Supplementary Information

Net charge of antibody complementarity-determining regions is a key predictor of specificity

Lilia A. Rabia, Yulei Zhang, Seth D. Ludwig, Mark C. Julian, Peter M. Tessier

Table S1. Impact of CDR net charge on individual biophysical properties of clinical-stage antibodies. Logistic regression analysis of individual biophysical properties for 137 clinical-stage mAbs as a function of CDR net charge (pH 7.4). Each mAb was evaluated using previously reported thresholds for each biophysical assay (PSR >0.27, AC-SINS >11.8 nm, CSI >0.01 RU, CIC >10.1 min, HIC >11.7 min, SMAC >12.8 min, SGAC-SINS <370 mM, BVP >4.3 signal/background, ELISA >1.9 signal/background and AS >0.08% monomer loss/day) (Jain *et al.*, 2017). The *p*-value for the coefficient of the independent variable of the logistic function and the area under the curve (AUC) for the receiver operating characteristic curve are reported. A negative relationship with CDR net charge means that the biophysical property is worsened as CDR charge is increased, and vice versa for the positive relationship.

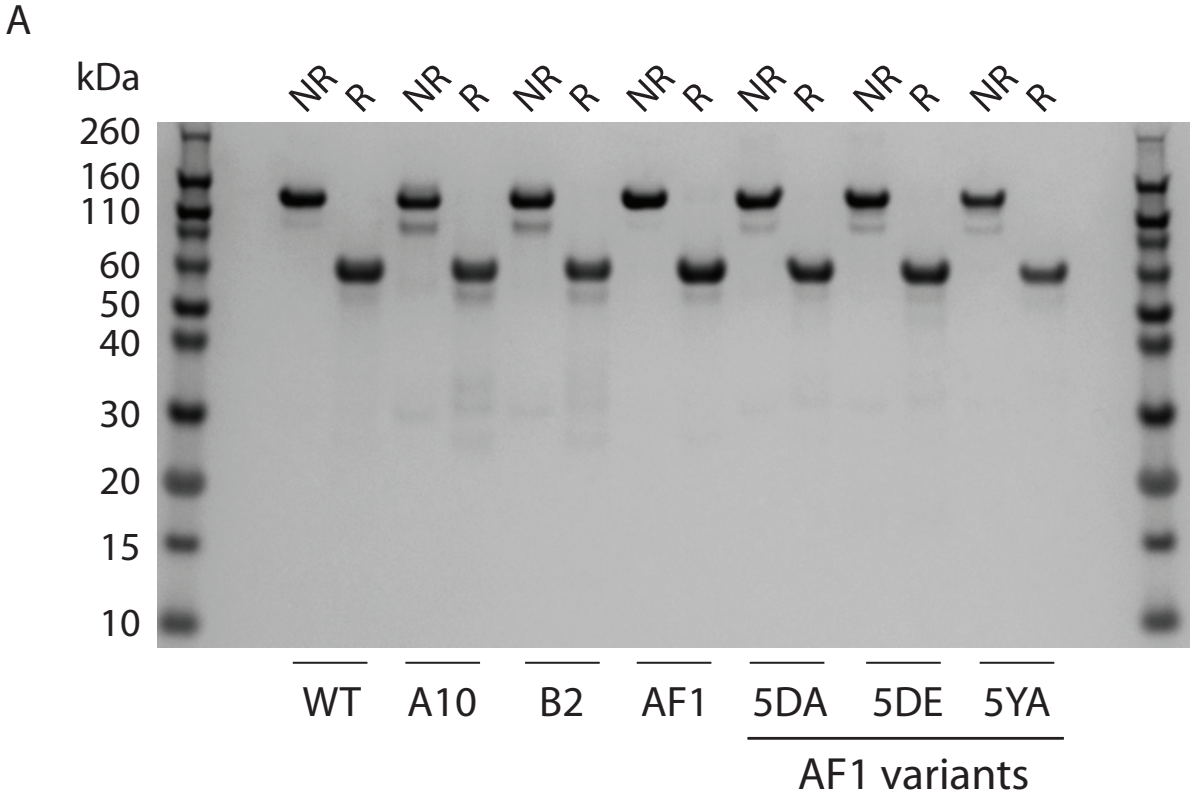
Biophysical flag		***p*-value**	**Area under curve (AUC)**	**Direction of relationship, CDR charge vs. %mAbs without flag**
Specificity	ELISA	0.00025	0.74	-
	BVP	0.00043	0.72	-
	PSR	0.00032	0.74	-
	CIC	0.094	0.59	-
Self-association	AC-SINS	0.0095	0.67	-
	CSI	0.010	0.66	-
Hydrophobicity	HIC	0.46	0.60	+
	SMAC	0.59	0.58	+
	SGAC-SINS	0.090	0.61	-
Aggregation	AS	0.34	0.57	-

Table S2. Increasing CDR net charge is associated with increased risk of poor overall biophysical properties for clinical-stage antibodies. Logistic regression analysis of the relationship between CDR net charge and the biophysical properties of 137 clinical-stage mAbs. Each mAb was assigned up to ten biophysical flags (PSR > 0.27, AC-SINS > 11.8 nm, CSI > 0.01 RU, CIC > 10.1 min, HIC > 11.7 min, SMAC > 12.8 min, SGAC-SINS < 370 mM, BVP > 4.3 signal/background, ELISA > 1.9 signal/background and AS > 0.08% monomer loss/day) (Jain *et al.*, 2017). Logistic regression was performed for the relationship between CDR net charge (pH 7.4) and <2 biophysical flags. The combinations of biophysical flags that were best correlated with CDR net charge are shown. The *p*-value for the coefficient of the independent variable of the logistic function and the area under the curve (AUC) for the receiver operating characteristic curve are reported**.**

Number of flags	Biophysical flags	*p*-value	Area under the curve (AUC)
4	CIC, CSI, ELISA, BVP	5.76E-05	0.76
6	AS, PSR, AC-SINS, CIC, ELISA, BVP	6.85E-05	0.73
4	AS, PSR, ELISA, BVP	7.11E-05	0.76
5	AS, PSR, CIC, ELISA, BVP	8.53E-05	0.74
5	AS, PSR, CSI, ELISA, BVP	9.58E-05	0.74
2	ELISA, BVP	9.62E-05	0.78
6	SGAC-SINS, AS, PSR, CIC, ELISA, BVP	9.87E-05	0.72
5	SGAC-SINS, PSR, CIC, ELISA, BVP	1.01E-04	0.74
5	PSR, AC-SINS, CIC, ELISA, BVP	1.01E-04	0.74
6	AS, PSR, CIC, CSI, ELISA, BVP	1.03E-04	0.72
5	PSR, CIC, CSI, ELISA, BVP	1.03E-04	0.74
3	SGAC-SINS, ELISA, BVP	1.05E-04	0.76
3	AC-SINS, ELISA, BVP	1.05E-04	0.76
4	AS, PSR, CSI, ELISA	1.07E-04	0.75
3	CIC, ELISA, BVP	1.07E-04	0.76
3	CSI, ELISA, BVP	1.08E-04	0.76
5	SMAC, AS, PSR, ELISA, BVP	1.09E-04	0.73
7	SGAC-SINS, AS, PSR, CIC, CSI, ELISA, BVP	1.15E-04	0.71
7	AS, PSR, AC-SINS, CIC, CSI, ELISA, BVP	1.15E-04	0.71
8	SGAC-SINS, AS, PSR, AC-SINS, CIC, CSI, ELISA, BVP	1.15E-04	0.71
4	SGAC-SINS, CIC, ELISA, BVP	1.19E-04	0.74
4	AC-SINS, CIC, ELISA, BVP	1.19E-04	0.74
3	SMAC, ELISA, BVP	1.24E-04	0.77
4	AS, PSR, CSI, BVP	1.25E-04	0.74
5	AS, PSR, AC-SINS, ELISA, BVP	1.28E-04	0.73
4	PSR, CSI, ELISA, BVP	1.34E-04	0.74
4	SMAC, PSR, ELISA, BVP	1.36E-04	0.75
6	PSR, AC-SINS, CIC, CSI, ELISA, BVP	1.41E-04	0.72
3	PSR, ELISA, BVP	1.42E-04	0.75
3	AS, ELISA, BVP	1.45E-04	0.75
5	AS, PSR, CIC, CSI, BVP	1.46E-04	0.72
4	HIC, AC-SINS, ELISA, BVP	1.48E-04	0.74
5	AC-SINS, CIC, CSI, ELISA, BVP	1.52E-04	0.72

Table S3. Effect of charged CDR residues on the overall biophysical properties of clinical-stage antibodies. Logistic regression analysis was performed in a similar manner as in Table S2 for the relationship between the number of single or combined charged residues and <2 biophysical flags. The mAbs were assigned either up to ten biophysical flags (PSR > 0.27, AC-SINS > 11.8 nm, CSI > 0.01 RU, CIC > 10.1 min, HIC > 11.7 min, SMAC > 12.8 min, SGAC-SINS < 370 mM, BVP > 4.3 signal/background, ELISA > 1.9 signal/background and AS > 0.08% monomer loss/day) or up to four biophysical flags (CSI > 0.01 RU, CIC > 10.1 min, BVP > 4.3 signal/background and ELISA > 1.9 signal/background) (Jain *et al.*, 2017). The *p*-value for the coefficient of the independent variable of the logistic function and the area under the curve (AUC) for the receiver operating characteristic curve are reported.

		Ten biophysical flags			**Four biophysical flags**		
Amino acids	**Average fraction** (± standard deviation)	***p*-value**	**Area under curve (AUC)**	**Direction of relationship, # of amino acids vs. favorable biophysical properties**	***p*-value**	**Area under curve (AUC)**	**Direction of relationship, # of amino acids vs. favorable biophysical properties**
R	0.039 ± 0.020	0.056	0.58	-	0.0010	0.68	-
K	0.031 ± 0.018	0.91	0.52	-	0.40	0.55	-
H	0.024 ± 0.018	0.054	0.58	-	0.097	0.56	-
R, K	0.070 ± 0.026	0.11	0.56	-	0.0016	0.67	-
R, K, H	0.095 ± 0.034	0.028	0.60	-	0.0013	0.68	-
D	0.061 ± 0.026	0.034	0.60	+	0.029	0.62	+
E	0.018 ± 0.020	0.040	0.62	+	0.061	0.64	+
D, E	0.079 ± 0.031	0.0026	0.67	+	0.0031	0.71	+



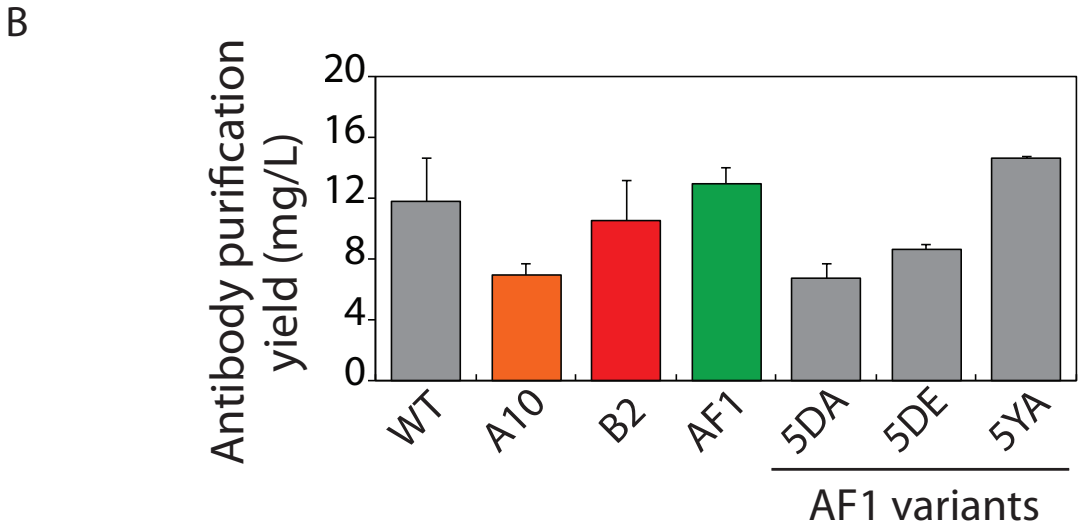


Fig. S1. Purity and purification yields of scFv-Fc antibodies. (A) SDS-PAGE analysis of scFv-Fc antibodies after mammalian expression and purification. Samples were either both reduced and boiled (R) or untreated (NR) prior to loading. (B) Purification yields of scFv-Fc antibodies. The yield was calculated as the mass of purified protein divided by the volume of the expression culture. An average of two replicates is shown, and the error bars are standard errors.

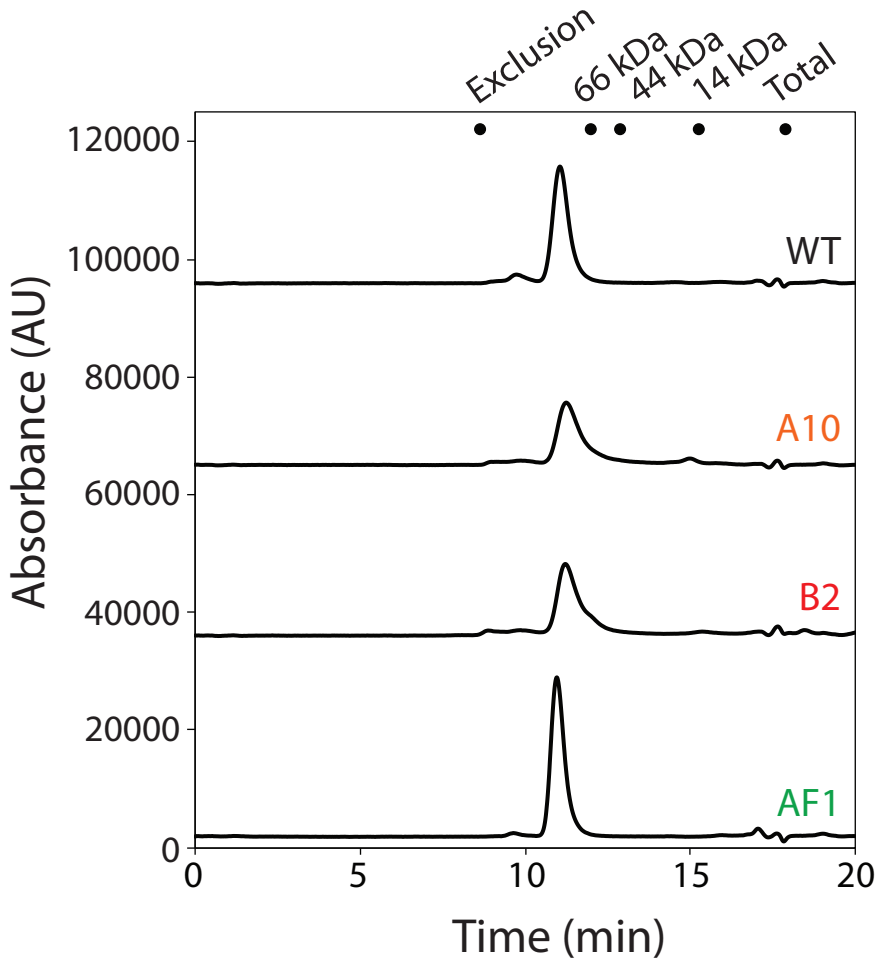


Fig. S2. Size-exclusion chromatography analysis of scFv-Fc antibodies. The antibodies (6 µg) were evaluated in PBS supplemented with 0.2 M arginine (pH 7.4).

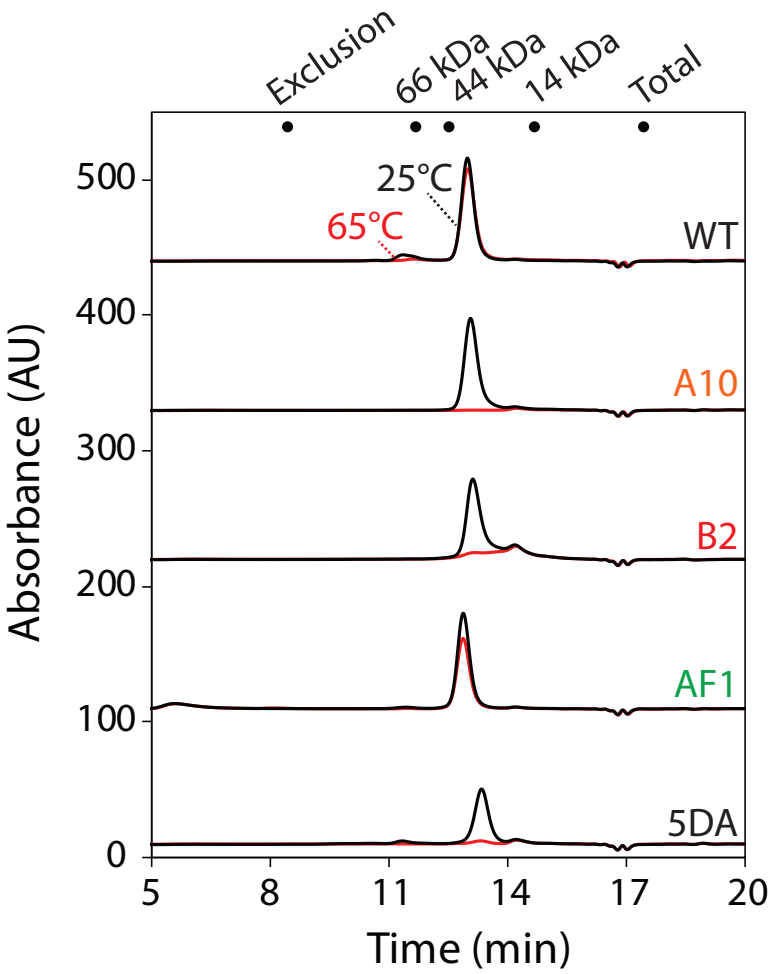


Fig. S3. Size-exclusion chromatography analysis of single-chain antibodies. The scFv antibodies were compared before and after being heated at 65 °C for 4 h (concentration of 0.25 mg/mL) followed by incubation at 4 °C overnight. The running buffer was PBS supplemented with 0.2 M arginine at pH 7.4. As part of the purification, the antibodies were first heated at 55 °C for 1 h to remove aggregate. A representative experiment of three independent experiments is shown.

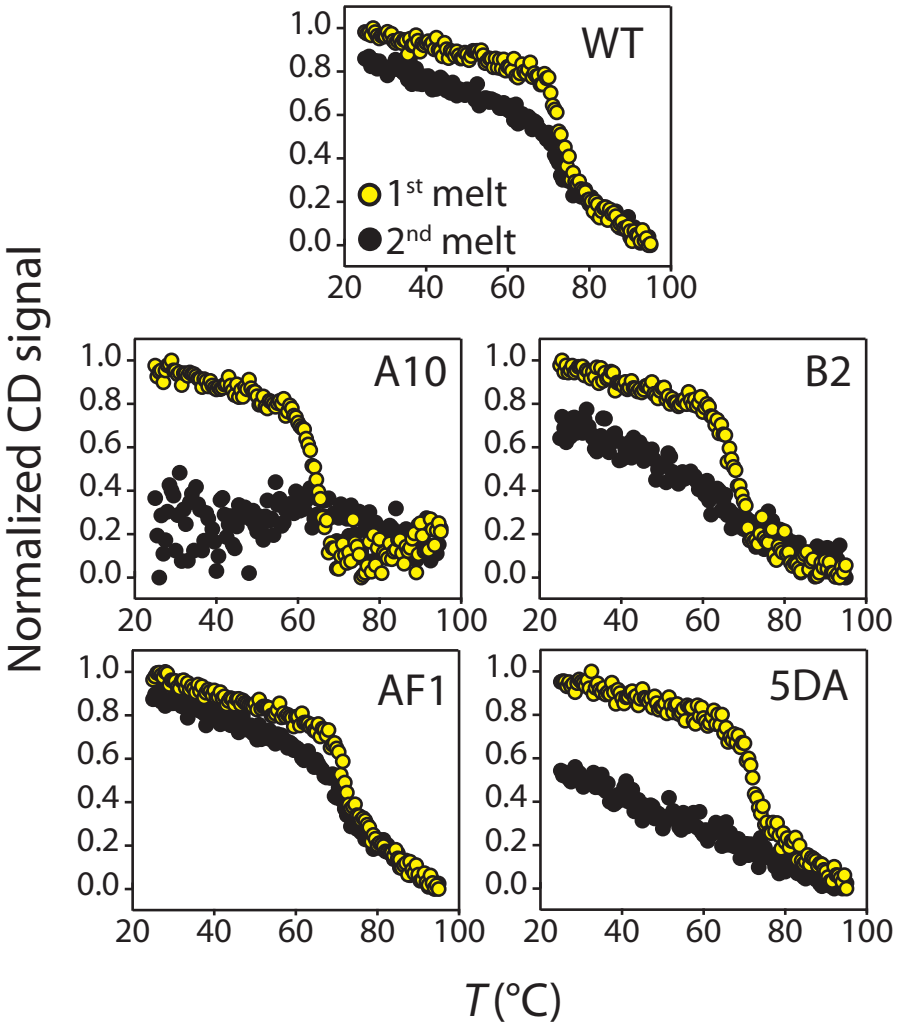


Fig. S4. Analysis of the reversibility of unfolding for single-chain antibodies with high and low specificity. Ellipticity at 235 nm was monitored while heating the scFv antibodies from 25 to 95 °C. After the first melt (yellow), the antibodies were cooled to 25 °C for 10 min before heating again (black). Two independent experiments were performed for each antibody, and representative melt curves are shown.