Supplementary Table 1 Rosetta multi-state designs and rational structure-based designs of both the C_H1/C_L and the V_H/V_L interfaces.

Design Paradigm[a]	#Explicit Designs Tested in Each Family	Average # Residues mutated	Design Pheno-type[b]	Expression[c]	Expression Mismatch WT HC/ Design LC[c]	Expression Mismatch WT LC/ Design HC[c]	Stability[d]
CH1/Cλ Designs							
1	2	3	Repack	+	+	+	+
2 (CRD2β)	6	3	Repack	+	-	-	+
3	1	2	Repack	+	+	+	-
4	1	2	Charge	+	+	+	+
5	4	2	Charge	+	+	-	+
6	2	3	Repack	+	+	+	-
7	4	2	Charge	-	n.d.[e]	n.d.	n.d.
8	3	6	Repack	+/-	-	-	+
9	1	7	Repack	+	-	-	-
10	1	8	Repack	-	n.d.	n.d.	n.d.
11	1	6	Repack	-	n.d.	n.d.	n.d.
12	1	10	Both	-	n.d.	n.d.	n.d.
13	2	5	Both	+	+	-	-
14	1	18	Both	-	n.d.	n.d.	n.d.
15	2	18	Both	-	n.d.	n.d.	n.d.
16	1	19	Both	-	n.d.	n.d.	n.d.
17	2	18	Both	-	n.d.	n.d.	n.d.
18	2	9	Both	+	+	+	+
19	4	16	Both	-	n.d.	n.d.	n.d.
20	1	13	Both	-	n.d.	n.d.	n.d.
1+5 (CRD1)	1	5	Both	+	+	-	+
CH1/Cλ CRD2β optimization							
1	2	4	Repack	+	-	-	+
2	2	3	Repack	+	-	-	+
3	4	4	Repack	+	-	-	++
VH/VL Designs							
1	4	2	Charge	+	-	+	+
2	8	2	Charge	+	-	+	+
3	1	2	Repack	-	n.d.	n.d.	n.d.
4	1	2	Repack	-	n.d.	n.d.	n.d.
5	3	2	Repack	-	n.d.	n.d.	n.d.
6 (VRD2)	3	2	Charge	+	n.d.	n.d.	+
1+2 (VRD1)	1	4	Charge	+	-	-	+

[a]Each paradigm is based on a similar design principle focused around a set of residues where variations within each paradigm may include variances in one or more amino acid changes to search for the best combination. [b]Design phenotypes could rudimentarily be categorized as a repacking of residues within the C_H1/C_L or V_H/V_L interface (repack) or as a swap or introduction of charge-charge interactions (charge) or both (both). [c]Expression and assembly similar to (+) or worse than (-) WT protein and if worse, then additional measurements were not performed. [d]Stability near WT based on residual activity following thermal challenge at temperatures ranging from 50-80 °C. [e]Not determined (n.d.).

Nature Biotechnology: doi:10.1038/nbt.2797

Supplementary Table 2 Ratio of the relative mass spectrometric intensity of two light chains when expressed with a single heavy chain

LC1	LC2	HC	%Assembly[a] (LC1/HC)	%Assembly[a] (LC2/HC)	Expression (μg/mL)
C_H1/C_L Specificity Designs in IgG-lacking V_H/V_L					
$C\lambda_{CRD1}$ LC	$C\lambda_{WT}$ LC	$CH1_{WT}$ HC	36	64	24
$C\lambda_{CRD1}$ LC	$C\lambda_{WT}$ LC	$CH1_{CRD1}$ HC	96	4	18
$C\lambda_{CRD1}$ LC	$C\kappa_{WT}$ LC	$CH1_{WT}$ HC	58	42	28
$C\lambda_{CRD1}$ LC	$C\kappa_{WT}$ LC	$CH1_{CRD1}$ HC	100	0	61
$C\lambda_{CRD2}$ LC	$C\lambda_{WT}$ LC	$CH1_{WT}$ HC	7±3	93±3	29±18
$C\lambda_{CRD2}$ LC	$C\lambda_{WT}$ LC	$CH1_{CRD2}$ HC	99.7±0.3	0.3±0.3	46±30
$C\lambda_{CRD2}$ LC	$C\kappa_{WT}$ LC	$CH1_{WT}$ HC	5±5	95±5	47±27
$C\lambda_{CRD2}$ LC	$C\kappa_{WT}$ LC	$CH1_{CRD2}$ HC	100±0.1	0±0.1	60±30
C_H1/C_L Specificity Designs in IgG with V_H/V_L					
$VL_{WT}C\lambda_{CRD1}$ LC	$VL_{WT}C\kappa_{WT}$ LC	$VH_{WT}CH1_{WT}$ HC	6	100	87
$VL_{WT}C\lambda_{CRD1}$ LC	$VL_{WT}C\kappa_{WT}$ LC	$VH_{WT}CH1_{CRD1}$ HC	15	100	104
$VL_{WT}C\lambda_{CRD2}$ LC	$VL_{WT}C\lambda_{WT}$ LC	$VH_{WT}CH1_{WT}$ HC	50	50	29
$VL_{WT}C\lambda_{CRD2}$ LC	$VL_{WT}C\lambda_{WT}$ LC	$VH_{WT}CH1_{CRD2}$ HC	79	21	28
$VL_{WT}C\lambda_{CRD2}$ LC	$VL_{WT}C\kappa_{WT}$ LC	$VH_{WT}CH1_{WT}$ HC	78	22	31
$VL_{WT}C\lambda_{CRD2}$ LC	$VL_{WT}C\kappa_{WT}$ LC	$VH_{WT}CH1_{CRD2}$ HC	79	21	19
V_H/V_L Specificity Designs					
$VL_{WT}C\lambda_{WT}$ LC	$VL_{WT}C\kappa_{WT}$ LC	$VH_{WT}CH1_{WT}$ HC	18	82	69
$VL_{VRD1}C\lambda_{WT}$ LC	$VL_{WT}C\kappa_{WT}$ LC	$VH_{WT}CH1_{WT}$ HC	10	90	73
$VL_{VRD1}C\lambda_{WT}$ LC	$VL_{WT}C\kappa_{WT}$ LC	$VH_{VRD1}CH1_{WT}$ HC	61	39	108
$VL_{VRD1}C\kappa_{WT}$ LC	$VL_{WT}C\lambda_{WT}$ LC	$VH_{WT}CH1_{WT}$ HC	40	60	132
$VL_{VRD1}C\kappa_{WT}$ LC	$VL_{WT}C\lambda_{WT}$ LC	$VH_{VRD1}CH1_{WT}$ HC	69	31	112
$VL_{VRD2}C\lambda_{WT}$ LC	$VL_{WT}C\kappa_{WT}$ LC	$VH_{WT}CH1_{WT}$ HC	39	61	94
$VL_{VRD2}C\lambda_{WT}$ LC	$VL_{WT}C\kappa_{WT}$ LC	$VH_{VRD2}CH1_{WT}$ HC	56	44	105
$VL_{VRD2}C\lambda_{WT}$ LC	$VL_{VRD1}C\kappa_{WT}$ LC	$VH_{VRD1}CH1_{WT}$ HC	23	77	95
$VL_{VRD2}C\lambda_{WT}$ LC	$VL_{VRD1}C\kappa_{WT}$ LC	$VH_{VRD2}CH1_{WT}$ HC	71	29	71
Combination of V_H/V_L and C_H1/C_L Specificity Designs					
$VL_{VRD1}C\lambda_{CRD2}$ LC	$VL_{WT}C\lambda_{WT}$ LC	$VH_{WT}CH1_{WT}$ HC	15±18	85±18	42±11
$VL_{VRD1}C\lambda_{CRD2}$ LC	$VL_{WT}C\lambda_{WT}$ LC	$VH_{VRD1}CH1_{CRD2}$ HC	73±12	27±12	55±26
$VL_{VRD2}C\lambda_{CRD2}$ LC	$VL_{WT}C\lambda_{WT}$ LC	$VH_{WT}CH1_{WT}$ HC	30±24	70±24	49±17
$VL_{VRD2}C\lambda_{CRD2}$ LC	$VL_{WT}C\lambda_{WT}$ LC	$VH_{VRD2}CH1_{CRD2}$ HC	84±10	16±10	51±28
$VL_{VRD1}C\lambda_{CRD2}$ LC	$VL_{VRD2}C\kappa_{WT}$ LC	$VH_{VRD2}CH1_{WT}$ HC	26±23	74±23	53±29
$VL_{VRD1}C\lambda_{CRD2}$ LC	$VL_{VRD2}C\kappa_{WT}$ LC	$VH_{VRD1}CH1_{CRD2}$ HC	73±6	27±6	73±41
$VL_{VRD1}C\lambda_{CRD2}$ LC	$VL_{VRD2}C\lambda_{WT}$ LC	$VH_{VRD2}CH1_{WT}$ HC	9±1	91±1	75±4
$VL_{VRD1}C\lambda_{CRD2}$ LC	$VL_{VRD2}C\lambda_{WT}$ LC	$VH_{VRD1}CH1_{CRD2}$ HC	89±4	11±4	76±16

[a]The percent assembly was calculated based on the relative area under the deconvoluted mass spectrometry peaks (i.e., proportional to the number of counts hitting the detector) of each of the LCs co-purified bound to the HC prior to mass spectrometry analysis. Purified samples were reduced with DTT prior to analysis. [b]Green highlighted cells indicate the matched LC and HC pairs. The mass spectrometry results are highlighted in green if the matched pair was >50% of

Nature Biotechnology: doi:10.1038/nbt.2797

the assembled product when expressed with the mismatched LC and red if the specific assembly was ≤50%.

Nature Biotechnology: doi:10.1038/nbt.2797

Supplemental Table 3 EpiMatrix Immunogenicity Prediction[a]

Designs	Protein Sequence[b]	Length (#residues)	EpiMatrix Hits[c]	EpiMatrix Score[d]
C_H1 Designs				
Wild-Type	IgG1 (-V_H)	329	131	-12.1
CRD1	IgG1 (-V_H)	329	130	-12.3
CRD2	IgG1 (-V_H)	329	128	-14.3
Cλ Designs				
Wild-Type	Cλ (-V_L)	105	29	-33.7
CRD1	Cλ (-V_L)	105	22	-53.4
CRD2	Cλ (-V_L)	105	26	-42.1
V_H Designs				
Wild-Type	IgG1	448	204	5.7
VRD1	IgG1	448	202	4.6
VRD2	IgG1	448	202	4.2
V_L Designs				
Wild-Type	Kappa LC	212	83	-6.2
VRD1	Kappa LC	212	80	-9.9
VRD2	Kappa LC	212	85	-4.8

[a]A standard set of human major histocompatibility alleles were used for the analysis. [b]Protein Sequence indicates the entire protein that was entered into the program for analysis. IgG1 (-V_H) and Cλ (-V_L) indicates that only the constant region IgG and Cλ sequences were input into the program for analysis. [c]EpiMatrix Hits indicate the number of EpiMatrix Z scores above a certain threshold criteria.[d]EpiMatrix Score is derived from the number and intensity of the EpiMatrix Hits normalized by the protein length. Based on the EpiMatrix manual, numbers above 20 or 30 are considered to have a high immunogenicity risk, while numbers below -20 are deemed to have a low immunogenicity risk.

Nature Biotechnology: doi:10.1038/nbt.2797

Supplementary Table 4 Structure determination statistics

	Wild-type C_H1/C_λ (4LLD)	CRD1 C_H1/C_λ (4LLM)	CRD2β C_H1/C_λ (4LLQ)	Wild-type Pertuzumab Fab (C_λ) (4LLU)	VRD2 Pertuzumab Fab (C_λ) (4LLW)	VRD1_CRD2 Pertuzumab Fab (C_λ) (4LLY)[a]
Crystallization conditions	35% PEG 4K	40% PEG 6K, 10mM tri-Sodium Citrate	39% PEG 6K, 5mM tri-Sodium Citrate	100mM Sodium Acetate pH 4.5, 10% MPD, 30% PEG 2000 MME, 200mM Ammonium Sulfate	15% PEG 8K, 200mM Ammonium Sulfate	100mM Na Citrate tribasic pH 5, 30% Jeffamine ED-2001 pH 7
Cryoprotectant	20% PEG 400	20% Ethylene glycol	20% PEG 400	16% Glycerol	20% Ethylene glycol	20% Glycerol
Data collection						
Resolution range (outer shell) (Å)	66.93-1.19 (1.25-1.19)[b]	71.50-1.75 (1.85-1.75)	71.22-1.42 (1.50-1.42)	19.90-2.16 (2.28-2.16)	36.32-1.95 (2.06-1.95)	86.84-1.60 (1.69-1.60)
Space Group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	C2	$P2_1$	P1
a (Å)	43.08	44.51	42.74	210.02	85.60	52.83
b (Å)	65.84	61.39	62.59	69.74	71.18	52.91
c (Å)	66.93	71.50	71.22	70.83	90.92	89.58
α (°)	90	90	90	90	90	76.89
β (°)	90	90	90	98.75	99.28	84.34
γ (°)	90	90	90	90	90	85.54
Completeness(%)	97.6 (95.6)	99.8 (99.9)	99.3 (98.6)	97.4 (94.5)	97.9 (97.2)	95.9 (94.3)
Mean redundancy	7.2 (7.0)	7.0 (7.2)	7.2 (7.2)	3.6 (3.6)	3.8 (3.8)	4.0 (3.8)
R_{sym} (%)[c]	4.3 (52.8)	7.6 (50.3)	5.8 (52.7)	17.5 (67.0)	11.3 (71.9)	8.2 (54.9)
Mean I/sd(I)	23.8 (3.6)	19.1 (5.6)	19.1 (3.7)	5.0 (1.9)	6.0 (1.7)	9.8 (1.8)
Refinement						
Resolution (Å)	46.93-1.19	37.78-1.75	47.01-1.42	19.90-2.16	35.59-1.95	43.42-1.60
R (%)[d]	17.9	17.9	19.0	22.5	21.2	19.9
R_{free} (%)[e]	19.3	22.0	22.0	27.3	26.2	22.9
rmsd bond lengths (Å)	0.007	0.009	0.012	0.011	0.011	0.022
rmsd angles (°)	1.3	1.3	1.3	1.3	1.5	1.5
Core Ramachandran	98.0	98.5	98.5	95.5	96.6	97.2

Nature Biotechnology: doi:10.1038/nbt.2797

(%)						
Disallowed Ramachandran (%)	0.0	0.0	0.0	0.0	0.1	0.0
Amino Acids	200	202	200	856	843	858
Chains	A,B	A,B	A,B	A,B,C,D	A,B,C,D	A,B,C,D
Chemical Components				Acetate, Sulfate	Sulfate	Magnesium, Glycerol
Waters	330	139	203	245	360	581

[a]Structure was a merger of two datasets. [b]Values in parentheses are for the highest resolution shell. [c]$R_{sym} = \Sigma | I-\langle I \rangle | / \Sigma I$, where I is the observed intensity and <I> is the average intensity over all observations of symmetry-related reflections. [d]$R = \Sigma || F_{obs} | - | F_{calc} || / \Sigma | F_{obs} |$, where $|F_{obs}|$ and $|F_{calc}|$ are the observed and calculated structure factor amplitudes, respectively. [e]R_{free} was calculated from a randomly chosen subset of 5% of the reflections.

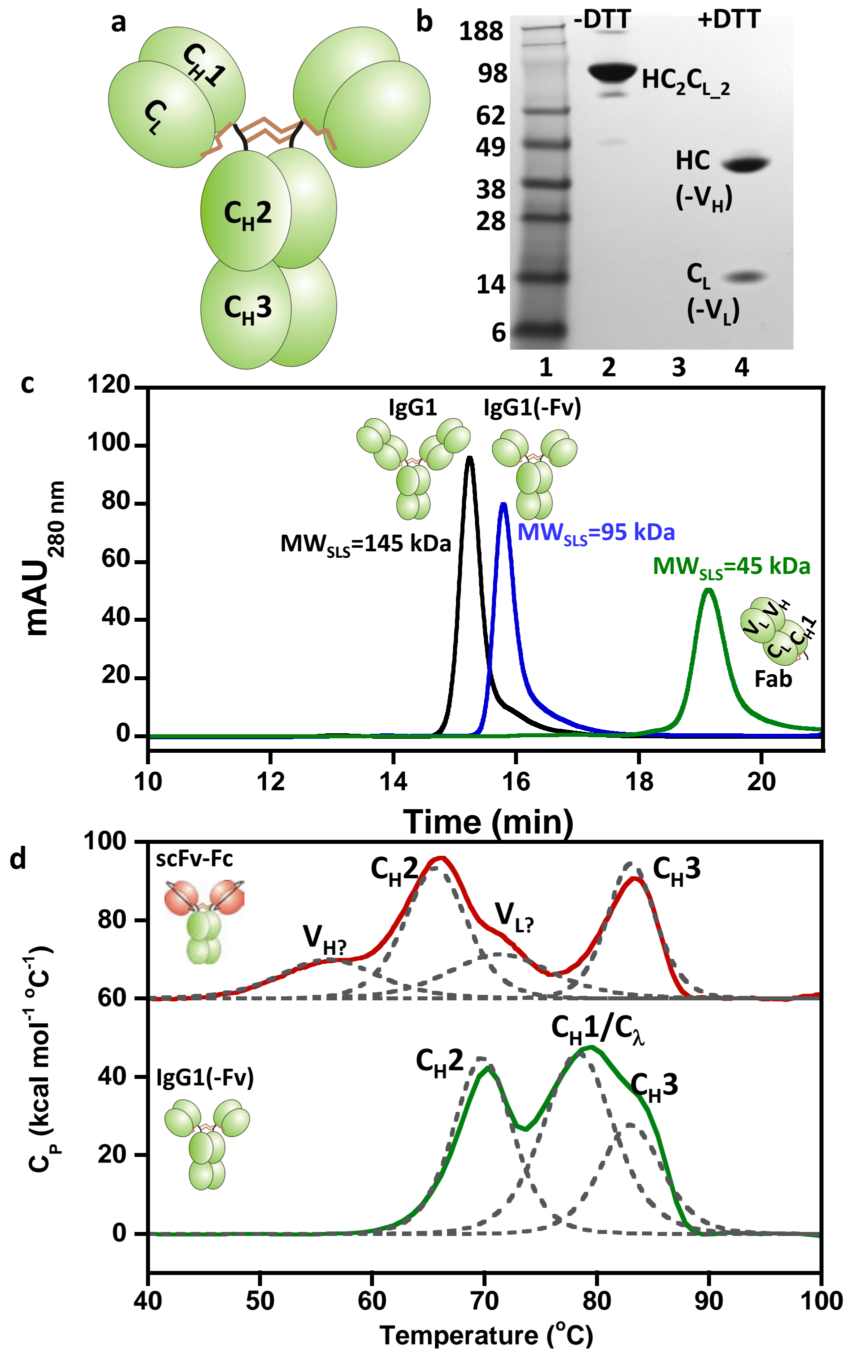
Nature Biotechnology: doi:10.1038/nbt.2797

Supplementary Table 5 Interface positions considered for design

C_H1 position	C_L position	V_H position	V_L position
120	116	37	36
122	118	39	38
124	121	44	43
139	123	45	44
141	124	47	46
143	129	89	49
145	131	91	85
146	133	93	87
169	135	103	98
172	137	106	100
174	160		101
175	162		
177	165		
186	167		
188	168		
190	174		
192	176		
221	178		
	180		

This table lists positions tested for mutation in multistate design. The first column denotes C_H1 residues from PDB 3TV3 that were allowed to mutate in at least one multistate design experiment.

Nature Biotechnology: doi:10.1038/nbt.2797



Nature Biotechnology: doi:10.1038/nbt.2797

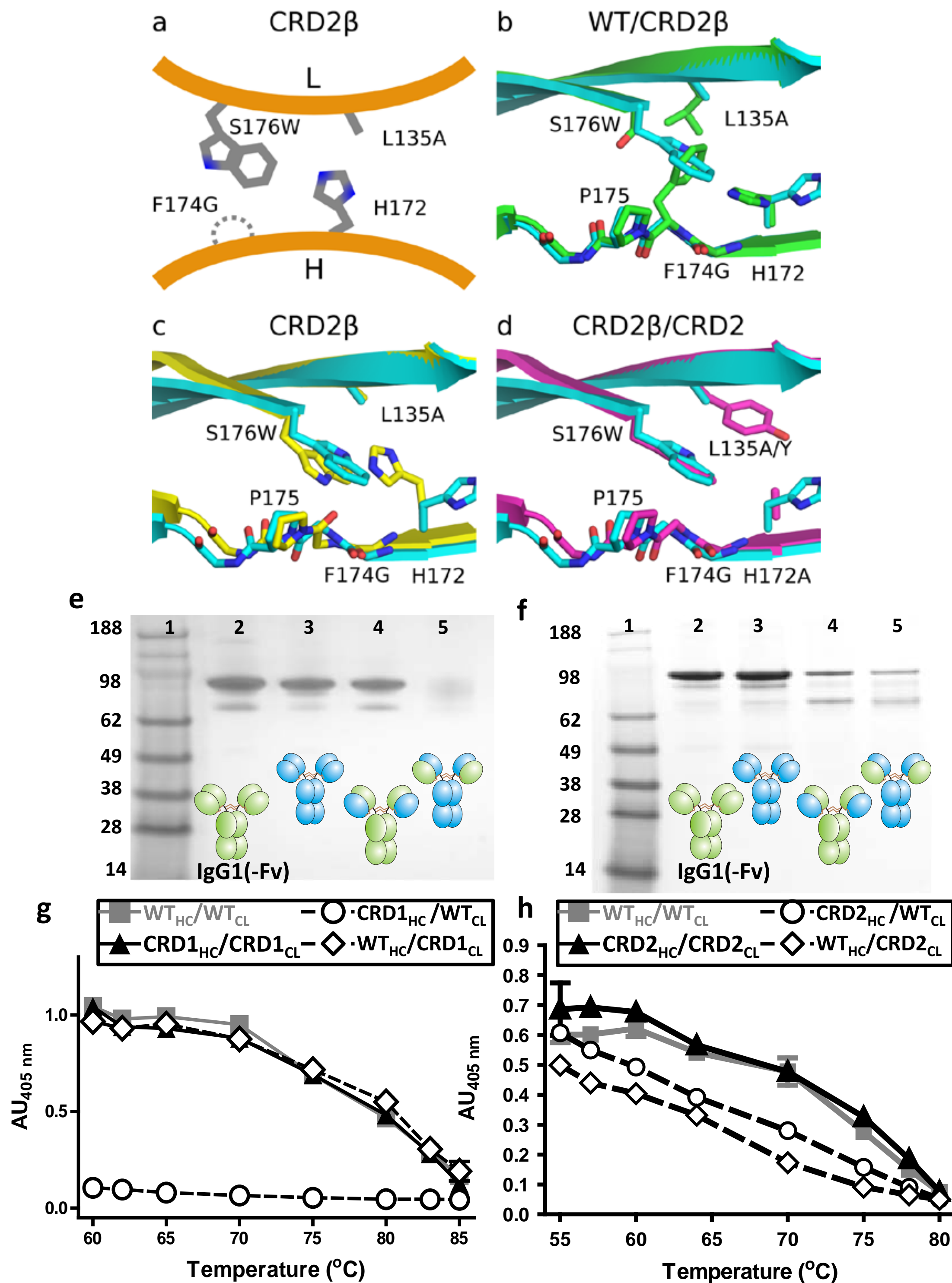
Supplementary Figure 1 Properties of the IgG1(-Fv) protein that lacks V_H and V_L domains and was used for screening the C_H1/C_L interface redesigns. (a) Schematic diagram of the domain architecture of the IgG1(-Fv) protein. Ovals represent the individual Ig-domains and the brown lines represent the disulfide bonds formed in the hinge and between the C_H1 and C_L domains. (b) SDS-PAGE analysis of the non-reduced heterotetramer (95 kDa – Lane 2) protein and the reduced HC (37.5 kDa) and C_L (12 kDa) proteins (Lane 4). (c) Analytical SEC of the WT pertuzumab IgG1, IgG1(-Fv), and pertuzumab IgG1 Fab with the solution molecular weights measured by static light scattering. (d) DSC analysis of the IgG1(-Fv) protein indicating the stability and cooperativity of folding of the C_H1/C_L heterodimeric unit.

Nature Biotechnology: doi:10.1038/nbt.2797

```
                                                  39                            62
WT V_H     EVQLVESGGGLVQPGGSLRLSCAASGFTFTDYTMDWVRQAPGKGLEWVADVNPNSGGSIYNQRFKGRFTLSVDRSKNTLYLQMNSLRAEDTAVYYC
VRD1 V_H   EVQLVESGGGLVQPGGSLRLSCAASGFTFTDYTMDWVRKAPGKGLEWVADVNPNSGGSIYNQEFKGRFTLSVDRSKNTLYLQMNSLRAEDTAVYYC
VRD2 V_H   EVQLVESGGGLVQPGGSLRLSCAASGFTFTDYTMDWVRYAPGKGLEWVADVNPNSGGSIYNQRFKGRFTLSVDRSKNTLYLQMNSLRAEDTAVYYC

           1                                    38
WT V_L     DIQMTQSPSSLSASVGDRVTITCKASQDVSIGVAWYQQKPGKAPKLLIYSASYRYTGVPSRFSGSGSGTDFTLTISSLQPEDFATYYC
VRD1 V_L   RIQMTQSPSSLSASVGDRVTITCKASQDVSIGVAWYQDKPGKAPKLLIYSASYRYTGVPSRFSGSGSGTDFTLTISSLQPEDFATYYC
VRD2 V_L   DIQMTQSPSSLSASVGDRVTITCKASQDVSIGVAWYQRKPGKAPKLLIYSASYRYTGVPSRFSGSGSGTDFTLTISSLQPEDFATYYC

                                          146                        172 174             190
WT C_H1    ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKV
CRD1 C_H1  ASTKGPSVFPLAPSSKSTSGGTAALGCLVKKYFPEPVTVSWNSGALTSGVHTTPAVLQSSGLYSLSSFVTVPSSSLGTQTYICNVNHKPSNTKVDKKV
CRD2 C_H1  ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVATGPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKV

                                  129   135                                      176
WT C_λ     GQPKAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTEC
CRD1 C_λ   GQPKAAPSVTLFPPSSEELQANDATLVCFISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTEC
CRD2 C_λ   GQPKAAPSVTLFPPSSEELQANKATLVCYISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASWYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTEC
```

Supplementary Figure 2 Sequence alignment of WT, VRD1_CRD2, and VRD2 Fab HCs and LCs. The designed residues are highlighted in yellow and labeled.

Nature Biotechnology: doi:10.1038/nbt.2797

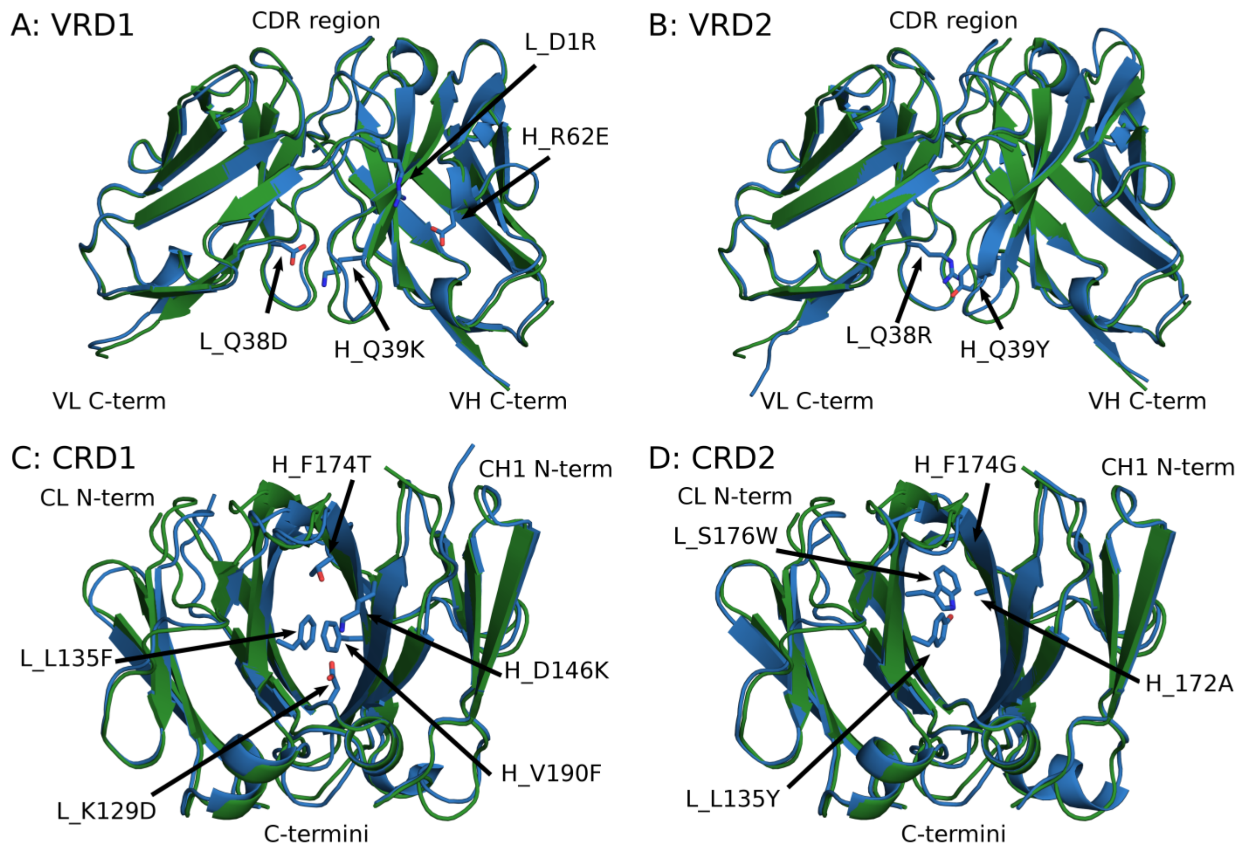


Supplementary Figure 3 Schematic demonstrating the pre-prolyl trans-cis peptide bond isomerization between C_H1_F174G and C_H1_P175 present in the crystal structure of design CRD2β. Panel a shows a schematic of CRD2β (compare Figure 2, panels c and d). Panel b shows, in green, a WT C_H1/C_L interface (repacked, from 3TV3) and, in blue, the design model for CRD2β. Panel c shows the same design model along with the crystal structure of CRD2β (4LLQ) in yellow. Notice the peptide bond isomerization between C_H1_F174G and C_H1_P175 and the incorrectly predicted C_L_S176W rotamer. Panel d shows the

Nature Biotechnology: doi:10.1038/nbt.2797

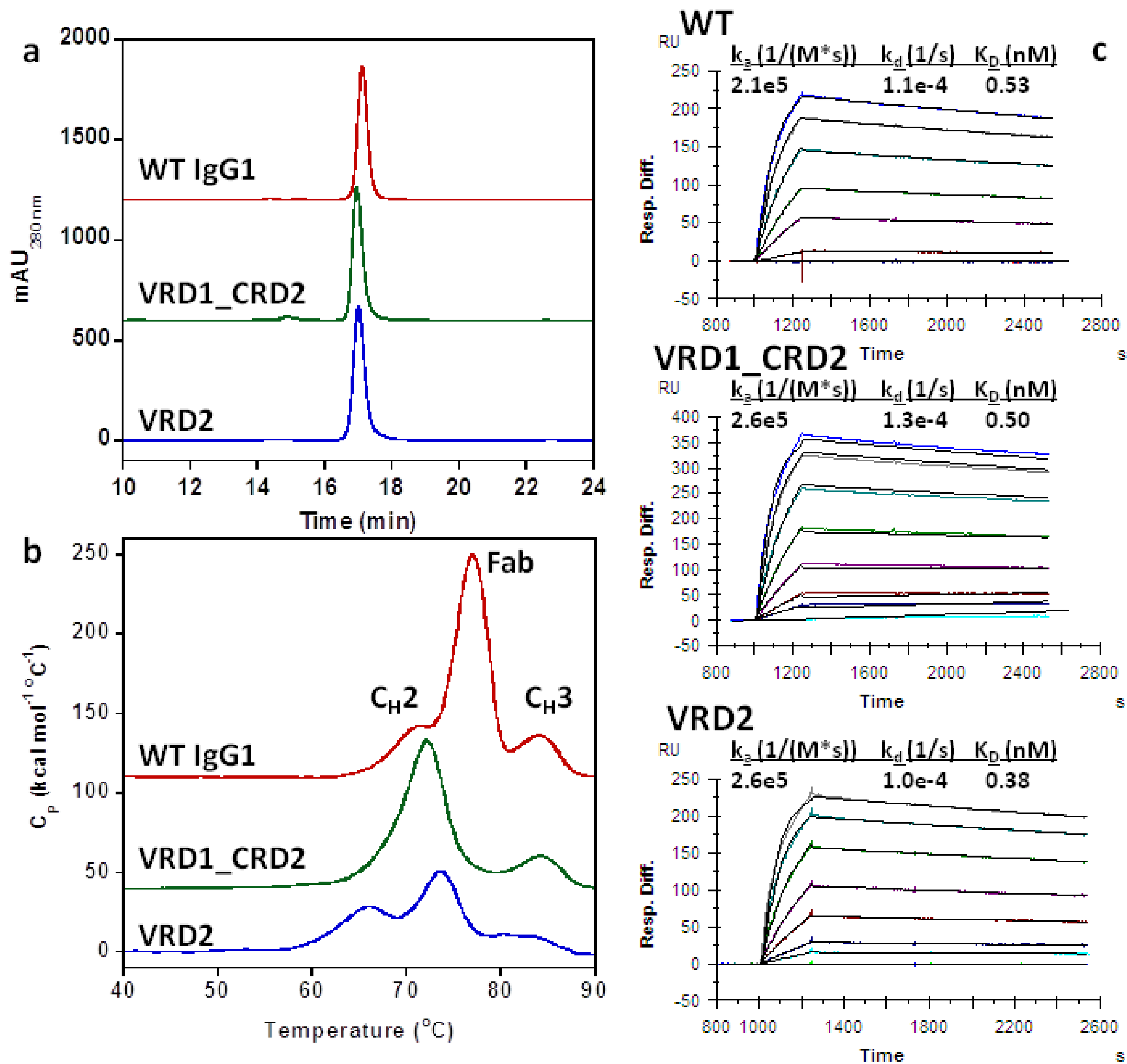
final crystal structure CRD2 (4LLY) in magenta compared with the CRD2β design model. The mutations added to CRD2 bring the C_L_S176W rotamer and C_H1_F174G-C_H1_P175 backbone back in line with the design model. Notice that Rosetta could not have predicted the backbone isomerization because that bond was not free to isomerize during modeling. (e-g) Additional characterization of the specificity afforded by CRD1 and CRD2 in the IgG1(-Fv) format. SDS-PAGE analysis of WT IgG1(-Fv) (depicted in green, lane 1), IgG1(-Fv) with the CRD1 (e) or CRD2 (f) design in C_H1/C_L (depicted in blue, lane 2), and the mismatched pairs (lanes 3-4). Relative stability using a thermal challenge assay of WT IgG1(-Fv) (grey squares), IgG1(-Fv) with CRD1 (g) or CRD2 (h) (black triangles) and mismatched pairs (open circles = CRD HC/WT LC, open diamonds = WT HC/CRD LC). In the thermal challenge assay, the protein samples were heated for 1 h at the specified temperatures, cooled, and assessed for their presence using an anti-Fc/anti-C_L ELISA.

Nature Biotechnology: doi:10.1038/nbt.2797



Supplementary Figure 4 In each panel, our WT Fab structure is in green, and a mutant structure is in blue. Panels A and B are V_H/V_L dimers (with the CDR loop region at the top), and Panels C and D are the C_H1/C_L dimer. Mutations are highlighted. In Panel A, we compare the VRD1 mutation in the V_H/V_L dimer. There is very little global structural change (0.37 Å RMSD), with a small change to the loop containing the Q39K mutation, which is opposite the CDR region. Panel B shows design VRD2, again with a small change in the loop conformation containing Q39Y, and little global change (0.39 Å RMSD). Panel C shows design CRD1, with a small 0.72 Å RMSD due to a minor rotation of the rigid body orientation of the two halves of the dimer. Panel D shows design CRD2, with a global RMSD of 0.67 Å. The apparent change in the helix in the lower right corner of Panel D is due to a crystal contact in the CRD2 design crystal (blue). Some of the sidechain atoms were modeled in D1R in VRD1 and D146K in CRD1 due to missing electron density.

Nature Biotechnology: doi:10.1038/nbt.2797



Supplementary Figure 5 Analytical SEC (a) and DSC (b) of WT pertuzumab IgG1 and pertuzumab variants containing both VRD1 and CRD2 or VRD2. (c) Kinetic biacore analyses of WT (top), VRD2 (middle), and VRD1_CRD2 (bottom) pertuzumab Fabs binding to immobilized HER-2 antigen. Fab concentrations were 50, 35, 20, 10, 5, 2, 1 nM. Results of fits to a 1:1 kinetic binding model are shown above each experiment. Except for the DSC, Cλ was in both LCs.

Nature Biotechnology: doi:10.1038/nbt.2797

Supplemental Figure 6 Sequence alignments of all the HC and LC sequences utilized in the generation of the IgG1 BsAbs described in **Table 1**. The variable domains (including D and J linkers) are highlighted in yellow. The constant domains are highlighted in blue. Mutations that enable HC/LC specificity are highlighted in red. The V_H and Vκ germline family from which each variable domain is derived is listed next to the wild-type sequences. The constant domain (CLambda or CKappa) that was used in each LC is also listed.

Pertuzumab_HC

```
Pertuzumab_WT_VH3          EVQLVESGGGLVQPGGSLRLSCAASGFTFTDYTMDWVRQAPGKGLEWVADVNPNSGGSIY
Pertuzumab_VRD1CRD2        EVQLVESGGGLVQPGGSLRLSCAASGFTFTDYTMDWVRKAPGKGLEWVADVNPNSGGSIY
Pertuzumab_VRD2            EVQLVESGGGLVQPGGSLRLSCAASGFTFTDYTMDWVRYAPGKGLEWVADVNPNSGGSIY

Pertuzumab_WT_VH3          NQRFKGRFTLSVDRSKNTLYLQMNSLRAEDTAVYYCARNLGPSFYFDYWGQGTLVTVSSA
Pertuzumab_VRD1CRD2        NQEFKGRFTLSVDRSKNTLYLQMNSLRAEDTAVYYCARNLGPSFYFDYWGQGTLVTVSSA
Pertuzumab_VRD2            NQRFKGRFTLSVDRSKNTLYLQMNSLRAEDTAVYYCARNLGPSFYFDYWGQGTLVTVSSA

Pertuzumab_WT_VH3          STKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSG
Pertuzumab_VRD1CRD2        STKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVATGPAVLQSSG
Pertuzumab_VRD2            STKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSG

Pertuzumab_WT_VH3          LYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSC
Pertuzumab_VRD1CRD2        LYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSC
Pertuzumab_VRD2            LYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSC
```

Pertuzumab_LC

```
Pertuzumab_WT_VK1CLambda      DIQMTQSPSSLSASVGDRVTITCKASQDVSIGVAWYQQKPGKAPKLLIYS
Pertuzumab_VRD1CRD2           RIQMTQSPSSLSASVGDRVTITCKASQDVSIGVAWYQDKPGKAPKLLIYS
Pertuzumab_VRD2_VK1CKappa     DIQMTQSPSSLSASVGDRVTITCKASQDVSIGVAWYQRKPGKAPKLLIYS

Pertuzumab_WT_VK1CLambda      ASYRYTGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQYYIYPYTFGQ
Pertuzumab_VRD1CRD2           ASYRYTGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQYYIYPYTFGQ
Pertuzumab_VRD2_VK1CKappa     ASYRYTGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQYYIYPYTFGQ

Pertuzumab_WT_VK1CLambda      GTKVEIKGQPKAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWK
Pertuzumab_VRD1CRD2           GTKVEIKGQPKAAPSVTLFPPSSEELQANKATLVCYISDFYPGAVTVAWK
Pertuzumab_VRD2_VK1CKappa     GTKVEIK-RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWK

Pertuzumab_WT_VK1CLambda      ADSSPVKAG-VETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHE
Pertuzumab_VRD1CRD2           ADSSPVKAG-VETTTPSKQSNNKYAAWSYLSLTPEQWKSHRSYSCQVTHE
Pertuzumab_VRD2_VK1CKappa     VDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQ

Pertuzumab_WT_VK1CLambda      G--STVEKTVAPTEC
Pertuzumab_VRD1CRD2           G--STVEKTVAPTEC
Pertuzumab_VRD2_VK1CKappa     GLSSPVTKSFNRGEC
```

Matuzumab_HC

```
Matuzumab_WT_VH1          QVQLVQSGAEVKKPGASVKVSCKASGYTFTSHWMHWVRQAPGQGLEWIGEFNPSNGRTNY
Matuzumab_VRD1CRD2        QVQLVQSGAEVKKPGASVKVSCKASGYTFTSHWMHWVRKAPGQGLEWIGEFNPSNGRTNY
Matuzumab_VRD2            QVQLVQSGAEVKKPGASVKVSCKASGYTFTSHWMHWVRYAPGQGLEWIGEFNPSNGRTNY

Matuzumab_WT_VH1          NEKFKSKATMTVDTSTNTAYMELSSLRSEDTAVYYCASRDYDYDGRYFDYWGQGTLVTVS
Matuzumab_VRD1CRD2        NEEFKSKATMTVDTSTNTAYMELSSLRSEDTAVYYCASRDYDYDGRYFDYWGQGTLVTVS
Matuzumab_VRD2            NEKFKSKATMTVDTSTNTAYMELSSLRSEDTAVYYCASRDYDYDGRYFDYWGQGTLVTVS

Matuzumab_WT_VH1          SASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQS
Matuzumab_VRD1CRD2        SASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVATGPAVLQS
Matuzumab_VRD2            SASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQS

Matuzumab_WT_VH1          SGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSC
Matuzumab_VRD1CRD2        SGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSC
Matuzumab_VRD2            SGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSC
```


Nature Biotechnology: doi:10.1038/nbt.2797

Matuzumab_LC

```
Matuzumab_WT_VK1CLambda      DIQMTQSPSSLSASVGDRVTITCSASSSVTYMYWYQQKPGKAPKLLIYDT
Matuzumab_VRD1CRD2           RIQMTQSPSSLSASVGDRVTITCSASSSVTYMYWYQDKPGKAPKLLIYDT
Matuzumab_VRD2               DIQMTQSPSSLSASVGDRVTITCSASSSVTYMYWYQRKPGKAPKLLIYDT

Matuzumab_WT_VK1CLambda      SNLASGVPSRFSGSGSGTDYTFTISSLQPEDIATYYCQQWSSHIFTFGQG
Matuzumab_VRD1CRD2           SNLASGVPSRFSGSGSGTDYTFTISSLQPEDIATYYCQQWSSHIFTFGQG
Matuzumab_VRD2               SNLASGVPSRFSGSGSGTDYTFTISSLQPEDIATYYCQQWSSHIFTFGQG

Matuzumab_WT_VK1CLambda      TKVEIKGQPKAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKA
Matuzumab_VRD1CRD2           TKVEIKGQPKAAPSVTLFPPSSEELQANKATLVCYISDFYPGAVTVAWKA
Matuzumab_VRD2               TKVEIKGQPKAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKA

Matuzumab_WT_VK1CLambda      DSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGS
Matuzumab_VRD1CRD2           DSSPVKAGVETTTPSKQSNNKYAAWSYLSLTPEQWKSHRSYSCQVTHEGS
Matuzumab_VRD2               DSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGS

Matuzumab_WT_VK1CLambda      TVEKTVAPTEC
Matuzumab_VRD1CRD2           TVEKTVAPTEC
Matuzumab_VRD2               TVEKTVAPTEC
```

METMAb_HC

```
MetMAb_WT_VH3          EVQLVESGGGLVQPGGSLRLSCAASGYTFTSYWLHWVRQAPGKGLEWVGMIDPSNSDTRF
MetMAb_VRD1CRD2        EVQLVESGGGLVQPGGSLRLSCAASGYTFTSYWLHWVRKAPGKGLEWVGMIDPSNSDTRF

MetMAb_WT_VH3          NPNFKDRFTISADTSKNTAYLQMNSLRAEDTAVYYCATYRSYVTPLDYWGQGTLVTVSSA
MetMAb_VRD1CRD2        NPEFKDRFTISADTSKNTAYLQMNSLRAEDTAVYYCATYRSYVTPLDYWGQGTLVTVSSA

MetMAb_WT_VH3          STKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSG
MetMAb_VRD1CRD2        STKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVATGPAVLQSSG

MetMAb_WT_VH3          LYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSC
MetMAb_VRD1CRD2        LYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSC
```

METMAb_LC

```
MetMAb_WT_VK1CLambda      DIQMTQSPSSLSASVGDRVTITCKSSQSLLYTSSQKNYLAWYQQKPGKAPKLLIYWASTR
MetMAb_VRD1CRD2           RIQMTQSPSSLSASVGDRVTITCKSSQSLLYTSSQKNYLAWYQDKPGKAPKLLIYWASTR

MetMAb_WT_VK1CLambda      ESGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQYYAYPWTFGQGTKVEIKGQPKAAP
MetMAb_VRD1CRD2           ESGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQYYAYPWTFGQGTKVEIKGQPKAAP

MetMAb_WT_VK1CLambda      SVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYA
MetMAb_VRD1CRD2           SVTLFPPSSEELQANKATLVCYISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYA

MetMAb_WT_VK1CLambda      ASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTEC
MetMAb_VRD1CRD2           AWSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTEC
```

Anti-Axl_HC

```
Anti_Axl_WT_VH3        EVQLVESGGGLVQPGGSLRLSCAASGFSLSGSWIHWVRQAPGKGLEWVGWINPYRGYAYY
Anti_Axl_VRD2          EVQLVESGGGLVQPGGSLRLSCAASGFSLSGSWIHWVRYAPGKGLEWVGWINPYRGYAYY

Anti_Axl_WT_VH3        ADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCAREYSGWGGSSVGYAMDYWGQGTL
Anti_Axl_VRD2          ADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCAREYSGWGGSSVGYAMDYWGQGTL

Anti_Axl_WT_VH3        VTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPA
Anti_Axl_VRD2          VTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPA

Anti_Axl_WT_VH3        VLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSC
Anti_Axl_VRD2          VLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSC
```

Anti-Axl_LC

```
Anti_Axl_WT_VK1CLambda      DIQMTQSPSSLSASVGDRVTITCRASQDVSTAVAWYQQKPGKAPKLLIYS
```


Nature Biotechnology: doi:10.1038/nbt.2797

```
Anti_Axl_VRD2                 DIQMTQSPSSLSASVGDRVTITCRASQDVSTAVAWYQRKPGKAPKLLIYS

Anti_Axl_WT_VK1CLambda        ASFLYSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYTTPPTFGQ
Anti_Axl_VRD2                 ASFLYSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYTTPPTFGQ

Anti_Axl_WT_VK1CLambda        GTKVEIKGQPKAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWK
Anti_Axl_VRD2                 GTKVEIKGQPKAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWK

Anti_Axl_WT_VK1CLambda        ADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEG
Anti_Axl_VRD2                 ADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEG

Anti_Axl_WT_VK1CLambda        STVEKTVAPTEC
Anti_Axl_VRD2                 STVEKTVAPTEC
```

BHA10_HC

```
BHA10_WT_VH1         QVQLVQSGAEVKKPGSSVKVSCKASGYTFTTYYLHWVRQAPGQGLEWMGWIYPGNVHAQY
BHA10_VRD2           QVQLVQSGAEVKKPGSSVKVSCKASGYTFTTYYLHWVRYAPGQGLEWMGWIYPGNVHAQY

BHA10_WT_VH1         NEKFKGRVTITADKSTSTAYMELSSLRSEDTAVYYCARSWEGFPYWGQGTTVTVSSASTK
BHA10_VRD2           NEKFKGRVTITADKSTSTAYMELSSLRSEDTAVYYCARSWEGFPYWGQGTTVTVSSASTK

BHA10_WT_VH1         GPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYS
BHA10_VRD2           GPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYS

BHA10_WT_VH1         LSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSC
BHA10_VRD2           LSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSC
```

BHA10_LC

```
BHA10_WT_VK1CKappa       DIQMTQSPSSLSASVGDRVTITCKASQNVGINVAWYQQKPGKAPKSLISSASYRYSGVPS
BHA10_VRD2               DIQMTQSPSSLSASVGDRVTITCKASQNVGINVAWYQRKPGKAPKSLISSASYRYSGVPS

BHA10_WT_VK1CKappa       RFSGSGSGTDFTLTISSLQPEDFATYFCQQYDTYPFTFGQGTKVEIKRTVAAPSVFIFPP
BHA10_VRD2               RFSGSGSGTDFTLTISSLQPEDFATYFCQQYDTYPFTFGQGTKVEIKRTVAAPSVFIFPP

BHA10_WT_VK1CKappa       SDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLT
BHA10_VRD2               SDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLT

BHA10_WT_VK1CKappa       LSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
BHA10_VRD2               LSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
```

Trastuzumab_HC

```
Trastuzumab_WT_VH3        EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYTRY
Trastuzumab_VRD1CRD2      EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRKAPGKGLEWVARIYPTNGYTRY

Trastuzumab_WT_VH3        ADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGTLVTVSS
Trastuzumab_VRD1CRD2      ADEVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGTLVTVSS

Trastuzumab_WT_VH3        ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSS
Trastuzumab_VRD1CRD2      ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVATGPAVLQSS

Trastuzumab_WT_VH3        GLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSC
Trastuzumab_VRD1CRD2      GLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSC
```

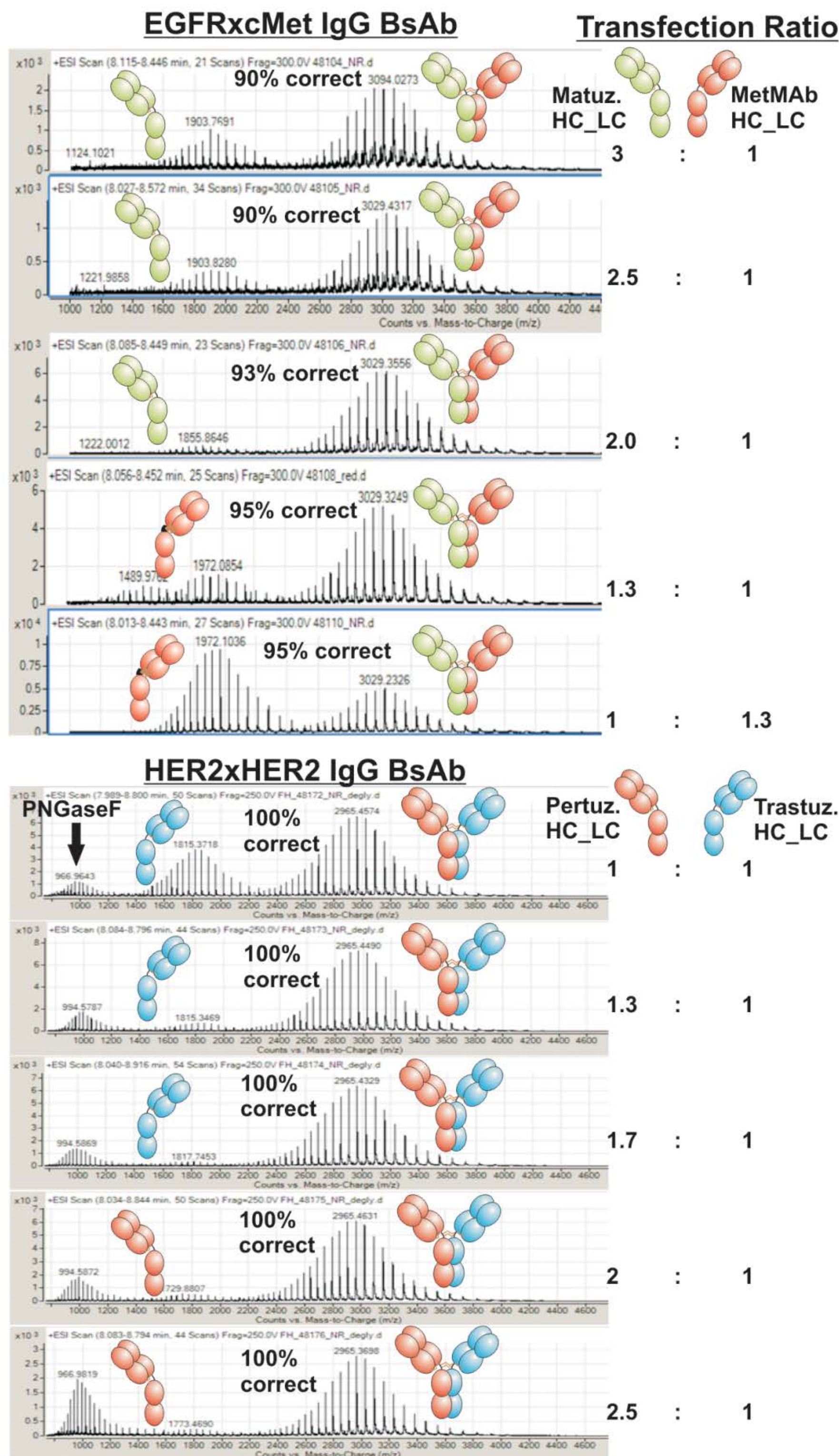
Trastuzumab_LC

```
Trastuzumab_WT_VK1CKappa      DIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYS
Trastuzumab_VRD1CRD2          RIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWFQDKPGKAPKLLIYS

Trastuzumab_WT_VK1CKappa      ASFLYSGVPSRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQ
Trastuzumab_VRD1CRD2          ASFLYSGVPSRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQ
```


Nature Biotechnology: doi:10.1038/nbt.2797

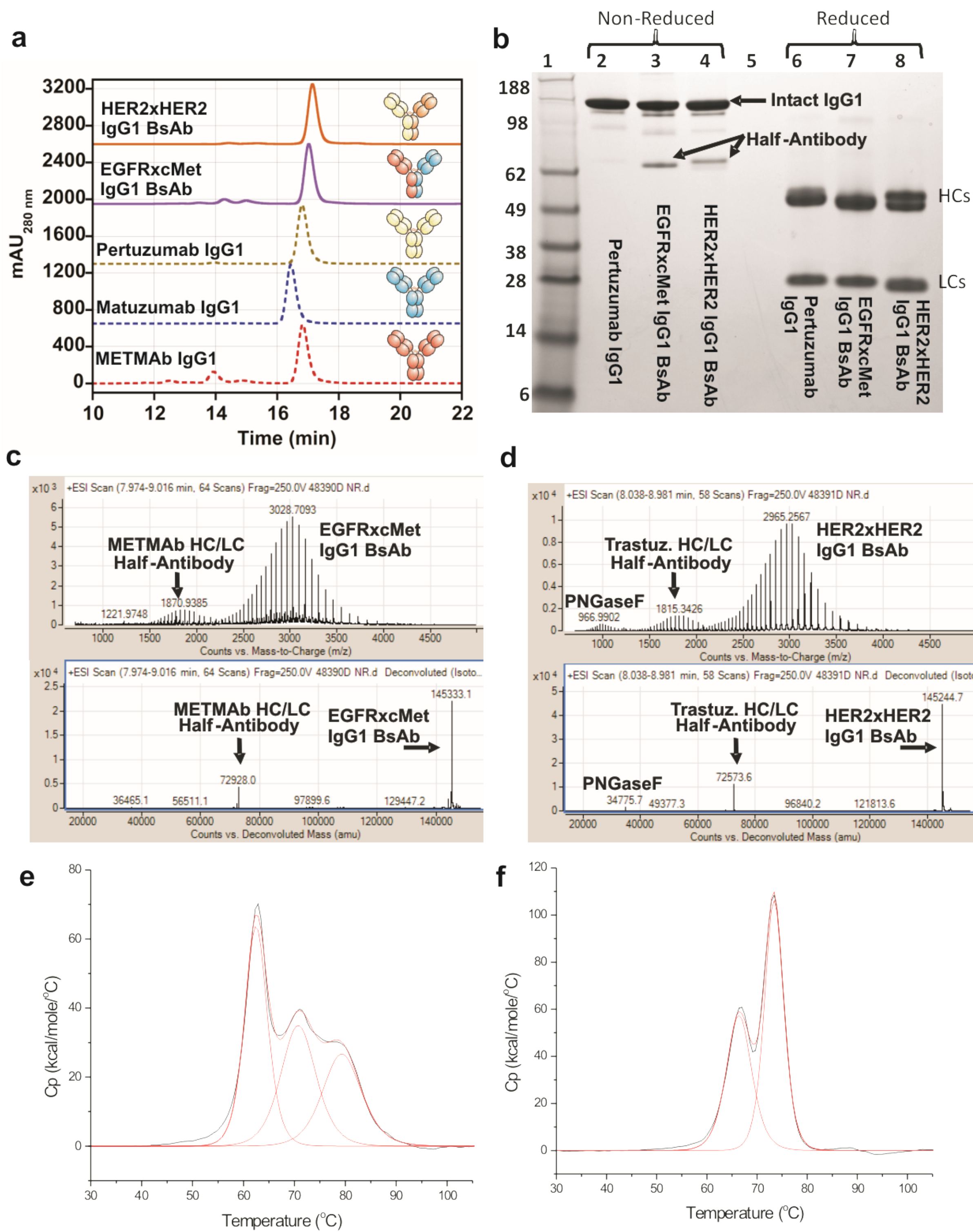
```
Trastuzumab_WT_VK1CKappa      GTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKV
Trastuzumab_VRD1CRD2          GTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCYLNNFYPREAKVQWKV

Trastuzumab_WT_VK1CKappa      DNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQG
Trastuzumab_VRD1CRD2          DNALQSGNSQESVTEQDSKDSTYSLWSTLTLSKADYEKHKVYACEVTHQG

Trastuzumab_WT_VK1CKappa      LSSPVTKSFNRGEC
Trastuzumab_VRD1CRD2          LSSPVTKSFNRGEC
```


Nature Biotechnology: doi:10.1038/nbt.2797



Supplemental Figure 7 Non-deconvoluted mass mpectra of two IgG BsAbs, EGFRxcMET (top half) and (HER2^{P}xHER2^{T} (bottom half), expressed while varying transfection ratios of the HC_LC pairs within each BsAb. The transfections were performed in HEK293F cells using four separate expression plasmids, one expressing each HC and LC within the BsAb. The HC:LC ratio of each half of the IgG BsAb was maintained at a constant 1:2 transfection ratio. Only the ratio of each HC_LC pair was varied. Overexpression of either half of the IgG BsAb led to the accumulation of "Half-Antibody" (a HC_LC pair without a disulfide bonded partner) that was visible at a lower molecular weight in the Mass Spectra as well as in SDS-PAGE analyses. Transfection ratios could be found for minimizing the level of Half-Antibody for all the BsAbs that were tested. The level of correctly assembled IgG BsAb is shown to the left of each IgG BsAb mass envelopes. Interestingly, varying the transfection ratios the HC_LC pairs did not impact significantly the level of correctly assembled IgG BsAb.

Nature Biotechnology: doi:10.1038/nbt.2797

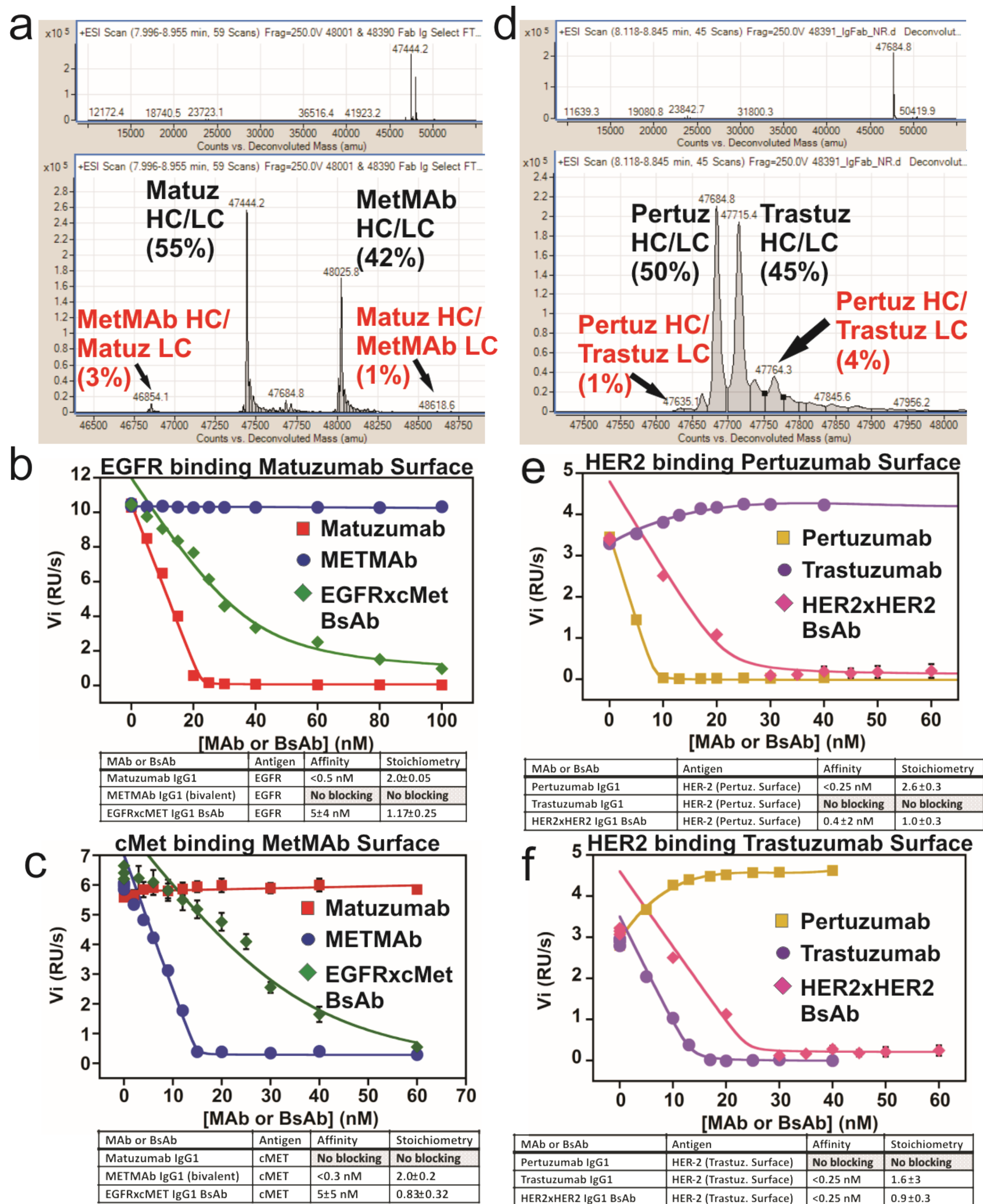


Supplementary Figure 8 Characterization of the EGFRxcMet and HER2^{P}xHER2^{T} IgG BsAbs from 1 L cultures. Five days after transient transfection in HEK293F cells with the BsAbs, the cells were spun down and filtered. BsAbs proteins in the cell supernatants were captured onto a MAbSelect (protein A, GE Healthcare) affinity column using an AKTA Explorer FPLC, washed

Nature Biotechnology: doi:10.1038/nbt.2797

extensively, and eluted using 0.1 M glycine pH 3.0. The proteins were neutralized using 1 M TrisHCl, pH 8.5 and dialyzed exhaustively against PBS. Panel (a) represents analytical SEC traces of parental monovalent IgG1 MAbs and the EGFR×cMet and $HER2^P \times HER2^T$ IgG1 BsAbs. The data indicates that both BsAbs produced at the 1 L scale are predominately monomeric similar to the parental MAbs. The two BsAbs were concentrated to 20 or 30 mg/mL, respectively, with no significant increases in aggregation or loss of protein using VIVASPIN500, 10 K MWCO filtration devices (Sartorius). Panel (b) shows SDS-PAGE analysis of the IgG BsAbs under non-reducing and reducing (1 mM DTT) conditions. For the analysis, proteins (3-5 µg each) were run onto a NuPAGE 4-12% Bis-Tris gel with MES running buffer (Life Technologies) according to manufacturer protocols. Panels (c) and (d) are the raw (top) and deconvoluted (bottom) mass spectra of the EGFR×cMet and $HER2^P \times HER2^T$ IgG1 BsAbs. Proteins were analyzed as described in the methods. The integrated area of the 'Half-Antibody' was 8% and 11%, respectively, for each sample. The average level of correct IgG BsAb assembly (i.e., correct HC/LC pairings) is provided in **Table 1**. Panels (e) and (f) are DSC thermograms of the EGFR×cMet and $HER2^P \times HER2^T$ IgG1 BsAbs, respectively. For the EGFR×cMet BsAb (pairing Matuzumab and MetMAb), combining Vκ1 domains with Cλ domains resulted in reduced stability of the Fab peak (T_m~62 °C). The $HER2^P \times HER2^T$ BsAb (Pertuzumab and Trastuzumab) was produced using fully kappa LCs.

Nature Biotechnology: doi:10.1038/nbt.2797

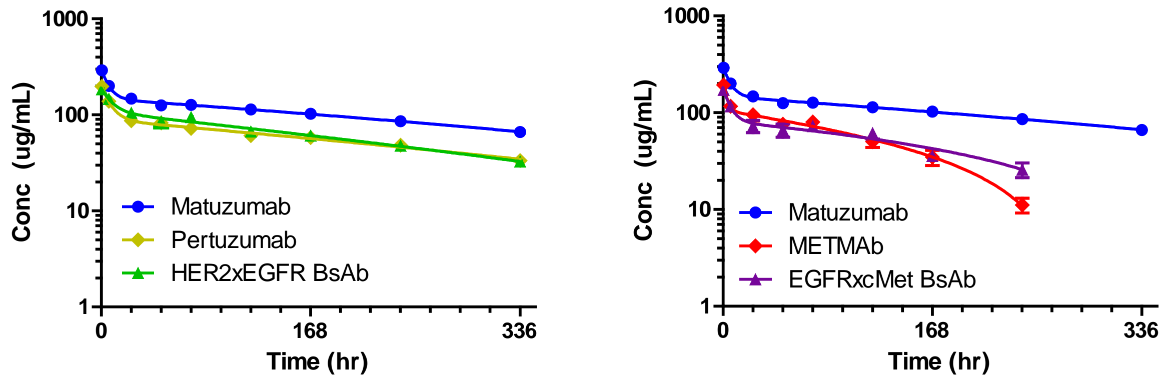


Supplementary Figure 9 Further characterization of the assembly of the EGFRxcMet (*a-c*) and $HER2^{P}\times HER2^{T}$ (*d-f*) IgG1 BsAbs. Panels *(a)* and *(d)* are deconvoluted mass spectra of the Fabs liberated from the EGFR×cMet and $HER2^{P}\times HER2^{T}$ IgG1 BsAbs, respectively, by papain digestion. For the digestions, 5 mg of each BsAb was dialyzed into 100 mM Tris, 2 mM EDTA pH 7.0 (no salt) and 20 mg papain (Roche cat#10814) was added. The reaction was allowed to

Nature Biotechnology: doi:10.1038/nbt.2797

continue for 1.5 hrs at room temperature with gentle rotary mixing. The reaction solution was then passed over an IgSelect column (GE Healthcare) that selectively binds only the Fc fragment. The papain liberated Fabs were passed through the column and were collected, while undigested and Fc-containing material bound the solid support. The collected solutions were immediately analyzed by mass spectrometry as described in the methods. Papain selectively cuts the hinge region of IgG1 MAbs (and BsAbs) between the His and Thr residues (…CDKTH_cut_TCPPC…) enabling theoretical mass calculation of both correctly and incorrectly assembled HC/LC pairs. The top panels are the entire deconvoluted spectra (10-55 kDa) while the bottom panels are the regions of interest for determining Fab assembly. Panels *(b)*, *(c)*, *(e)*, and *(f)* show solution-based antibody-antigen binding studies using surface plasmon resonance for detection. The method is described in (Day et al., 2005 *Biochemistry* 44, 1919-31). The MAbs and BsAbs are titrated into a solution with constant antigen. All antigens were purchased as Fc-fusion proteins from R&D systems. EGFR-Fc was 30 nM; cMet-Fc was 45 nM; and HER2-Fc was 20 nM. The titrated solutions were passed over sensorchip surfaces containing high levels of each parental MAb. The initial slope (Vi) of antigen association is linearly dependent on the concentration of antigen and can be utilized to determine $[\text{antigen}]_{\text{free}}$. The data obtained is sensitive to both binding affinity and stoichiometry. Results of fitting the data for the affinity and the number of binding arms of each MAb or BsAb are shown in the table below each set of curves. The affinity of some of the molecules was too strong to be measured accurately; however, the X-intercept could be used to determine stoichiometry in all cases. The MAbs demonstrated bivalency toward their respective antigens, while the BsAbs demonstrated monovalency.

Nature Biotechnology: doi:10.1038/nbt.2797



Supplemental Figure 10 In vivo pharmacokinetic analysis of the HER2×EGFR and EGFR×cMet IgG BsAbs in Balb/c mice. Plotted values represent mean +/- standard error (n= 3 or 4); error bars are plotted for all data points.

mAb or BsAb	Elimination half-life $t_{1/2}\,\beta$ (hr)	AUC (hr*ug/mL)	Clearance (mL/hr)	Vss (mL)
Pertuzumab	241.6 ± 15.5	32,374 ± 1,315	0.007 ± 0.0003	2.36 ± 0.07
Matuzumab	294.8 ± 15.3	64,215 ± 2,315	0.0035 ± 0.0001	1.47 ± 0.03
MetMab	82.5*	15,977*	0.014*	1.72*
HER2×EGFR BsAb	197.3 ± 15.2	31,429 ± 1,390	0.007 ± 0.0003	1.99 ± 0.09
cMet×EGFR BsAb	137.6 ± 21.2	18,224 ± 1,613	0.012 ± 0.001	2.39 ± 0.23

Pharmacokinetic parameters are shown for MAbs and BsAbs. Data fitted to the IV bolus model represent estimated mean ± standard error. *Data fitted by non-compartmental analysis represent estimated mean.

Nature Biotechnology: doi:10.1038/nbt.2797

Supplementary Protocol 1. PREPACK. This is a Rosetta3 command line which will perform the prepacking rotamer relaxation used in this work. This is expected to take a few minutes on a single processor.

Command line:

```
/path/to/rosetta/source/bin/fixbb.<os><mode>release @options
```

The options file consists of:

```
#begin options
#database path
-database /path/to/rosetta/database
#input path
-s input.pdb
#use extra rotamers
-ex1
-ex2
-ex_cutoff 0
#use the crystal rotamer if it is superior
-use_input_sc
#do not design
-packing:repack_only
#perform 10 repacks but only output the best
-ndruns 10
#end options
```


Nature Biotechnology: doi:10.1038/nbt.2797

Supplementary Protocol 2. Multistate design script. Specifications necessary to run heterodimer-to-orthogonal-heterodimer multistate design calculations via Rosetta's mpi_msd application. This example is drawn from the experiment used to create design CRD1Î². Extensive reference to Rosetta's mpi_msd documentation and examples published with the code is strongly suggested. Use in MPI is strongly suggested. The number of processors should be an even divisor of (preferably equal to) the number of states (10 in this example).

```
Command line:
/path/to/rosetta/source/bin/mpi_msd.<os><mode>release @options
```

The options file consists of:

```
#begin options

#mpi_msd options - refer to its documentation
-entity_resfile entity.resfile
-fitness_file fitness.daf
-ms::pop_size 100
-ms::generations 240
-ms::numresults 10
-use_input_sc
-ms::fraction_by_recombination 0.025
#-msd::double_lazy_ig_mem_limit 800

-database /path/to/rosetta/database

#write the rosetta version and options to the log file
-options:user
-run:version

#write each processor to its own file, for organization
-mpi_tracer_to_file proc

#this exclusion reduces memory use but is not important
-chemical:exclude_patches LowerDNA  UpperDNA Cterm_amidation SpecialRotamer
protein_cutpoint_upper protein_cutpoint_lower VirtualBB ShoveBB
VirtualDNAPhosphate VirtualNTerm CTermConnect sc_orbitals
pro_hydroxylated_case1 pro_hydroxylated_case2 ser_phosphorylated
thr_phosphorylated tyr_phosphorylated tyr_sulfated lys_dimethylated
lys_monomethylated lys_trimethylated lys_acetylated glu_carboxylated
cys_acetylated tyr_diiodinated N_acetylated C_methylamidated
MethylatedProteinCterm

#scorefunction corrections, notice these have been superseded in the time
before publication
-corrections::score::score12prime
-no_his_his_pairE

#end options
```

File fitness.daf. This is the main setup file for mpi_msd. The nomenclature "LpHp", with p for "prime", means light chain and heavy chain mutant. Lack of prime means the wild type sequence - the undesired mixed products.

```
#begin fitness.daf
```


Nature Biotechnology: doi:10.1038/nbt.2797

```
#Light chain, mutant, by itself
STATE_VECTOR Lp Lp.states
#Heavy chain, mutant, by itself
STATE_VECTOR Hp Hp.states
#mutant pair
STATE_VECTOR LpHp LpHp.states
#cross pair
STATE_VECTOR LHp LHp.states
#cross pair
STATE_VECTOR LpH LpH.states

#constant states do not need to be modeled
POSE_ENERGY bestL 3TV3_CL.pdb
POSE_ENERGY bestH 3TV3_CH1.pdb
POSE_ENERGY bestLH 3TV3_CH1CL.pdb

# extract the lowest-energy from each vector of state energies
SCALAR_EXPRESSION bestLp = vmin( Lp )
SCALAR_EXPRESSION bestHp = vmin( Hp )
SCALAR_EXPRESSION bestLpHp = vmin( LpHp )
SCALAR_EXPRESSION bestLHp = vmin( LHp )
SCALAR_EXPRESSION bestLpH = vmin( LpH )

# now compute the binding energies
SCALAR_EXPRESSION dGbind_LpHp = bestLpHp - bestLp - bestHp
SCALAR_EXPRESSION dGbind_LHp = bestLHp - bestL - bestHp
SCALAR_EXPRESSION dGbind_LpH = bestLpH - bestLp - bestH
SCALAR_EXPRESSION dGbind_LH = bestLH - bestL - bestH

#-12 is the energy cap, determined empirically as the rough value of redocked
binding energy for the worst mutations
SCALAR_EXPRESSION clipped_dGbindLHp = ite( lt( dGbind_LHp, -12.0 ),
dGbind_LHp, ( -12.0 + 0 * dGbind_LHp) )
SCALAR_EXPRESSION clipped_dGbindLpH = ite( lt( dGbind_LpH, -12.0 ),
dGbind_LpH, ( -12.0 + 0 * dGbind_LpH) )

#this file counts the number of mutations and penalizes for too many
ENTITY_FUNCTION cstE 3TV3.entfunc

#this is the final fitness function
FITNESS 1.0 * bestLpHp - 0 * dGbind_LpHp + 0.5 * ( clipped_dGbindLpH +
clipped_dGbindLHp ) + 1.0 * cstE
#end fitness.daf
```

File 3TV3.entfunc. This file, referenced in fitness.daf, controls for the number of mutations.

```
#begin 3TV3.entfunc
#ee_1 is L-L135
SET_CONDITION ee1nat = ee_1 in { L }

#ee_2 is L-S165
SET_CONDITION ee2nat = ee_2 in { S }

#ee_3 is L-A174
SET_CONDITION ee3nat = ee_3 in { A }

#ee_4 is L-S176
```


Nature Biotechnology: doi:10.1038/nbt.2797

```
SET_CONDITION ee4nat = ee_4 in { S }

#ee_5 is H-H172
SET_CONDITION ee5nat = ee_5 in { H }

#ee_6 is H-F174
SET_CONDITION ee6nat = ee_6 in { F }

#ee_7 is H-P175
SET_CONDITION ee7nat = ee_7 in { P }

#ee_8 is H-S188
SET_CONDITION ee8nat = ee_8 in { S }

#ee_9 is H-V190
SET_CONDITION ee9nat = ee_9 in { V }


#count the number of mutations
SUB_EXPRESSION nnat_L = 0 + ee5nat + ee6nat + ee7nat + ee8nat + ee9nat
SUB_EXPRESSION nnat_H = 0 + ee1nat + ee2nat + ee3nat + ee4nat

#penalize after 3 mutations per chain;
SUB_EXPRESSION mut_H_penalty = ite( lt( nnat_H, 2 ), 2 - nnat_H, 0 )
SUB_EXPRESSION mut_L_penalty = ite( lt( nnat_L, 1 ), 1 - nnat_L, 0 )

SCORE 3 * ( mut_H_penalty + mut_L_penalty )
#end 3TV3.entfunc
```

The states files, specified in fitness.daf, correlate what states exist, which input PDB files provide the backbone for each state, and how to correlate the mutable positions onto each backbone. Each file contains two clones of the same setup, to ensure that rare failures of simulated annealing do not poison the results. These input PDBs are not provided.

```
#begin Hp.states
3TV3_CH1.pdb Hp.corr Hx.2resfile
3TV3_CH1.pdb Hp.corr Hx.2resfile
#end Hp.states

#begin LHp.states
3TV3_CH1CL.pdb LHp.corr LxHx.2resfile
3TV3_CH1CL.pdb LHp.corr LxHx.2resfile
#end LHp.states

#begin LpHp.states
3TV3_CH1CL.pdb LpHp.corr LxHx.2resfile
3TV3_CH1CL.pdb LpHp.corr LxHx.2resfile
#end LpHp.states

#begin LpH.states
3TV3_CH1CL.pdb LpH.corr LxHx.2resfile
3TV3_CH1CL.pdb LpH.corr LxHx.2resfile
#end LpH.states

#begin Lp.states
3TV3_CL.pdb Lp.corr Lx.2resfile
```


Nature Biotechnology: doi:10.1038/nbt.2797

```
3TV3_CL.pdb Lp.corr Lx.2resfile
#end Lp.states
```

The corr (correspondence) files determine which positions on the backbone map to the mutable positions defined in entity.resfile. Wild-type positions in the mixed wild-type/mutant and mutant/wild-type structures are unrepresented in these files - that is why they do not mutate.

```
#begin Hp.corr
5  172 H
6  174 H
7  175 H
8  188 H
9  190 H
#end Hp.corr

#begin LHp.corr
5  172 H
6  174 H
7  175 H
8  188 H
9  190 H
#end LHp.corr

#begin Lp.corr
1  135 L
2  165 L
3  174 L
4  176 L
#end Lp.corr

#begin LpH.corr
1  135 L
2  165 L
3  174 L
4  176 L
#end LpH.corr

#begin LpHp.corr
1  135 L
2  165 L
3  174 L
4  176 L
5  172 H
6  174 H
7  175 H
8  188 H
9  190 H
#end LpHp.corr
```

File LxHx.2resfile is a Rosetta resfile that represents a listing of all repackable positions in the desired interface. It contains extra packing instructions encoded in the resfile; these could equivalently be encoded in the options file. Files Lx.2resfile and Hx.2resfile, which are necessary, are not reproduced here for brevity. They represent only the header and L-chain or H-chain portions of LxHx.2resfile.

Nature Biotechnology: doi:10.1038/nbt.2797

```
#begin LxHx.2resfile

NATRO
start
#L
114 L NATAA EX 1 EX 2 USE_INPUT_SC
115 L NATAA EX 1 EX 2 USE_INPUT_SC
116 L NATAA EX 1 EX 2 USE_INPUT_SC
117 L NATAA EX 1 EX 2 USE_INPUT_SC
118 L NATAA EX 1 EX 2 USE_INPUT_SC
119 L NATAA EX 1 EX 2 USE_INPUT_SC
120 L NATAA EX 1 EX 2 USE_INPUT_SC
121 L NATAA EX 1 EX 2 USE_INPUT_SC
122 L NATAA EX 1 EX 2 USE_INPUT_SC
123 L NATAA EX 1 EX 2 USE_INPUT_SC
124 L NATAA EX 1 EX 2 USE_INPUT_SC
125 L NATAA EX 1 EX 2 USE_INPUT_SC
126 L NATAA EX 1 EX 2 USE_INPUT_SC
127 L NATAA EX 1 EX 2 USE_INPUT_SC
129 L NATAA EX 1 EX 2 USE_INPUT_SC
131 L NATAA EX 1 EX 2 USE_INPUT_SC
133 L NATAA EX 1 EX 2 USE_INPUT_SC
134 L NATAA EX 1 EX 2 USE_INPUT_SC
135 L NATAA EX 1 EX 2 USE_INPUT_SC
136 L NATAA EX 1 EX 2 USE_INPUT_SC
137 L NATAA EX 1 EX 2 USE_INPUT_SC
138 L NATAA EX 1 EX 2 USE_INPUT_SC
158 L NATAA EX 1 EX 2 USE_INPUT_SC
160 L NATAA EX 1 EX 2 USE_INPUT_SC
161 L NATAA EX 1 EX 2 USE_INPUT_SC
162 L NATAA EX 1 EX 2 USE_INPUT_SC
163 L NATAA EX 1 EX 2 USE_INPUT_SC
164 L NATAA EX 1 EX 2 USE_INPUT_SC
165 L NATAA EX 1 EX 2 USE_INPUT_SC
166 L NATAA EX 1 EX 2 USE_INPUT_SC
167 L NATAA EX 1 EX 2 USE_INPUT_SC
168 L NATAA EX 1 EX 2 USE_INPUT_SC
170 L NATAA EX 1 EX 2 USE_INPUT_SC
173 L NATAA EX 1 EX 2 USE_INPUT_SC
174 L NATAA EX 1 EX 2 USE_INPUT_SC
175 L NATAA EX 1 EX 2 USE_INPUT_SC
176 L NATAA EX 1 EX 2 USE_INPUT_SC
177 L NATAA EX 1 EX 2 USE_INPUT_SC
178 L NATAA EX 1 EX 2 USE_INPUT_SC
180 L NATAA EX 1 EX 2 USE_INPUT_SC

#H
120 H NATAA EX 1 EX 2 USE_INPUT_SC
121 H NATAA EX 1 EX 2 USE_INPUT_SC
122 H NATAA EX 1 EX 2 USE_INPUT_SC
123 H NATAA EX 1 EX 2 USE_INPUT_SC
124 H NATAA EX 1 EX 2 USE_INPUT_SC
125 H NATAA EX 1 EX 2 USE_INPUT_SC
126 H NATAA EX 1 EX 2 USE_INPUT_SC
134 H NATAA EX 1 EX 2 USE_INPUT_SC
137 H NATAA EX 1 EX 2 USE_INPUT_SC
138 H NATAA EX 1 EX 2 USE_INPUT_SC
```


Nature Biotechnology: doi:10.1038/nbt.2797

```
139 H NATAA EX 1 EX 2 USE_INPUT_SC
140 H NATAA EX 1 EX 2 USE_INPUT_SC
141 H NATAA EX 1 EX 2 USE_INPUT_SC
142 H NATAA EX 1 EX 2 USE_INPUT_SC
143 H NATAA EX 1 EX 2 USE_INPUT_SC
145 H NATAA EX 1 EX 2 USE_INPUT_SC
146 H NATAA EX 1 EX 2 USE_INPUT_SC
167 H NATAA EX 1 EX 2 USE_INPUT_SC
168 H NATAA EX 1 EX 2 USE_INPUT_SC
169 H NATAA EX 1 EX 2 USE_INPUT_SC
170 H NATAA EX 1 EX 2 USE_INPUT_SC
172 H NATAA EX 1 EX 2 USE_INPUT_SC
173 H NATAA EX 1 EX 2 USE_INPUT_SC
174 H NATAA EX 1 EX 2 USE_INPUT_SC
175 H NATAA EX 1 EX 2 USE_INPUT_SC
176 H NATAA EX 1 EX 2 USE_INPUT_SC
177 H NATAA EX 1 EX 2 USE_INPUT_SC
178 H NATAA EX 1 EX 2 USE_INPUT_SC
179 H NATAA EX 1 EX 2 USE_INPUT_SC
180 H NATAA EX 1 EX 2 USE_INPUT_SC
181 H NATAA EX 1 EX 2 USE_INPUT_SC
186 H NATAA EX 1 EX 2 USE_INPUT_SC
187 H NATAA EX 1 EX 2 USE_INPUT_SC
188 H NATAA EX 1 EX 2 USE_INPUT_SC
189 H NATAA EX 1 EX 2 USE_INPUT_SC
190 H NATAA EX 1 EX 2 USE_INPUT_SC
192 H NATAA EX 1 EX 2 USE_INPUT_SC
221 H NATAA EX 1 EX 2 USE_INPUT_SC
225 H NATAA EX 1 EX 2 USE_INPUT_SC
228 H NATAA EX 1 EX 2 USE_INPUT_SC
#end LxHx.2resfile
```

entity.resfile is a multistate design resfile which controls how many positions are mutable, and how they can be mutated.

```
#begin entity.resfile
9
ALLAAwc EX 1 EX 2 USE_INPUT_SC
start
#no body to the file - all 9 positions obey the header
#end entity.resfile
```


Nature Biotechnology: doi:10.1038/nbt.2797

Supplementary Protocol 3. Docking protocol. Individual docking trajectories are very fast (order of 30s per trajectory) and are independent, so many-processor MPI is suggested.

Rosetta command line:

```
/path/to/rosetta/source/bin/rosetta_scripts.<os><mode>release @options
```

The options file consists of:

```
#begin options
#database path
-database /path/to/rosetta/database
#input path; pdblist is an endline-delimited list of paths to pdb files (one PDB per line)
-l pdblist

#perform 2 repacks but only output the best each cycle
-ndruns 2

#allow 100 attempts to pass docking filters
-jd2:ntrials 100

#perform 750 trajectories per input
-nstruct 750

#docking settings
-docking:dock_pert 2 4

#minimization settings
-run::min_type dfpmin_armijo_nonmonotone

#packing settings
-ex1
-ex2
-use_input_sc

#scoring corrections, probably superseded by the time of publication
-no_his_his_pairE
-corrections::score::score12prime

#write the rosetta version and options to the log file
-options:user
-run:version

#this exclusion reduces memory use but is not important
-chemical:exclude_patches LowerDNA  UpperDNA Cterm_amidation SpecialRotamer protein_cutpoint_upper protein_cutpoint_lower VirtualBB ShoveBB VirtualDNAPhosphate VirtualNTerm CTermConnect sc_orbitals pro_hydroxylated_case1 pro_hydroxylated_case2 ser_phosphorylated thr_phosphorylated tyr_phosphorylated tyr_sulfated lys_dimethylated lys_monomethylated lys_trimethylated lys_acetylated glu_carboxylated cys_acetylated tyr_diiodinated N_acetylated C_methylamidated MethylatedProteinCterm

#output controls
-out:file:silent docking_lrf.out
```

```
-out:file:silent_struct_type binary

#quiet mode
-mute all

#script to define docking
-parser:protocol lrf_docking.xml
#end options
```

The lrf_docking.xml file controls the docking protocol.

```
#begin lrf_docking.xml
<ROSETTASCRIPTS>
        <SCOREFXNS>
        <s12_prime weights="score12prime"/>
        </SCOREFXNS>
      <TASKOPERATIONS>
        <RestrictToInterfaceVector name=vectorTask chain1_num=1 chain2_num=2
CB_dist_cutoff=10.0 nearby_atom_cutoff=5.5 vector_angle_cutoff=65.0
vector_dist_cutoff=8.0/>
        <RestrictChainToRepacking name=repack1 chain=1/>
        <RestrictChainToRepacking name=repack2 chain=2/>
        <RestrictToRepacking name=repackonly/>
        <InitializeFromCommandline name=cmdTask/>
        <IncludeCurrent name=currentTask/>
        </TASKOPERATIONS>
        <FILTERS>
        <Rmsd name=rms threshold=50.0 confidence=1/>
      </FILTERS>
        <MOVERS>
        <Prepack name=prepack scorefxn=s12_prime jump_number=1 min_bb=0
task_operations=vectorTask,repackonly,cmdTask,currentTask/>
        <DockingProtocol name=dockprotocol docking_local_refine=1
docking_score_high=s12_prime ignore_default_docking_task=1
task_operations=vectorTask,repack1,repack2,cmdTask,currentTask/>
        <TaskAwareMinMover name=minmover scorefxn=s12_prime chi=1 bb=1 jump=1
task_operations=vectorTask,cmdTask,currentTask/>
        <InterfaceAnalyzerMover name=fullanalyze scorefxn=score12 packstat=0
pack_input=0 jump=1 tracer=0 use_jobname=1 resfile=0 />
        </MOVERS>
        <APPLY_TO_POSE>
        </APPLY_TO_POSE>
        <PROTOCOLS>
        <Add mover_name=prepack/>
        <Add mover_name=dockprotocol filter_name=rms/>
        <Add mover_name=minmover filter_name=rms/>
        <Add mover_name=fullanalyze/>
        </PROTOCOLS>
</ROSETTASCRIPTS>
#end lrf_docking.xml
```


Nature Biotechnology: doi:10.1038/nbt.2797