

Stabilizing the C_H2 domain of an Antibody by Engineering in an Enhanced Aromatic Sequon

Wentao Chen^{†,‡}, Leopold Kong[§], Stephen Connelly[§], Julia M. Dendle^{†,‡}, Yu Liu^{†,‡}, Ian A. Wilson^{§,#}, Evan T. Powers^{‡,*}, Jeffery W. Kelly^{†,‡,#,*}

[†]Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, CA 92037, USA.

[‡]Department of Chemistry, The Scripps Research Institute, La Jolla, CA 92037, USA.

[§]Department of Integrative, Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, USA.

[#]The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037, USA.

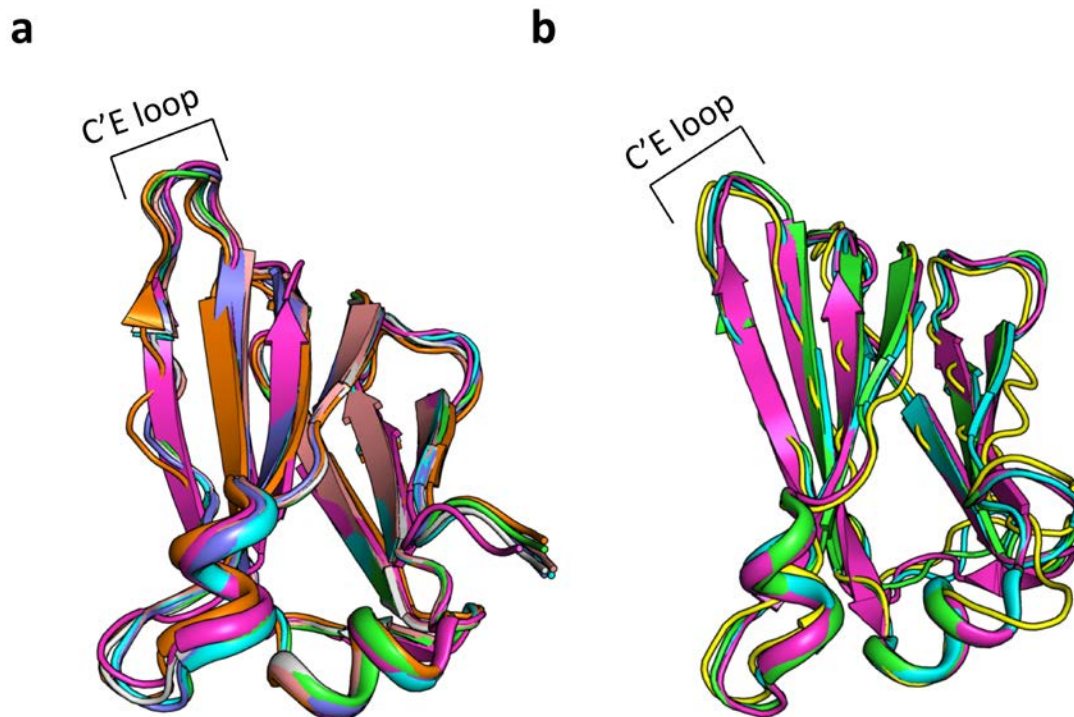


Figure S1. Alignment of the C_H2 domains of the other Fc structures. Structures are shown in ribbon format. Only the protein backbone atoms are used in the structural alignment. (a) C_H2 domains in PDBID 2WAH, 4CDH, 1HZH, 3AVE, 3SGJ, 4KU1 and 4X4M are shown in green, cyan, magenta, terra cotta, grey, slate blue and orange, respectively. The C'E loops have type I β -turns. (b) Alignment of the C_H2 domains from structures of the Fc fragment of allotype G1m3 (Cri). C_H2 domains in PDBID 1H3T, 1E4K, 1FC1 and 3RY6 are shown in green, cyan, magenta and yellow, respectively. The C'E loops have type IV β -turns.

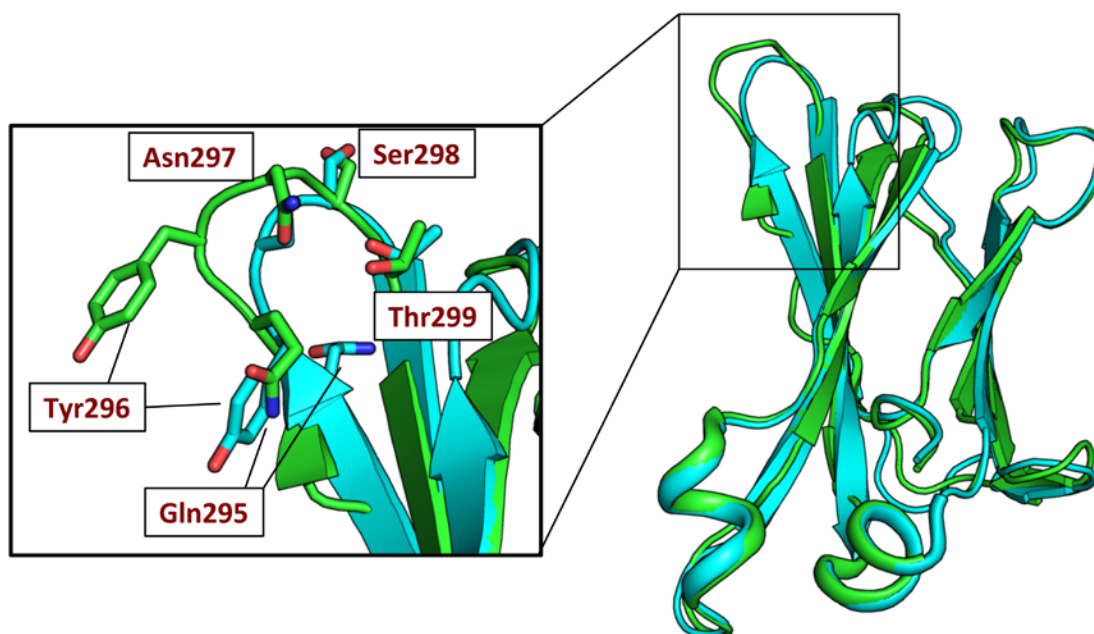


Figure S2. Alignment of the C_H2 domains of Fc fragment structures with turn structures. Only the protein backbone atoms from 3AVE (green, type I β-turn) and 1FC1 (cyan, type IV β-turn) are used in the structural alignment. The side chains of residues Gln295-Thr299 are shown in stick format.

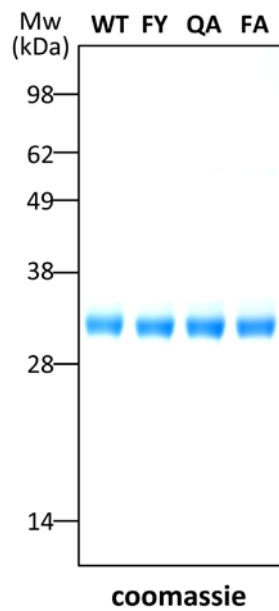


Figure S3. Coomassie stained SDS-PAGE of purified Fc fragment variants.

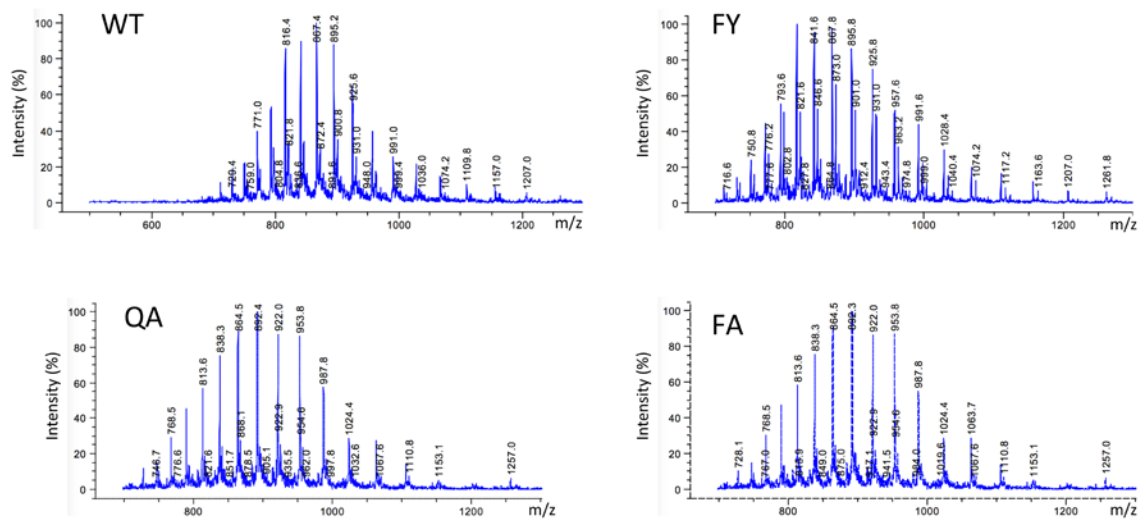


Figure S4. ESI-MS data of dithiothreitol-treated Fc fragment variants.











	Deconvoluted Mw	Observed Mw	Relative abundance	N-glycan
WT	27717	27724	1.00	 G0F
	27879	27883	0.41	 G1F
	27838	27839	0.19	 hybrid
FY	27736	27739	1.00	 G0F
	27898	27901	0.61	 G1F
	28060	28062	0.20	 G2F
QA	27625	27630	1.00	 G0F
	27746	27749	0.29	 hybrid
FA	27644	27648	1.00	 G0F
	27765	27766	0.25	 hybrid

Figure S5. Deconvoluted molecular weight (Mw) and relative abundance (normalized to the most abundant species) of ESI-MS data for each dithiothreitol-treated Fc fragment variant. Only those glycoforms that have relative abundance above 0.15 are shown. G# refers to the number of Gal on the two arms and F to fucosylation of the initial GlcNAc residue.

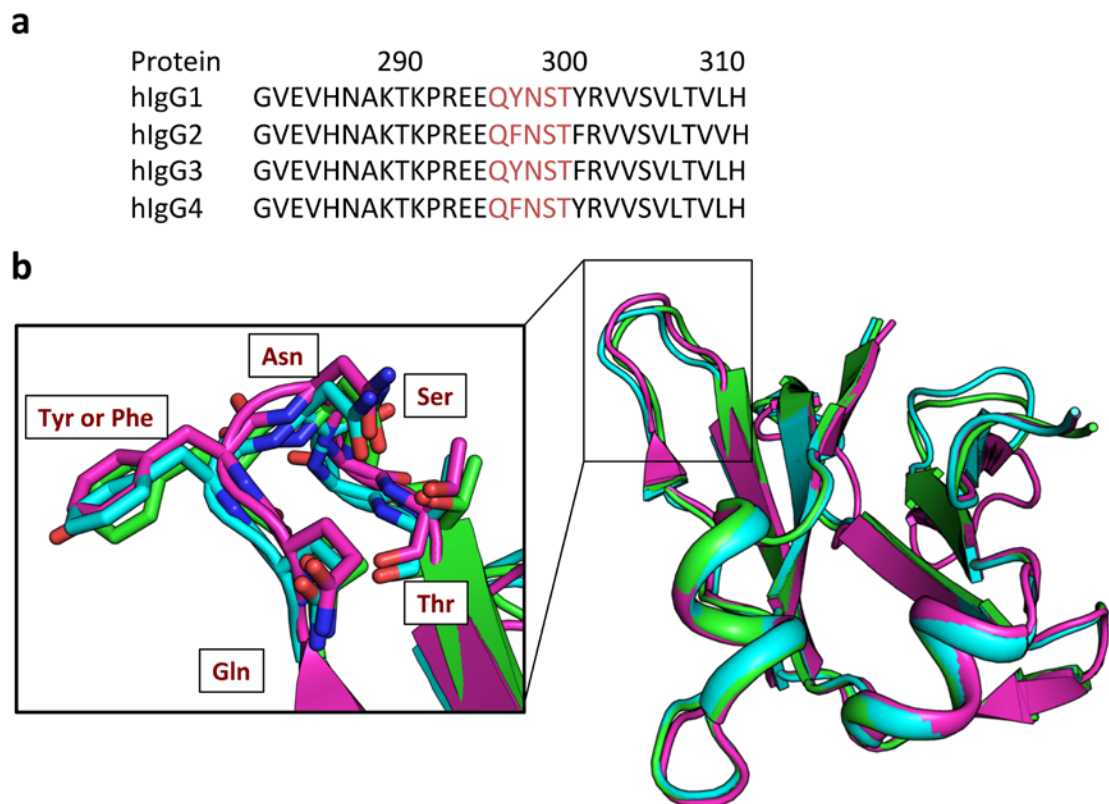


Figure S6. Sequence alignment and structure of human IgG1-4 around the C'E loop. (a) Sequence alignment of human IgG1-4 around the C'E loop. Sequences of the C'E loops are highlighted in red. (b) Alignment of the C_H2 domains from Fc structures of IgG1 (3AVE, green), IgG2 (4HAF, cyan) and IgG4 (4C55, magenta). Only the protein backbone atoms are used in the structural alignment. Inset: The side chains of the residues in the C'E loop (Gln295-Thr299) are shown in stick format.

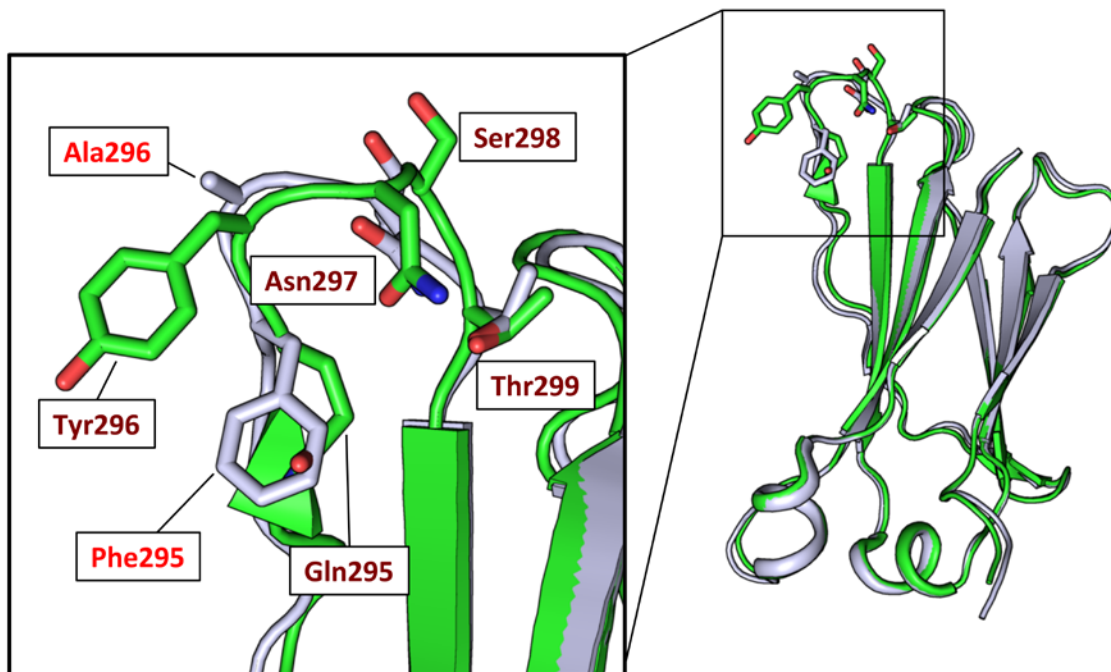


Figure S7. Alignment of the C_H2 domains of the Fc structures of wild-type (3AVE, green) and EAS-stabilized Fc (grey). The side chains of Gln295-Thr299 in the wild-type and the Phe295-Thr299 in the EAS-stabilized Fc are indicated and shown in stick format. The N-glycans are not shown.

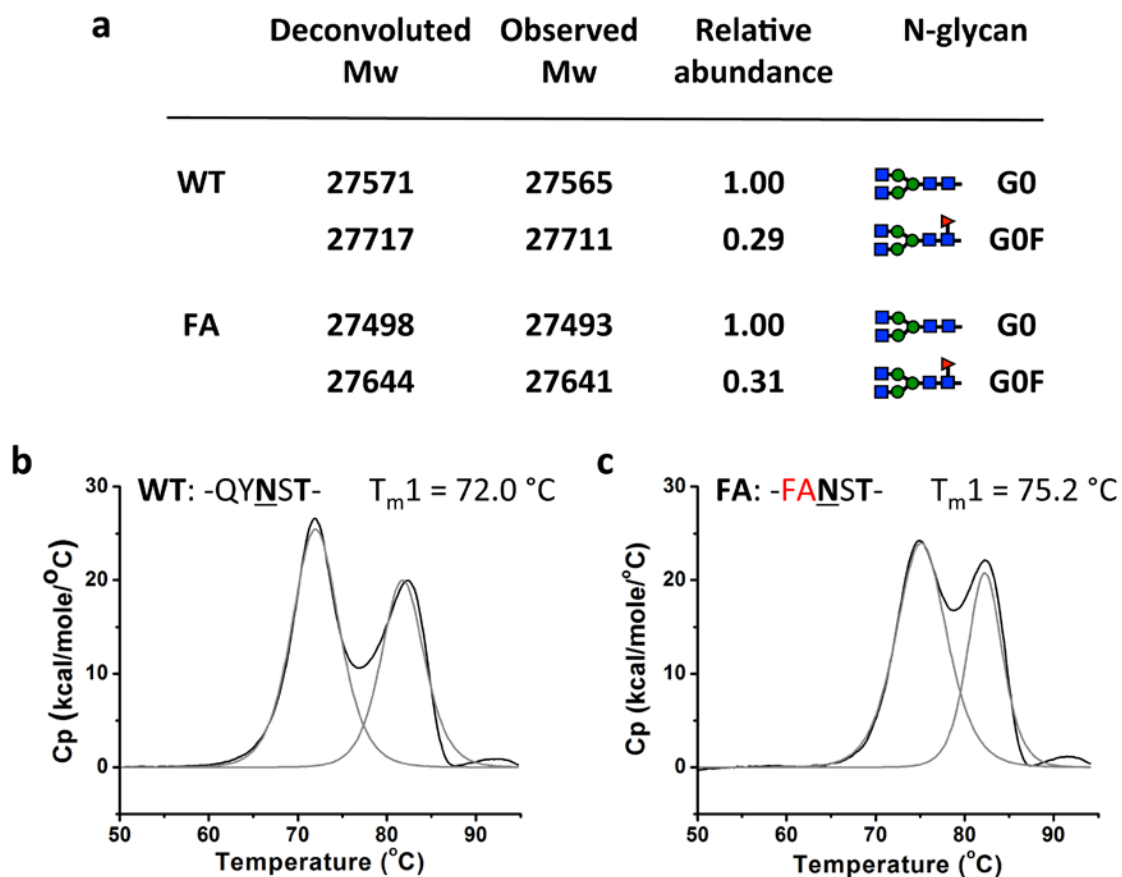


Figure S8. ESI-MS and DSC analysis of WT and FA Fc variants expressed with inhibition of core fucosylation. (a) Deconvoluted molecular weight (Mw) and relative abundance (normalized to the most abundant species) of ESI-MS data for each dithiothreitol-treated Fc fragment variant expressed under conditions of core fucosylation inhibition employing 2-deoxy-2-fluoro-L-fucose (200 μM). (b) The baseline-adjusted DSC trace (black line) showed a two-peak feature for each variant which could be deconvoluted (grey lines) to the melting of C_{H2} (lower temperature, T_{m1}) and C_{H3} (higher temperature, T_{m2}) domains. The protein primary sequence of the C'E loop and T_{m1} for each variant are shown above the DSC curves.

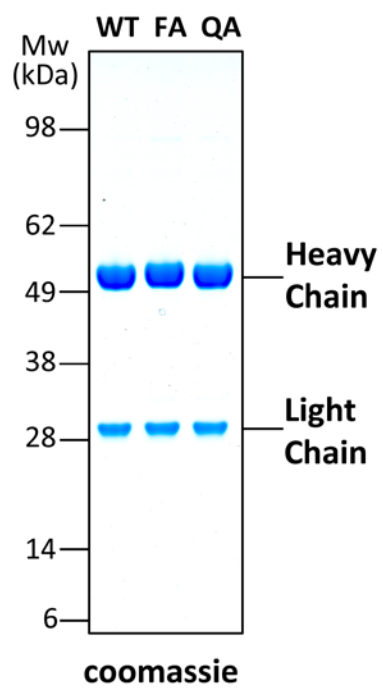


Figure S9. Coomassie stained SDS-PAGE of purified 5J8 antibody variants.

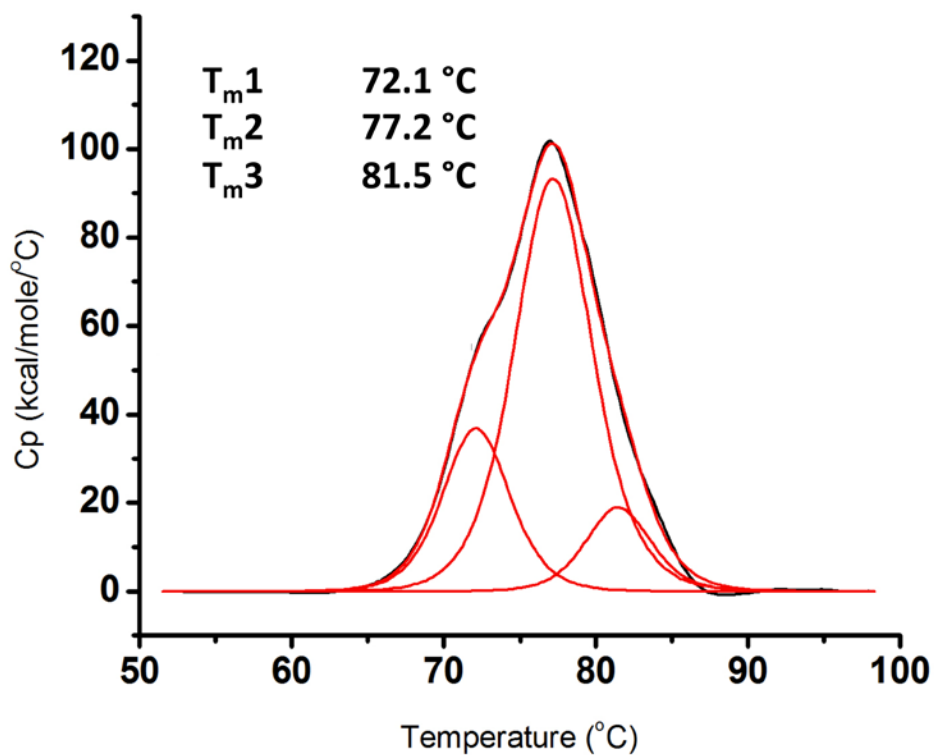


Figure S10. DSC analysis of the wild-type 5J8 antibody. The baseline-adjusted DSC trace is shown in black and the deconvoluted peaks, which correspond to the melting of C_H2 (T_{m1}), Fab (T_{m2}) and C_H3 (T_{m3}) domains, are shown in red. The combined signal (also shown in red) of the deconvoluted peaks matches closely to the baseline-adjusted DSC trace.

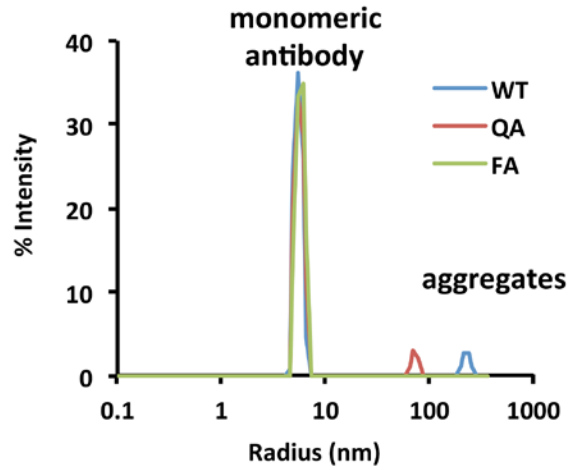


Figure S11. DLS analysis of the 5J8 variants. DLS signal of the supernatant of 5J8 antibody samples for WT (blue), QA (red) and FA (green) in pH 3.5 buffer after $20,000 \times g$ centrifugation. Most of the signal for all three variants is due to the monomeric form ($R_h = 7$ nm), indicating that high speed centrifugation removes the $R_h = 100$ -300 nm aggregates from the solution.

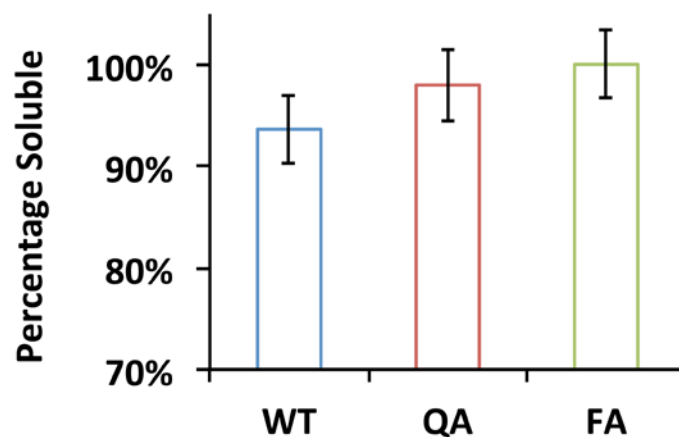


Figure S12. Quantification of the soluble fraction of 5J8 antibody variants upon diluting in pH 3.5 buffer. The percentage is calculated by dividing the protein concentration measured by microBCA assay after high speed centrifugation by the protein concentration before high speed centrifugation. Less than 10% of the total protein for each variant is removed by centrifugation, indicating that even such small amounts of the aggregates formed in the WT and QA variants of 5J8 is enough to completely suppress the DLS signal for the monomeric antibodies.

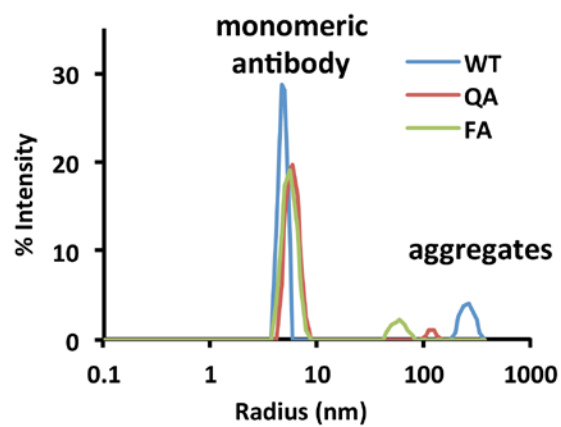


Figure S13. DLS signal of 5J8 antibody samples for WT (blue), QA (red) and FA (green) variants after dilution in PBS. Most of the signal for all three variants is due to the monomeric form ($R_h = 7$ nm).

Table S1. Protein primary sequence differences and C'E loop turn types for different PDB structures.

PDB code	Residue 272	Residue 283	Residue 294	Residue 312	Residue 315	Residue 356	Residue 358	Residue 393	turn type	turn residues
1H3T	Q	Q	Q	N	D	E	M	T	IV	YNST
1E4K	Q	Q	Q	N	D	E	M	T	IV	YNST
1FC1	Q	Q	Q	N	D	E	M	T	IV	YNST
3RY6	Q	Q	Q	N	D	E	M	T	IV	YNST
2WAH	E	E	E	D	N	D	L	A	I	QYNS
4CDH	E	E	E	D	N	D	L	T	I	QYNS
1HZH	E	E	E	D	N	D	L	T	I	QYNS
3AVE	E	E	E	D	N	D	L	T	I	QYNS
3SGJ	E	E	E	D	N	D	L	T	I	QYNS
4KU1	E	E	E	D	N	D	L	T	IV	QYNS
4X4M	E	E	E	D	N	D	L	T	I	QYNS

Table S2. X-ray data collection and refinement statistics.

Data collection	Human IgG Fc domain with enhanced aromatic sequon
X-ray Source	SSRL11-1
Wavelength (Å)	0.9795
Space group	C2
Unit cell parameters	a = 154.1, b = 49.2, c = 75.5 Å $\alpha = \gamma = 90.0^\circ$, $\beta = 104.2^\circ$
Resolution (Å)	50.0-3.00 (3.05-3.00) ^a
Observations	22,488
Unique reflections	10,017(471) ^a
Redundancy	2.2 (2.1) ^a
Completeness (%)	88.5 (83.1) ^a
$\langle I/\sigma \rangle^b$	12.9 (1.8) ^a
R_{sym}^c	0.08 (0.33) ^a
R_{pim}^c	0.06 (0.26) ^a
CC _{1/2}	(0.78) ^a
Refinement statistics	
Resolution (Å)	46.85-3.00 (3.29-3.00) ^a
Reflections (work)	9,487 (2,229) ^a
Reflections (test)	518 (135) ^a
$R_{\text{cryst}}(\%)^c$	29.8 (33.0) ^a
$R_{\text{free}}(\%)^d$	32.6 (36.5) ^a
Average B-value (Å ²)	77
Fc protein chain A	76
Fc glycan chain A	238
Fc protein chain B	71
Fc glycan chain B	101
Wilson B-value (Å ²)	69
RMSD from ideal geometry	
Bond length (Å)	0.004
Bond angles (°)	0.846
Ramachandran statistics (%) ^f	
Favored	95.0
Allowed	4.0
Outliers	1.0
PDB ID	4QGT

^a Numbers in parentheses refer to the highest resolution shell.

^b Calculated as $\text{average}(I)/\text{average}(\sigma I)$

^c $R_{\text{sym}} = \sum_{hkl} \sum_i |I_{hkl,i} - \langle I_{hkl} \rangle| / \sum_{hkl} \sum_i I_{hkl,i}$, where $I_{hkl,i}$ is the scaled intensity of the i^{th} measurement of reflection h, k, l, $\langle I_{hkl} \rangle$ is the average intensity for that reflection, and n is the redundancy. R_{pim} is a redundancy-independent measure of the quality of intensity measurements. $R_{\text{pim}} = \sum_{hkl} [1/(n-1)]^{1/2} \sum_i |I_{hkl,i} - \langle I_{hkl} \rangle| / \sum_{hkl} \sum_i I_{hkl,i}$, where $I_{hkl,i}$ is the scaled intensity of the i^{th} measurement of reflection h, k, l, $\langle I_{hkl} \rangle$ is the average intensity for that reflection, and n is the redundancy.

^d $R_{\text{cryst}} = \sum_{hkl} |F_o - F_c| / \sum_{hkl} |F_o| \times 100$

^e R_{free} was calculated as for R_{cryst} , but on a test set comprising 5% of the data excluded from refinement.

^f These values were calculated using MolProbity (<http://molprobity.biochem.duke.edu/>).

Table S3. Summary of the type I β -turn sequences and dihedral angles of the WT Fc (PDBID:3AVE, chain B) and FA variant (chain B) structures.

Fc variant	PDB/chain	Turn Sequence	Gln/Phe295		Tyr/Ala296		Asn297		Ser298		Thr299	
			phi	psi	phi	psi	phi	psi	phi	psi	phi	psi
WT	3AVE/B	Gln295-Ser298	-112.9	164.3	-63.1	-36.5	-84.6	23.1	73.8	23.6	-126.5	169.3
FA	4QGT/B	Phe295-Ser298	-99.6	173.7	-62.5	-32.8	-60.1	-37.8	84.9	4.4	-88.6	117.9