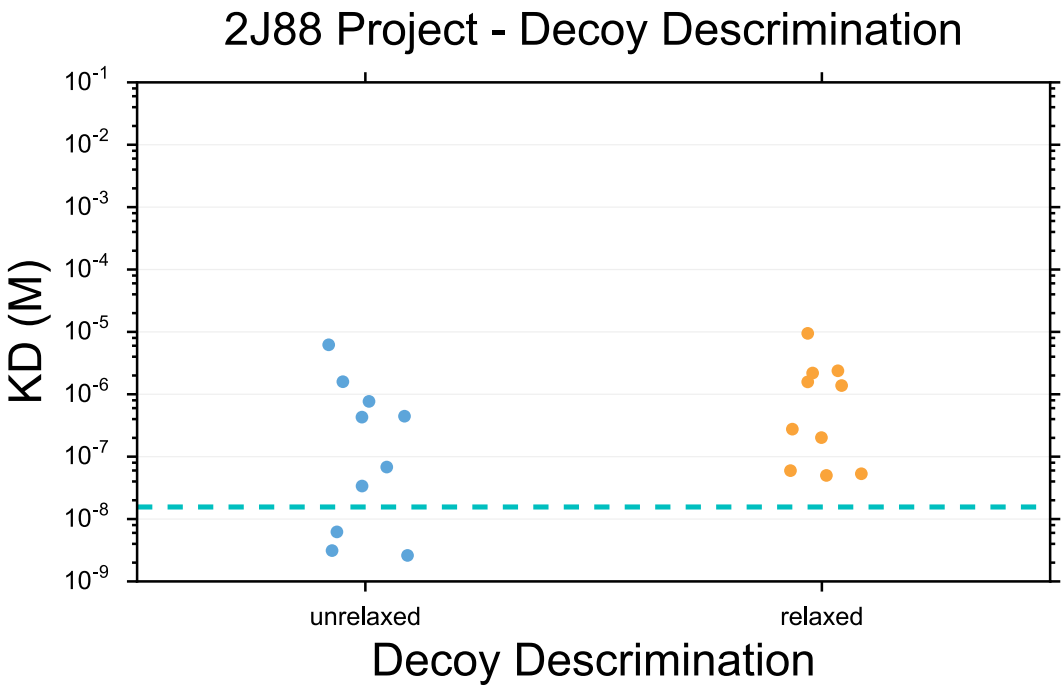
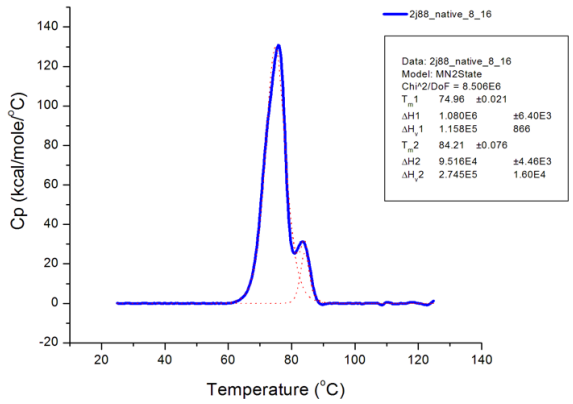
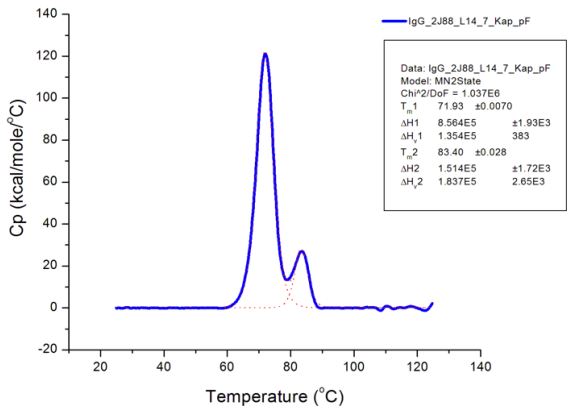


Figure S11: Binding affinity of designs in the Bee Hyaluronidase antibody binding experiment. Dotted line is the K_d of the native antibody from a Biacore 4000. All of the designed antibodies which bound better were chosen through unrelaxed versions of the decoy, output directly from the antibody design program (*unrelaxed*), as opposed to using a full dualspace-relax of all of the designs for decoy discrimination (*relaxed*).

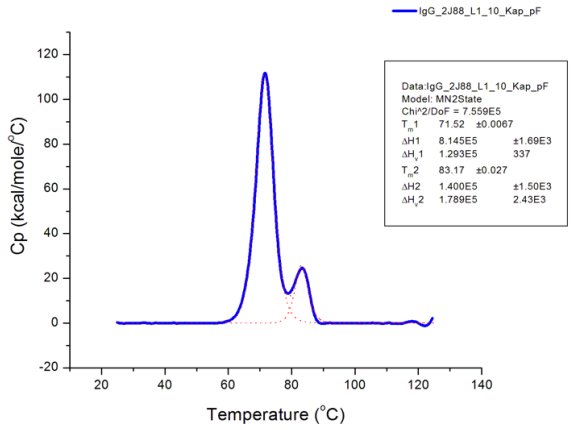
(a) WT 2J88



(b) Design L14_7



(c) Design L1_10



(d) Design L1_5

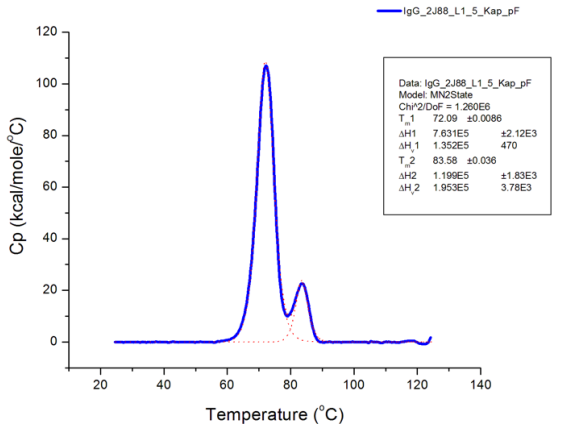


Figure S12: Thermostability measurements of WT 2J88 antibody and designs by Differential Scanning Calorimetry (DSC). **(a)** WT 2J88 thermostability; **(b)** L14_7 design thermostability; **(c)** L1_10 design thermostability; **(d)** L1_5 design thermostability.

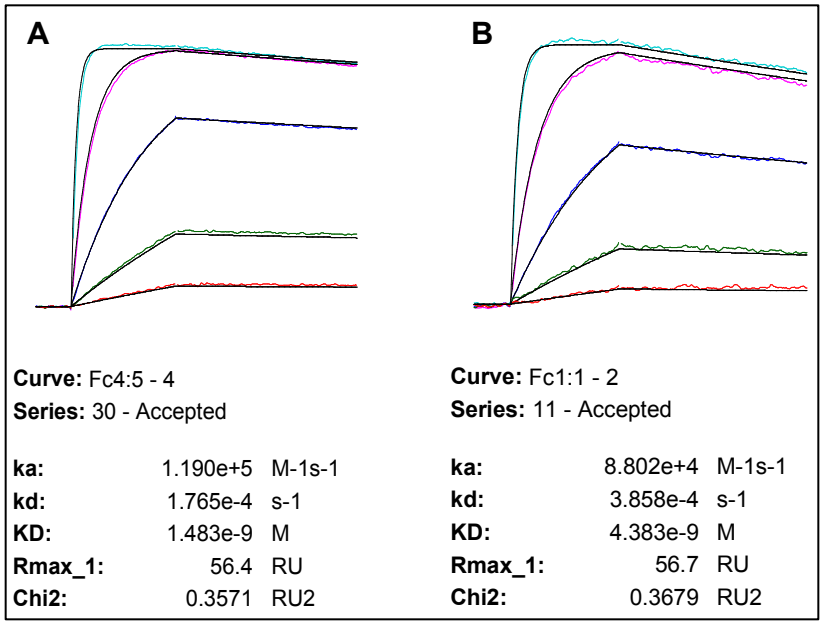


Figure S13: L14_7 Mutants with WT residue at position 38. **(a)** L14_7 **(b)** L14_7 K38Y

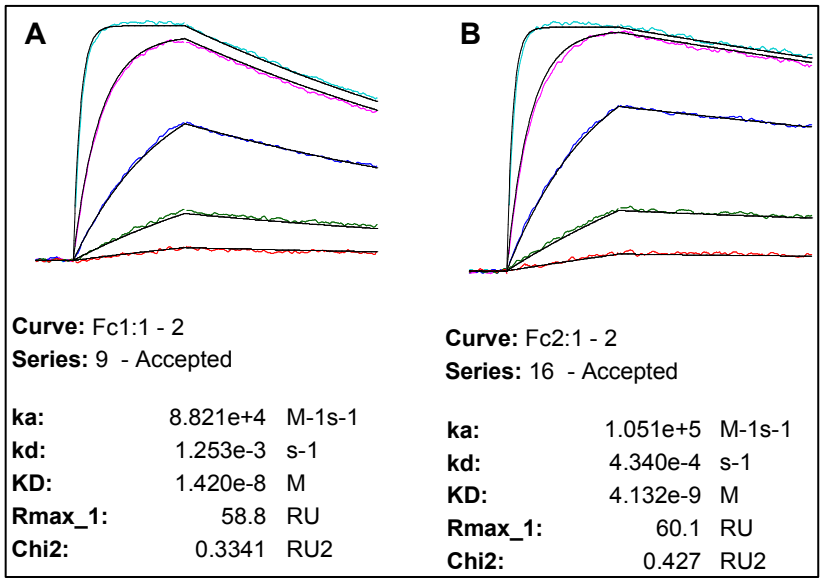


Figure S14: 2J88 WT Mutants with L14_7 residue at position 38. **(a)** 2J88 WT; **(b)** 2J88 Y38K

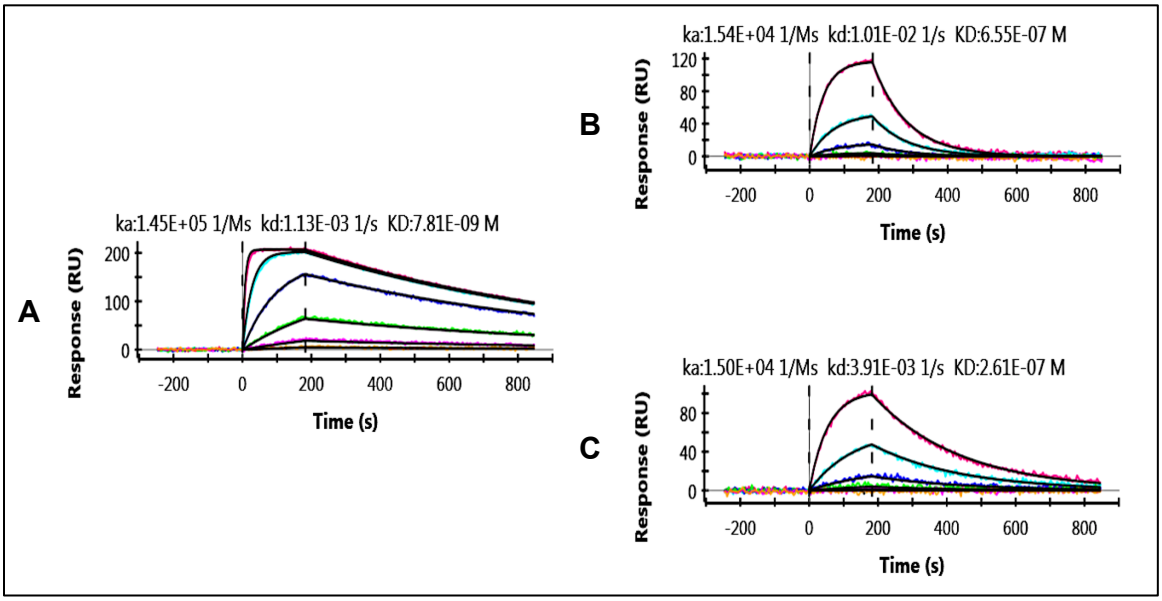
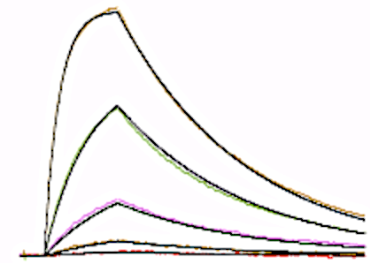
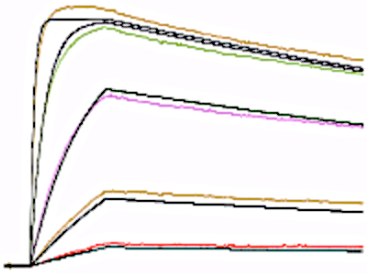
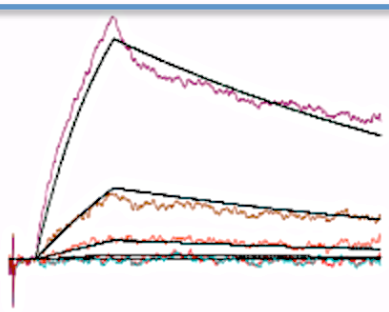
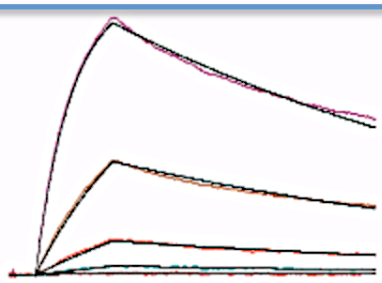
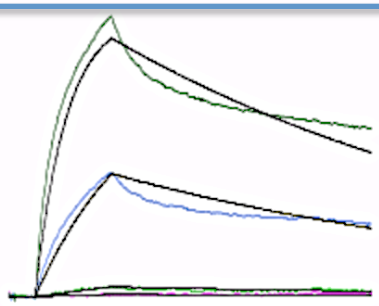
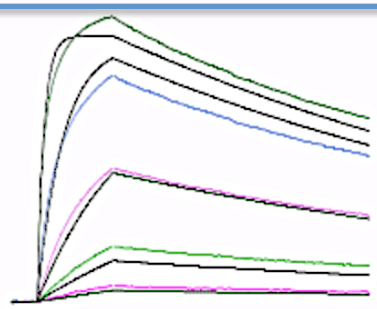
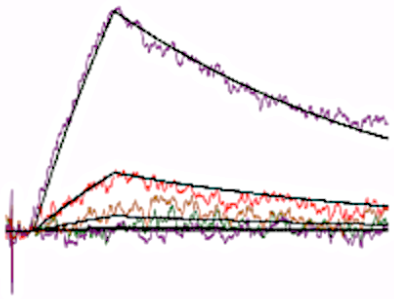
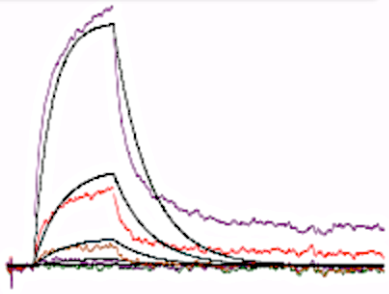


Figure S15: 2J88 WT Mutants with L14_7 residue at position 38. **(a)** 2J88 WT; **(b)** L1_4 Design; **(c)** L1_4 S36V.

	WT	H2_6 Design
HXB2	 **Curve:** Fc1:1 - 2 **Series:** 27 - Accepted **ka:** 6.022e+3 M-1s-1 **kd:** 3.152e-3 s-1 **KD:** 5.234e-7 M **Rmax_1:** 120.0 RU **Chi2:** 0.8653 RU2	 **Curve:** Fc1:1 - 2 **Series:** 25 - Accepted **ka:** 2.806e+4 M-1s-1 **kd:** 3.689e-4 s-1 **KD:** 1.315e-8 M **Rmax_1:** 133.7 RU **Chi2:** 10.57 RU2
GH069	 **Curve:** Fc1:1 - 2 **Series:** 51 - Accepted **ka:** 2.453e+3 M-1s-1 **kd:** 9.410e-4 s-1 **KD:** 3.837e-7 M **Rmax_1:** 37.4 RU **Chi2:** 0.3935 RU2	 **Curve:** Fc1:1 - 2 **Series:** 49 - Accepted **ka:** 6.222e+3 M-1s-1 **kd:** 8.805e-4 s-1 **KD:** 1.415e-7 M **Rmax_1:** 125.4 RU **Chi2:** 0.8368 RU2

	WT	H2_6 Design
YU2	 **Curve:** Fc1:1 - 2 **Series:** 75 - Accepted **ka:** 2.044e+3 M-1s-1 **kd:** 9.563e-4 s-1 **KD:** 4.678e-7 M **Rmax_1:** 98.7 RU **Chi2:** 9.859 RU2	 **Curve:** Fc1:1 - 2 **Series:** 73 - Accepted **ka:** 1.001e+4 M-1s-1 **kd:** 7.244e-4 s-1 **KD:** 7.236e-8 M **Rmax_1:** 178.0 RU **Chi2:** 47.5 RU2
AC10	 **Curve:** Fc1:1 - 2 **Series:** 3 - Accepted **ka:** 1.985e+2 M-1s-1 **kd:** 1.425e-3 s-1 **KD:** 7.178e-6 M **Rmax_1:** 77.4 RU **Chi2:** 0.1236 RU2	 **Curve:** Fc1:1 - 2 **Series:** 1 - Accepted **ka:** 2.976e+3 M-1s-1 **kd:** 1.268e-2 s-1 **KD:** 4.259e-6 M **Rmax_1:** 54.8 RU **Chi2:** 2.913 RU2

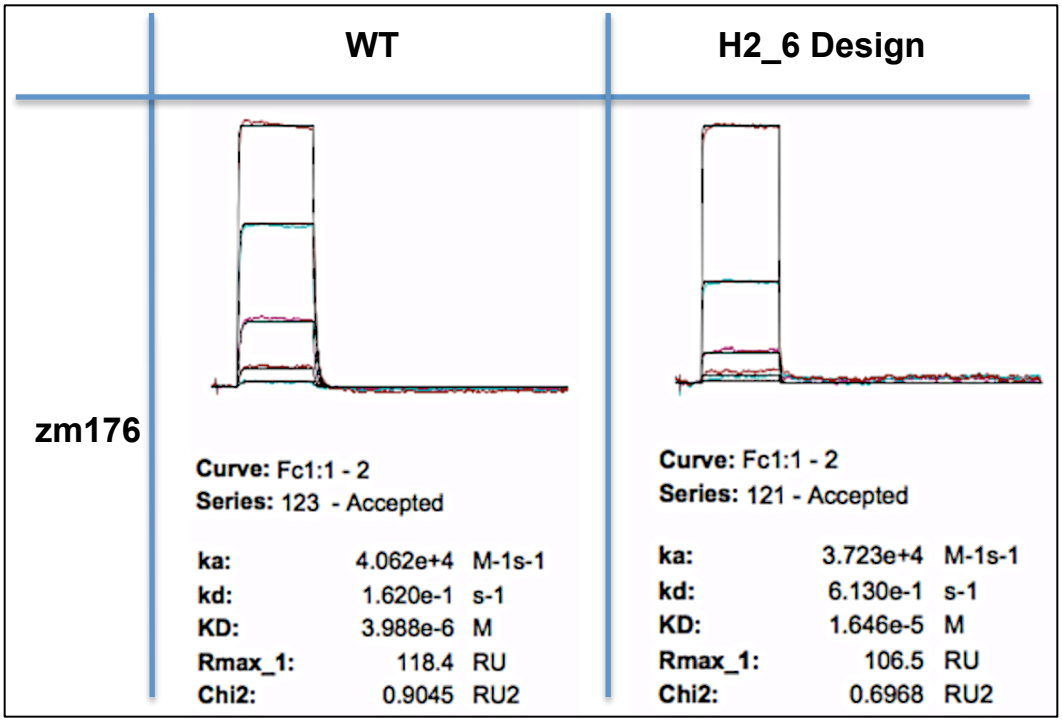


Figure S16: Kinetic sensorgrams of CH103 antibody and H2_6 design to a panel of GP120s. Binding studies were performed on a Biacore 4000 (See Online Methods).

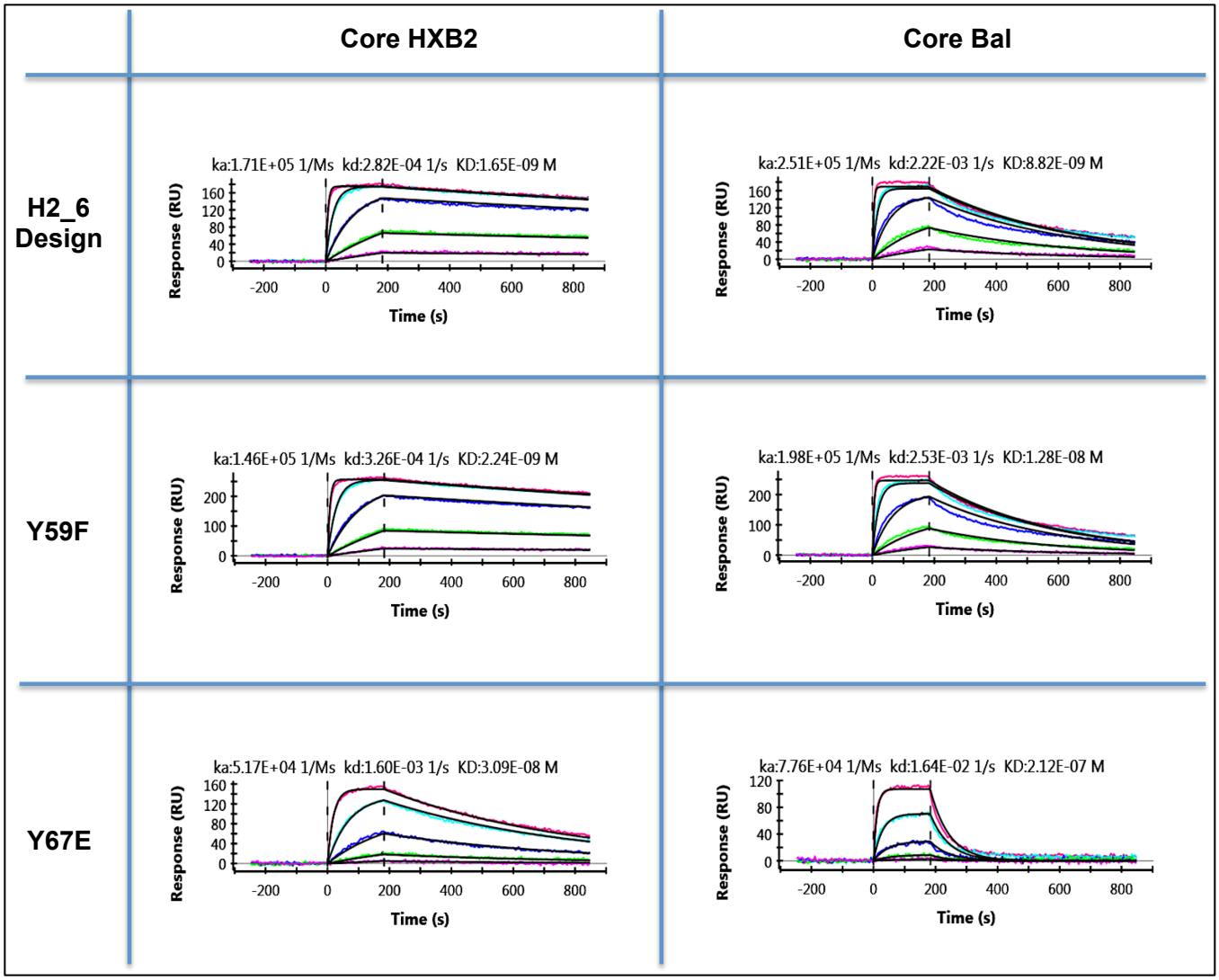


Figure S17: Kinetic Sensorgrams (ProteON XPR) of CH103, H2_6 design mutants with residues from CH103 WT.

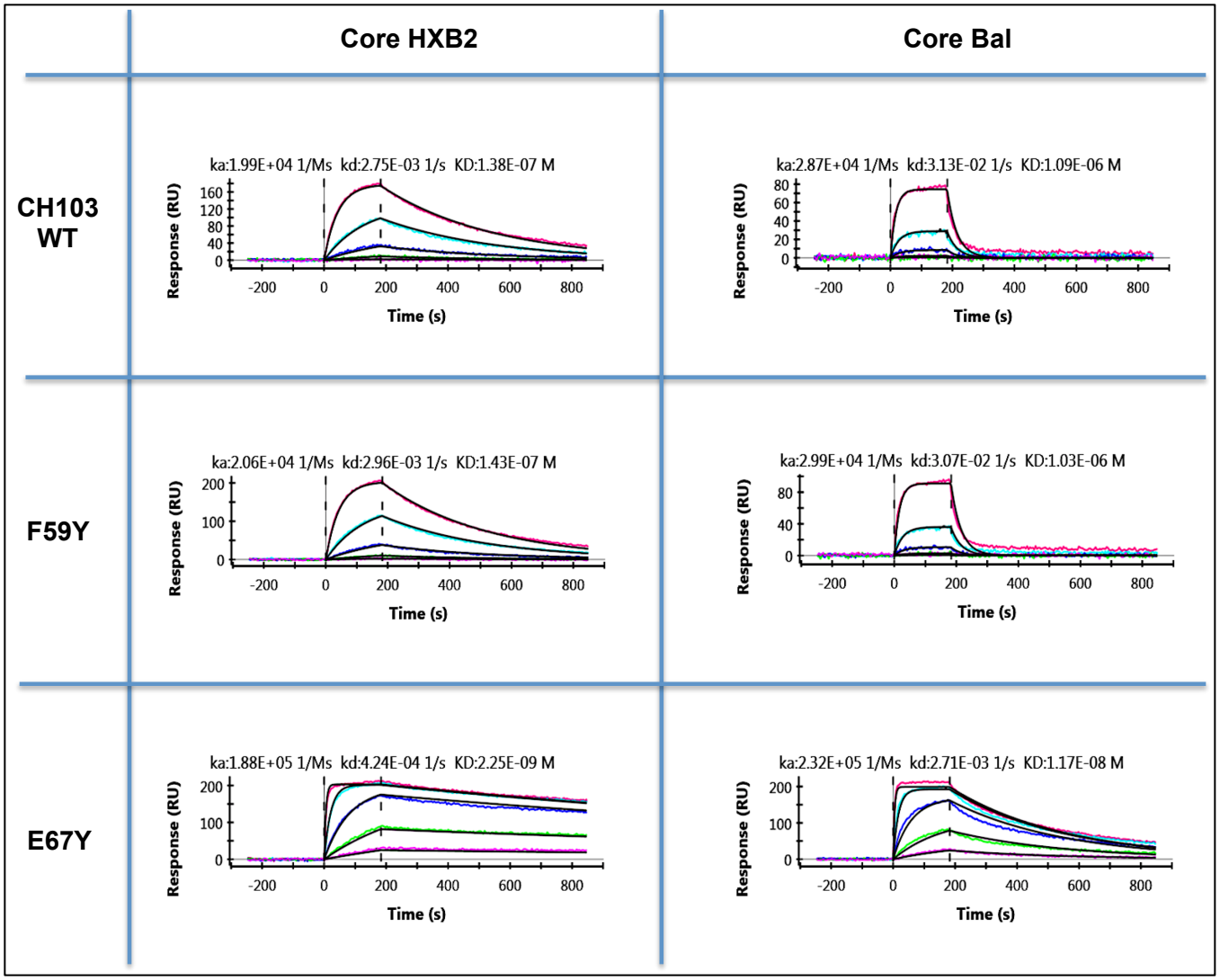


Figure S18: Kinetic Sensorgrams (ProteON XPR) of CH103 WT with H2_6 design residues grafted on.

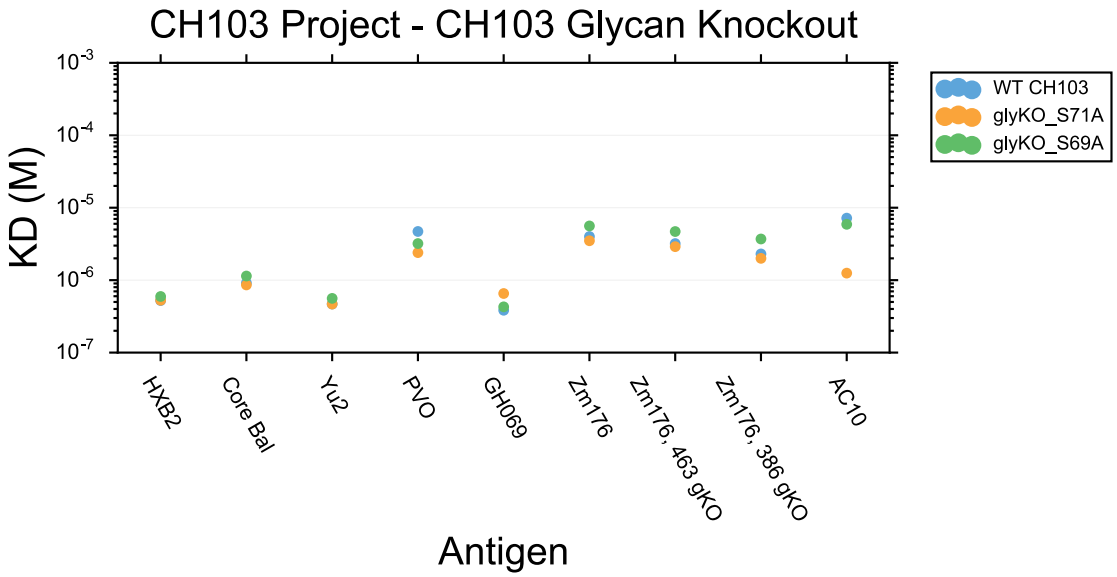


Figure S19: Two potential glycan sites were located in proximity to the antibody paratope for the CH103 antibody (PDB: 4JAN). These sites were mutated from serine to alanine mutations at position 69 and 71 in the heavy chain (AHo-Numbering). Binding affinity (K_d) is shown from Biacore experiments to nine expressed GP120s.

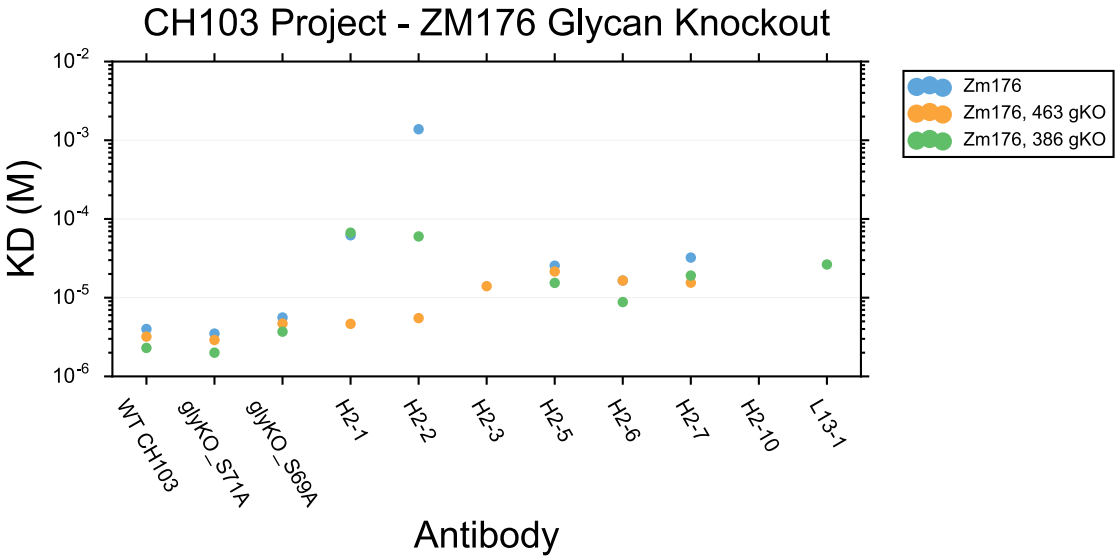


Figure S20: Potential glycan sites were located in proximity to the GP120-CH103 epitope at positions 386 and 463 of Zm176 (PDB ID 4JAN). Each of these sites were knocked out using Serine to Alanine mutations. Binding affinity (K_d) is shown from Biacore experiments to nine expressed GP120s.

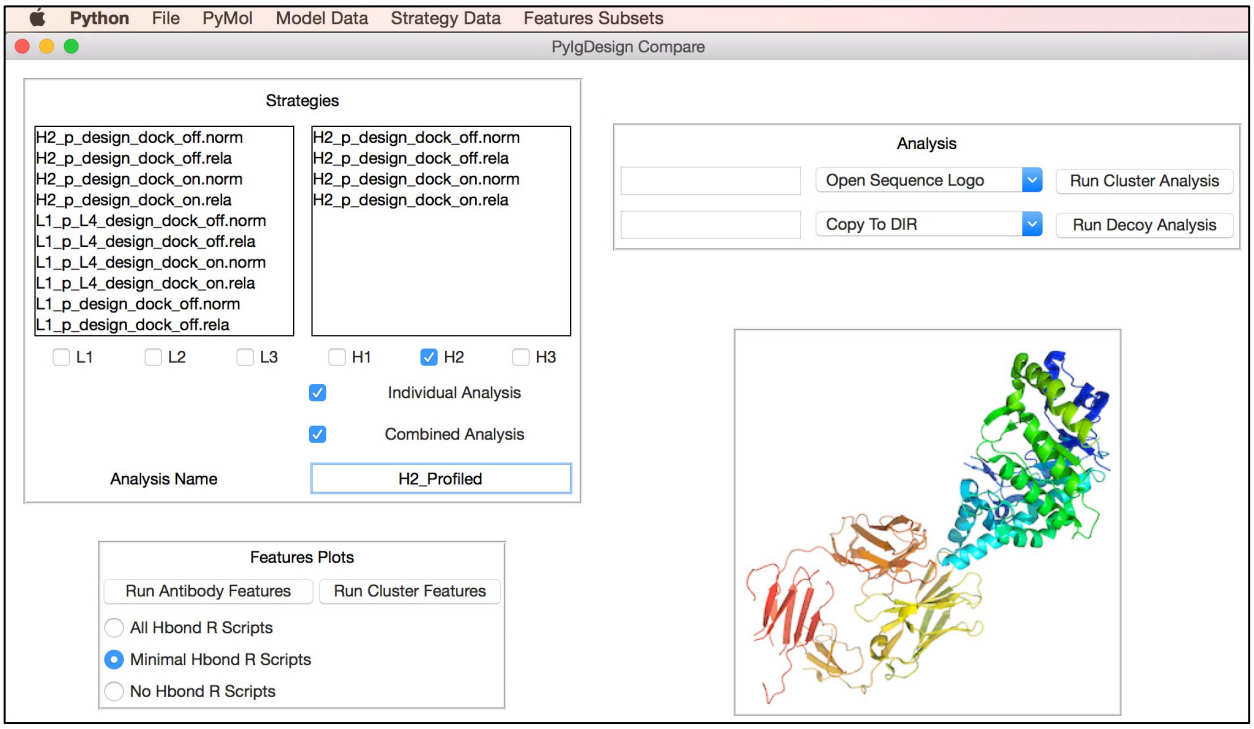


Figure S21: Jade Antibody Design analysis Graphical User Interface (GUI)

Table S1: Native benchmark physical characteristics. Data generated by the Rosetta *AntibodyFeatures* reporter (See Online Methods). Interface energy of separated – complex (**Δ**G) is in Rosetta Energy Units (REU). **Δ**SASA and **Δ**G were calculated through the InterfaceFeatures reporter (which uses the Rosetta InterfaceAnalyzer) by separating the antibody from the antigen and repacking the interface residues both together and separated. An interface distance of 6 Å was used.

	ΔSASA (Å^2)	**ΔG (REU)**	**delta_unsats**	**sc_value**	**packstat**
AVG	1889.45	-6.68	10.20	0.69	0.55
_SD	356.55	24.65	4.35	0.06	0.28
MIN	1371.69	-32.00	4	0.57	0.00
MAX	2571.81	51.83	19	0.77	0.77
1a2y	1371.69	-20.40	4	0.71	0.56
1fe8	1944.60	-19.17	14	0.71	0.70
1ic7	1685.71	-18.46	10	0.71	0.64
1n8z	1538.17	-8.19	11	0.77	0.71
1nsn	1865.76	51.83	11	0.57	0.64
4xnq	1594.52	-11.40	8	0.63	0.00
1tqb	1957.94	-8.69	13	0.72	0.75
2xwt	2571.81	-32.02	7	0.69	0.76
3h3b	1951.91	-27.67	5	0.76	0.00
3nh7	2412.40	26.39	19	0.62	0.63

Table S2: Native benchmark CDR characteristics. Data generated by the Rosetta *AntibodyFeatures* reporter (See Online Methods). Interface energy of separated – complex (dG) is in Rosetta Energy Units (REU). **Δ**SASA and was calculated through the InterfaceFeatures reporter (which uses the Rosetta InterfaceAnalyzer) by separating the antibody from the antigen and repacking the interface residues both together and separated. An interface distance of 6 Å was used.

PDB_CDR	ΔSASA (Å²)	#aromatics	PDB_CDR	ΔSASA (Å²)	#aromatics
1nsn_L1	228.58	2	**1ic7_L1**	110.11	0
1nsn_L2	106.35	1	**1ic7_L2**	126.19	1
1nsn_L3	195.66	2	**1ic7_L3**	123.33	2
1nsn_H1	111.04	3	**1ic7_H1**	193.27	2
1nsn_H2	267.21	2	**1ic7_H2**	241.54	3
1fe8_L1	208.62	1	**1n8z_L1**	131.03	0
1fe8_L2	125.24	2	**1n8z_L2**	110.92	3
1fe8_L3	100.36	1	**1n8z_L3**	163.30	1
1fe8_H1	66.70	2	**1n8z_H1**	21.42	2
1fe8_H2	120.76	1	**1n8z_H2**	165.57	2
4xnq_L1	204.31	5	**2xwt_L1**	204.71	1
4xnq_L2	31.82	1	**2xwt_L2**	118.55	1
4xnq_L3	37.50	2	**2xwt_L3**	175.37	1
4xnq_H1	35.93	1	**2xwt_H1**	124.68	2
4xnq_H2	82.49	4	**2xwt_H2**	134.17	1
1a2y_L1	87.35	1	**1tqb_L1**	210.50	1
1a2y_L2	102.33	2	**1tqb_L2**	206.87	1
1a2y_L3	112.15	2	**1tqb_L3**	58.58	2
1a2y_H1	97.01	2	**1tqb_H1**	160.19	3
1a2y_H2	113.35	1	**1tqb_H2**	227.59	2
3nh7_L1	154.05	1	**3h3b_L1**	232.02	2
3nh7_L2	89.95	1	**3h3b_L2**	1.54	2
3nh7_L3	73.94	1	**3h3b_L3**	84.98	3
3nh7_H1	192.10	3	**3h3b_H1**	72.95	4
3nh7_H2	106.98	1	**3h3b_H2**	237.15	1

Table S3: Binding affinity , CDR design identity, and selection strategies of the expressed and tested 2J88 designs from the Biacore 4000 results.

Name	Binding Affinity (M)	Dock?	Relaxed?	Selection Method	CDR	length	cluster	CDR_sequence
Native 2J88 WT	1.56E-08				L1	11	L1-11-1	RASENIYSYLT
Native 2J88 WT	1.56E-08				H2	9	H2-9-1	HIYWDDDKR
2J88_L14_7	***2.61E-09***	***dock_on***	***unrelaxed***	***dG_top_p_total***	***L1***	***11***	***L1-11-2***	***RSSRDIKDYIT***
2J88_L1_10	***3.12E-09***	***dock_on***	***unrelaxed***	***dG_top_p_total***	***L1***	***11***	***L1-11-1***	***RASKDISDYLT***
2J88_L1_5	***6.22E-09***	***dock_off***	***unrelaxed***	***delta_unsats_per_1000_dSASA***	***L1***	***11***	***L1-11-1***	***RASQDISNYLT***
2J88_H2_8	*Low Rmax*	*dock_on*	*unrelaxed*	*dG_top_p_total*	*H2*	*10*	*H2-10-1*	*AIYPSDGETR*
2J88_L14_3	*Low Rmax*	*dock_off*	*relaxed*	*dG_top_p_total*	*L1*	*15*	*L1-15-1*	*RASESVDSYGTNHIH*
2J88_L14_8	3.38E-08	dock_on	unrelaxed	dG_top_p_total	L1	11	L1-11-2	RSSRNIKDFIS
2J88_H2_4	Low Rmax	dock_off	unrelaxed	delta_unsats_per_1000_dSASA	H2	10	H2-10-2	YISWSGTVTS
2J88_L14_6	5.00E-08	dock_off	relaxed	delta_unsats_per_1000_dSASA	L1	11	L1-11-2	KSSEEIKNFIT
2J88_H2_9	5.32E-08	dock_on	relaxed	dG_top_p_total	H2	10	H2-10-2	EISSDGSRTY
2J88_L1_7	5.97E-08	dock_off	relaxed	delta_unsats_per_1000_dSASA	L1	11	L1-11-2	RASQDIKNNIT
2J88_L1_2	6.81E-08	dock_off	unrelaxed	dG_top_p_total	L1	15	L1-15-1	RASKSVDSYGSSFMS
2J88_H2_1	Low Rmax	dock_off	unrelaxed	dG_top_p_total	H2	10	H2-10-1	EIYPSDGDTR
2J88_H2_5	Low Rmax	dock_off	unrelaxed	delta_unsats_per_1000_dSASA	H2	10	H2-10-1	AIYGEDGETR
2J88_L1_11	2.02E-07	dock_on	relaxed	dG_top_p_total	L1	15	L1-15-1	RASKSVDSYGFSFMS
2J88_L14_9	2.76E-07	dock_on	relaxed	dG_top_p_total	L1	11	L1-11-2	RSSEDIKNFIS
2J88_L1_9	4.30E-07	dock_on	unrelaxed	dG_top_p_total	L1	11	L1-11-1	RASKDISKYIA
2J88_L1_1	4.45E-07	dock_off	unrelaxed	dG_top_p_total	L1	15	L1-15-1	RASESVESYGSSFIS
2J88_L1_6	7.69E-07	dock_off	unrelaxed	delta_unsats_per_1000_dSASA	L1	11	L1-11-1	QASQDVGDALT
2J88_L14_1	Low Rmax	dock_off	unrelaxed	dG_top_p_total	L1	17	L1-17-1	KSSHSLLNSSAQMNYLS
2J88_L1_4	1.38E-06	dock_off	relaxed	dG_top_p_total	L1	17	L1-17-1	KSSQSVLNSRSEKSYLT
2J88_L1_12	1.58E-06	dock_on	relaxed	dG_top_p_total	L1	15	L1-15-1	RASKSVESYGNSFIS
2J88_L14_4	1.59E-06	dock_off	unrelaxed	delta_unsats_per_1000_dSASA	L1	11	L1-11-1	RASQIVSYALS
2J88_L1_8	2.19E-06	dock_off	relaxed	delta_unsats_per_1000_dSASA	L1	15	L1-15-1	RASESVDSYGNSFIS
2J88_L1_3	2.38E-06	dock_off	relaxed	dG_top_p_total	L1	15	L1-15-1	RASESVESYGNSFMH
2J88_L14_2	6.21E-06	dock_off	unrelaxed	dG_top_p_total	L1	17	L1-17-1	KSSQSLLNSNAEKNYLT
2J88_H2_3	9.47E-06	dock_off	relaxed	dG_top_p_total	H2	12	H2-12-1	EIRSKADGSATH
2J88_L14_5	0	dock_off	unrelaxed	delta_unsats_per_1000_dSASA	L1	11	L1-11-2	RSSRDIKNAIS
2J88_H2_6	0	dock_off	relaxed	delta_unsats_per_1000_dSASA	H2	10	H2-10-6	WINLDGGSTS
2J88_H2_2	0	dock_off	unrelaxed	dG_top_p_total	H2	12	H2-12-1	EIGSKSFGGETK
2J88_H2_7	0	dock_on	unrelaxed	dG_top_p_total	H2	10	H2-10-1	MIYPSDGDTR

Table S4: Binding affinity of expressed and purified CH103 designs and their selection strategies.

Name	Antigen	Binding Affinity (M)	by_score_group	strategy
Native	HXB2	5.23E-07		
glyKO_S71A	HXB2	5.28E-07		
glyKO_S69A	HXB2	5.95E-07		
H2-1	HXB2	6.70E-06	dG_top_Ptotal	ch103_H2_all_p.norm
H2-2	HXB2	3.70E-06	dG_top_Ptotal	ch103_H2_all_p.norm
H2-3	HXB2	3.70E-06	dG_top_Ptotal	ch103_H2_all_p.norm
H2-4	HXB2	0	dG_top_Ptotal	ch103_H2_all.rela
H2-5	HXB2	**5.16E-07**	dG_top_Ptotal	ch103_H2_all.rela
H2-6	HXB2	***1.32E-08***	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-7	HXB2	**8.92E-07**	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-8	HXB2	0	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-9	HXB2	0	sc_value	ch103_H2_all_p.rela
H2-10	HXB2	1.67E-05	sc_value	ch103_H2_all.rela
L13-1	HXB2	0	dG_top_Ptotal	ch103_L1L3_all_p.norm
Native	Core Bal	**9.15E-07**		
glyKO_S71A	Core Bal	8.59E-07		
glyKO_S69A	Core Bal	1.14E-06		
H2-1	Core Bal	1.14E-05	dG_top_Ptotal	ch103_H2_all_p.norm
H2-2	Core Bal	7.00E-06	dG_top_Ptotal	ch103_H2_all_p.norm
H2-3	Core Bal	6.00E-06	dG_top_Ptotal	ch103_H2_all_p.norm
H2-4	Core Bal	0.00E+00	dG_top_Ptotal	ch103_H2_all.rela
H2-5	Core Bal	1.20E-06	dG_top_Ptotal	ch103_H2_all.rela
H2-6	Core Bal	***1.69E-08***	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-7	Core Bal	2.40E-06	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-8	Core Bal	0	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-9	Core Bal	0	sc_value	ch103_H2_all_p.rela
H2-10	Core Bal	5.50E-06	sc_value	ch103_H2_all.rela
L13-1	Core Bal	0	dG_top_Ptotal	ch103_L1L3_all_p.norm
Native	Yu2	**4.68E-07**		
glyKO_S71A	Yu2	4.68E-07		
glyKO_S69A	Yu2	5.60E-07		
H2-1	Yu2	0	dG_top_Ptotal	ch103_H2_all_p.norm
H2-2	Yu2	1.18E-05	dG_top_Ptotal	ch103_H2_all_p.norm
H2-3	Yu2	5.30E-06	dG_top_Ptotal	ch103_H2_all_p.norm
H2-4	Yu2	0	dG_top_Ptotal	ch103_H2_all.rela
H2-5	Yu2	6.10E-06	dG_top_Ptotal	ch103_H2_all.rela
H2-6	Yu2	***7.24E-08***	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-7	Yu2	3.70E-06	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela

H2-8	Yu2	0	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-9	Yu2	0	sc_value	ch103_H2_all_p.rela
H2-10	Yu2	6.50E-06	sc_value	ch103_H2_all.rela
L13-1	Yu2	0	dG_top_Ptotal	ch103_L1L3_all_p.norm
Native	PVO	**4.70E-06**		
glyKO_S71A	PVO	2.40E-06		
glyKO_S69A	PVO	3.20E-06		
H2-1	PVO	7.70E-05	dG_top_Ptotal	ch103_H2_all_p.norm
H2-2	PVO	0	dG_top_Ptotal	ch103_H2_all_p.norm
H2-3	PVO	9.20E-06	dG_top_Ptotal	ch103_H2_all_p.norm
H2-4	PVO	0	dG_top_Ptotal	ch103_H2_all.rela
H2-5	PVO	5.80E-06	dG_top_Ptotal	ch103_H2_all.rela
H2-6	PVO	**3.50E-07**	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-7	PVO	5.40E-06	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-8	PVO	0	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-9	PVO	0	sc_value	ch103_H2_all_p.rela
H2-10	PVO	0	sc_value	ch103_H2_all.rela
L13-1	PVO	0	dG_top_Ptotal	ch103_L1L3_all_p.norm
Native	GH069	3.84E-07		
glyKO_S71A	GH069	6.54E-07		
glyKO_S69A	GH069	4.28E-07		
H2-1	GH069	0	dG_top_Ptotal	ch103_H2_all_p.norm
H2-2	GH069	1.18E-06	dG_top_Ptotal	ch103_H2_all_p.norm
H2-3	GH069	5.97E-07	dG_top_Ptotal	ch103_H2_all_p.norm
H2-4	GH069	0	dG_top_Ptotal	ch103_H2_all.rela
H2-5	GH069	5.98E-07	dG_top_Ptotal	ch103_H2_all.rela
H2-6	GH069	**1.42E-07**	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-7	GH069	2.70E-06	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-8	GH069	0	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-9	GH069	0	sc_value	ch103_H2_all_p.rela
H2-10	GH069	0	sc_value	ch103_H2_all.rela
L13-1	GH069	0	dG_top_Ptotal	ch103_L1L3_all_p.norm
Native	Zm176	**4.00E-06**		
glyKO_S71A	Zm176	3.50E-06		
glyKO_S69A	Zm176	5.60E-06		
H2-1	Zm176	6.20E-05	dG_top_Ptotal	ch103_H2_all_p.norm
H2-2	Zm176	1.38E-03	dG_top_Ptotal	ch103_H2_all_p.norm
H2-3	Zm176	0	dG_top_Ptotal	ch103_H2_all_p.norm
H2-4	Zm176	0	dG_top_Ptotal	ch103_H2_all.rela
H2-5	Zm176	2.55E-05	dG_top_Ptotal	ch103_H2_all.rela
H2-6	Zm176	1.65E-05	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela

H2-7	Zm176	3.23E-05	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-8	Zm176	0	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-9	Zm176	0	sc_value	ch103_H2_all_p.rela
H2-10	Zm176	0	sc_value	ch103_H2_all.rela
L13-1	Zm176	0	dG_top_Ptotal	ch103_L1L3_all_p.norm
Native	Zm176, 463 gKO	**3.20E-06**		
glyKO_S71A	Zm176, 463 gKO	2.90E-06		
glyKO_S69A	Zm176, 463 gKO	4.69E-06		
H2-1	Zm176, 463 gKO	**4.65E-06**	dG_top_Ptotal	ch103_H2_all_p.norm
H2-2	Zm176, 463 gKO	**5.50E-06**	dG_top_Ptotal	ch103_H2_all_p.norm
H2-3	Zm176, 463 gKO	1.40E-05	dG_top_Ptotal	ch103_H2_all_p.norm
H2-4	Zm176, 463 gKO	0	dG_top_Ptotal	ch103_H2_all.rela
H2-5	Zm176, 463 gKO	2.15E-05	dG_top_Ptotal	ch103_H2_all.rela
H2-6	Zm176, 463 gKO	1.65E-05	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-7	Zm176, 463 gKO	1.55E-05	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-8	Zm176, 463 gKO	0	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-9	Zm176, 463 gKO	0	sc_value	ch103_H2_all_p.rela
H2-10	Zm176, 463 gKO	0	sc_value	ch103_H2_all.rela
L13-1	Zm176, 463 gKO	**0**	dG_top_Ptotal	ch103_L1L3_all_p.norm
Native	Zm176, 386 gKO	**2.30E-06**		
glyKO_S71A	Zm176, 386 gKO	2.00E-06		
glyKO_S69A	Zm176, 386 gKO	3.70E-06		
H2-1	Zm176, 386 gKO	6.70E-05	dG_top_Ptotal	ch103_H2_all_p.norm
H2-2	Zm176, 386 gKO	6.00E-05	dG_top_Ptotal	ch103_H2_all_p.norm
H2-3	Zm176, 386 gKO	0	dG_top_Ptotal	ch103_H2_all_p.norm
H2-4	Zm176, 386 gKO	0	dG_top_Ptotal	ch103_H2_all.rela
H2-5	Zm176, 386 gKO	1.54E-05	dG_top_Ptotal	ch103_H2_all.rela
H2-6	Zm176, 386 gKO	**8.80E-06**	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-7	Zm176, 386 gKO	1.91E-05	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-8	Zm176, 386 gKO	0	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-9	Zm176, 386 gKO	0	sc_value	ch103_H2_all_p.rela
H2-10	Zm176, 386 gKO	0	sc_value	ch103_H2_all.rela
L13-1	Zm176, 386 gKO	**2.64E-05**	dG_top_Ptotal	ch103_L1L3_all_p.norm

Native	AC10	**7.18E-06**		
glyKO_S71A	AC10	1.25E-06		
glyKO_S69A	AC10	5.90E-06		
H2-1	AC10	0	dG_top_Ptotal	ch103_H2_all_p.norm
H2-2	AC10	0	dG_top_Ptotal	ch103_H2_all_p.norm
H2-3	AC10	0	dG_top_Ptotal	ch103_H2_all_p.norm
H2-4	AC10	0	dG_top_Ptotal	ch103_H2_all.rela
H2-5	AC10	0	dG_top_Ptotal	ch103_H2_all.rela
H2-6	AC10	**4.26E-06**	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-7	AC10	0	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-8	AC10	0	delta_unsats_per_1000_dSASA	ch103_H2_all_p.rela
H2-9	AC10	0	sc_value	ch103_H2_all_p.rela
H2-10	AC10	0	sc_value	ch103_H2_all.rela
L13-1	AC10	0	dG_top_Ptotal	ch103_L1L3_all_p.norm

Table S5: *CDRClusterFeatures* reporter tables

Column	Description
Table cdr_clusters	
struct_id	Structure number – used to query other tables for same loaded structure
cdr_start	Starting number of CDR
cdr_end	Ending number of CDR
chain	Chain of CDR – L or H
CDR	Identity of CDR
length	Length of CDR
cluster	Identified cluster of CDR
dis	Calculated Dihedral Distance to cluster center
normDis	Dihedral Distance normalized by length (radians)
normDis_deg	Dihedral Distance normalized by length (degrees)
sequence	Amino acid sequence of the CDR

Table S6: *InterfaceFeatures* reporter tables

Column	Description
Table interfaces	
struct_id	Structure number – used to query other tables for same loaded structure
interface	Interface ID – The interface that is being analyzed. The ID lists the chains that make up the interface - such as L_A or LH_A. Used to query other tables
chains_side1	Interface Side 1 ID – chains that make up interface side 1 such as L or LH. Used to query other tables
chains_side2	Same as above, but for side 2
nchains_side1	Number of chains that make up side 1
nchains_side2	Number of chains that make up side 2
dSASA	Difference in Buried Solvent Accessible Surface Area as described above *(SASA separated - SASA together)*
dSASA_hphobic	Hydrophobic component of dSASA
dSASA_polar	Polar component of dSASA
dG	Difference in Rosetta energy as described above *(E together - E separated)*
dG_cross	Total cross interface energy of pair-wise terms. Close to dG.
dG_dev_dSASAx100	dG normalized by dSASA. Roughly an energy density calculation.
dG_cross_dev_dSASAx100	dG_cross normalized by dSASA. Roughly an energy density calculation.
hbond_E_fraction	Fraction of dG attributable to cross interface hydrogen bonds.
sc_value	Shape Complementarity value
packstat	Packing quality measure
nres_int	Total interface residues
nres_all	Total structure residues
complex_normalized	Total energy/Total residues
Table interface_sides	
struct_id	Structure number – used to query other tables for same loaded structure
interface	Interface ID – The interface that is being analyzed. The ID lists the chains that make up the interface - such as L_A or LH_A. Used to query other tables
side	The Side that this data pertains to. side1 or

	side2.
chains_side1	Interface Side 1 ID – chains that make up interface side 1 such as L or LH. Used to query other tables
chains_side2	Same as above, but for side 2
interface_nres	Total number of residues for interface side
dSASA	dSASA of interface side
dSASA_sc	dSASA of the side chains of the interface side
dhSASA	Hydrophobic dSASA of the interface side
dhSASA_sc	Hydrophobic dSASA of the side chains of the interface side
dG	dG of the interface side
energy_int	Energy of the interface residues complexed
energy_sep	Energy of the interface residues separated
avg_per_residue_energy_dG	Average per-residue dG of the side
avg_per_residue_energy_int	Average per-residue energy of the side while complexed
avg_per_residue_energy_sep	Average per-residue energy of the side while separated
avg_per_residue_dSASA	Average per-residue dSASAS of the side
avg_per_residue_SASA_int	Average per-residue total SASA of the side while complexed
avg_per_residue_SASA_sep	Average per-residue total SASA of the side while separated
aromatic_fraction	Fraction of total residues that are aromatic
aromatic_dSASA_fraction	Fraction of aromatic dSASA from total dSASA of the side
aromatic_dG_fraction	Fraction of aromatic dG from total dG of the side
interface_to_surface_fraction	Fraction of side interface residues to total surface residues
ss_sheet_fraction	Fraction of side interface residues that are identified as beta-sheet from DSSP
ss_helix_fraction	Fraction of side interface residues that are identified as helix from DSSP
ss_loop_fraction	Fraction of side interface residues that are identified as loop from DSSP
Table interface_residues	
struct_id	Structure number – used to query other tables for same loaded structure
interface	Interface ID – The interface that is being analyzed. The ID lists the chains that make up the interface - such as L_A or LH_A. Used to query other tables
resNum	Residue number of particular residue for which this data reports.
chains_side1	Interface Side 1 ID – chains that make up interface side 1 such as L or LH. Used to query other tables

chains_side2	Same as above, but for side 2
side	The Side that this residue is part of. Side1 or side2.
dSASA	dSASA of the residue
dSASA_sc	dSASA of the side-chain of the residue
dhSASA	Hydrophobic dSASA of the residue
dhSASA_sc	Hydrophobic dSASA of the side-chain of the residue
SASA_int	Total SASA of the residue while complexed
SASA_sep	Total SASA of the residue while separated
relative_dSASA_fraction	Fraction of dSASA of the residue relative to the total SASA separated. 0 indicates no burial, while 1 indicates total burial.
dG	dG of the residue
energy_int	energy of the residue while complexed
energy_sep	energy of the residue while separated

Table S7: Additional *AntibodyFeatures* reporter tables

Column	Description
Table ab_metrics	
struct_id	Structure number – used to query other tables for same loaded structure
numbering_scheme	The numbering scheme used to define the antibody
cdr_definition	The CDR Definition used to define the start and end points of the CDR
cdr_residues	The total number of CDR residues
antigen_present	Is an antigen present?
antigen_chains	If so, what are its chains?
net_charge	Net charge of the antibody
paratope_charge	Net charge of the paratope
paratope_SASA	Total SASA of the paratope
paratope_hSASA	Total Hydrophobic SASA of the paratope
paratope_pSASA	Total Polar SASA of the paratope
VL_VH_packing_angle	VH/VL Packing Angle
VL_VH_distance	VH/VL Distance
VL_VH_opening_angle	VH/VL Opening Angle
VL_VH_opposite_opening_angle	VH/VL Opposite Opening Angle
is_camelid	Is the antibody a Camelid antibody?
Table cdr_metrics	
struct_id	Structure number – used to query other tables for same loaded structure
CDR	The identity of the CDR
length	The length of the CDR
start	Starting resNum of the CDR
end	Ending resNum of the CDR
ag_ab_contacts_total	Total unique Antigen/Antibody contacts from this CDR. *An atomic contact is defined as at least 5 atoms of the antigen that are within 5 Angstroms of an antibody atom*
ag_ab_contacts_nres	Total number of residues making contact with Antigen
ag_ab_dSASA	dSASA of CDR/Antigen interface
ag_ab_dSASA_sc	dSASA of only sidechains of CDR/Antigen interface
ag_ab_dhSASA	Hydrophobic dSASA of CDR/Antigen interface
ag_ab_dhSASA_sc	Hydrophobic dSASA of only sidechains of CDR/Antigen interface
ag_ab_dG	dG of CDR/Antigen interface
SASA	Total SASA of the CDR
charge	Total charge of the CDR
energy	Total energy of the CDR

anchor_CN_distance	Distance from C of first residue to N of last residue of the CDR
aromatic_nres	Total aromatic residues of this CDR
Table cdr_residues	
struct_id	Structure number – used to query other tables for same loaded structure
resNum	Residue number of particular residue for which this data reports.
CDR	The CDR for which this residue belongs
position	The position (1 – length) for which this is in the CDR
ag_contacts	The total number of Antigen contacts this residue makes as described above.
Table ab_h3_kink_metrics	
struct_id	Structure number – used to query other tables for same loaded structure
kink_type	The identified Kink type – either Kinked or Extended.
begin	Start of the kink
end	End of the kink
anion_res	Identity of the anion residue of the Kink
cation_res	Identify of the cation residue of the Kink
RD_Hbond_dis	Distance of the sc-sc Hbond across the strands at the beginning of the H3 kink (typically Asp-Arg)
bb_Hbond_dis	Deturns of the bb-bb Hbond across the strands at the begining of the kink (typically Asp-Arg)
Trp_Hbond_dis	Distance of the Trp sc-bb Hbond across the H3 kink residues (n-1 to n+2)
qdis	Distance from the four kink residues of the H3 C-terminal end that make up the qdih below.
qdih	Dihedral angle from the four kink residues of the H3 C-terminal end

REFERENCES

1. Leaver-Fay A, O'Meara MJ, Tyka M, Jacak R, Song Y, Kellogg EH, Thompson J, Davis IW, Pache RA, Lyskov S, Gray JJ, Kortemme T, Richardson JS, Havranek JJ, Snoeyink J, Baker D, Kuhlman B (2013). Scientific benchmarks for guiding macromolecular energy function improvement. *Methods Enzymol* **523**: 109-143.
2. Ihaka R, Gentleman R (1996). R: A language for data analysis and graphics. *J Comp Graphical Stat* **5**: 299-314.
3. R Core Team (2015) R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
4. Adolf-Bryfogle J, Xu Q, North B, Lehmann A, Dunbrack RL, Jr. (2015). PyIgClassify: a database of antibody CDR structural classifications. *Nucleic Acids Res* **43**: D432-438.
5. Lewis SM, Kuhlman BA (2011). Anchored design of protein-protein interfaces. *PloS one* **6**: e20872.
6. Stranges PB, Kuhlman B (2013). A comparison of successful and failed protein interface designs highlights the challenges of designing buried hydrogen bonds. *Protein Science* **22**: 74-82.
7. Le Grand SM, Merz KM (1993). Rapid approximation to molecular surface area via the use of Boolean logic and look‐up tables. *Journal of Computational Chemistry* **14**: 349-352.
8. Lawrence MC, Colman PM (1993). Shape complementarity at protein/protein interfaces. *Journal of molecular biology* **234**: 946-950.
9. Sheffler W, Baker D (2009). RosettaHoles: rapid assessment of protein core packing for structure prediction, refinement, design, and validation. *Protein Sci* **18**: 229-239.
10. Hubbard SJ, Thornton JM (1993) NACCESS. London: Department of Biochemistry and Molecular Biology, University College London.
11. Chothia C (1976). The nature of the accessible and buried surfaces in proteins. *Journal of Molecular Biology* **105**: 1-12.
12. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FT, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* **102**: 3586-3616.
13. Word JM, Lovell SC, LaBean TH, Taylor HC, Zalis ME, Presley BK, Richardson JS, Richardson DC (1999). Visualizing and quantifying molecular goodness-of-fit: small-probe contact dots with explicit hydrogen atoms. *Journal of Molecular Biology* **285**: 1711-1733.
14. Bondi A (1964). van der Waals volumes and radii. *The Journal of physical chemistry* **68**: 441-451.
15. Gavezzotti A (1983). The calculation of molecular volumes and the use of volume analysis in the investigation of structured media and of solid-state organic reactivity. *Journal of the American Chemical Society* **105**: 5220-5225.
16. Abhinandan KR, Martin AC (2010). Analysis and prediction of VH/VL packing in antibodies. *Protein Eng Des Sel* **23**: 689-697.
17. Marze NA, Lyskov S, Gray JJ (2016). Improved prediction of antibody VL-VH orientation. *Protein Eng Des Sel* **29**: 409-418.
18. North B, Lehmann A, Dunbrack RL, Jr. (2011). A new clustering of antibody CDR loop conformations. *J Mol Biol* **406**: 228-256.
19. Chaudhury S, Lyskov S, Gray JJ (2010). PyRosetta: a script-based interface for implementing molecular modeling algorithms using Rosetta. *Bioinformatics* **26**: 689-691.