Supporting Information

SI Materials and Methods

Protein production. All EndoS constructs derived originally from pGEXndoS (Genbank entry: AF296340). All EndoS proteins were expressed in *Escherichia coli* and purified as described (1, 2). SeMet-labeled $EndoS_{D233Q}$(98-995) was produced in *E. coli* B834(DE3) cells by autoinduction (3). The purification procedure was identical to that used for the native protein, but with the addition of 1 mM dithiothreitol to all buffers in order to prevent selenium oxidation. The presence of 16 SeMet residues was verified by mass spectrometry. Fucosylated IgG1 Fc with homogenous asialo-biantennary complex type N-glycan was obtained from Rituximab papain digestion and chemoenzymatic glycoengineering as described previously (1).

Oligomerization states. CPD fusion proteins $EndoS_{WT}$(37-995) and $EndoS_{WT}$(98-995) (3 mg mL^{-1} each) were applied to a Superdex 200 10/300 GL SEC column equilibrated in 50 mM Tris-Cl, pH 8.0, with either 5 mM EDTA or 2 mM $CaCl_2$ running at 1 mL min^{-1}. Elution time was compared to those of molecular weight standards (Bio-Rad) to determine oligomerization states.

Hydrolytic activity. Two mixtures of Rituximab IgG (3.3 μM each) in 100 mM Tris-HCl pH 8.0 with or without 20 mM EDTA were incubated with 0.5 nM $EndoS_{WT}$(37-995) or 100 nM $EndoS_{WT}$(98-995) at 37˚C. Aliquots of each reaction were removed at timed intervals, immediately quenched in 2xSDS loading buffer and separated by SDS-PAGE, the data from which were analyzed by band densitometry using ImageQuant software. Hydrolytic activity for all EndoS mutants was evaluated for Ig2, IgG3 and IgG4, bearing kappa light chains (Sigma-Aldrich) using the same procedure.

Protein crystallization. Crystallization of native EndoS has been described in detail (2). SeMet-$EndoS_{D233Q}$(98-995) crystals were obtained by liquid-liquid diffusion using Crystal Formers (Microlytic) by micro-seeding with native crystals in the previously determined crystallization condition.

Structure determination and refinement. For data collection, crystals were flash cooled at 100 K in mother liquor containing 20% ethylene glycol. Diffraction data for SeMet-$EndoS_{D233Q}$(98-995) were collected using a Dectris PILATUS 6M detector at beam line 11-1 at Stanford Synchrotron Radiation Lightsource (SLAC National Accelerator Laboratory, CA); diffraction data for crystals of native $EndoS_{D233Q}$(98-995) and $EndoS_{WT}$(98-995) were collected using a MAR 300 CCD detector at beam line 23-ID-B at the Advanced Photon Source (Argonne National Laboratory, IL). Data were processed and indexed with XDS (4) and scaled with the Xscale (5). The structure of SeMet-$EndoS_{D233Q}$(98-995) was solved by multi-wavelength anomalous dispersion (MAD) at 3.2 Å using a MAD script by A. Gonzalez (with SHELX options based on a script by Qingping Xu) including the programs SHELX (6-8), SOLVE (9) and RESOLVE (10). $EndoS_{D233Q}$(98-995) and $EndoS_{WT}$(98-995) were solved at 1.9 Å and 2.6 Å resolution, respectively, by molecular replacement using Phaser (11), using the SeMet-$EndoS_{D233Q}$(98-995) structure as a model. All data collection and statistics are shown in Table S1. The structures were built and refined using the programs Coot and PHENIX (12), respectively. TLS-refinement (13) with PHENIX was used to refine the $EndoS_{D233Q}$(98-995) structure at 1.9 Å. Interfaces and buried surface areas were calculated using PISA server (14).

Small Angle X-ray Scattering (SAXS). SAXS experiments were performed at Bio-SAXS beamline BL4-2 at Stanford Synchrotron Radiation Lightsource (SSRL) (15). Data were collected using a MX225-HE CCD detector (Rayonix) with a 1.7 m sample-to-detector distance and beam energy of 11 keV (wavelength, λ = 1.127 Å). The momentum transfer (scattering vector) q measure in inverse Angstroms ($Å^{-1}$) was defined as $q = 4\pi\sin(\theta)/l$, where 2θ was scattering angle. The q scale was calibrated with silver behenate powder. Data were collected using the BL4-2 automatic sample-loading robot (16, 17). 30 μl of buffer and sample were exposed to the X-ray beam via a 1.5mm quartz capillary cell (Hampton Research), oscillated to reduce radiation damage.

Binding analysis. All SPR experiments were performed using a Biacore T100 instrument (GE Healthcare). IgG1 Fc in 10 mM sodium acetate, pH 4 was immobilized at a density of 500 RU to flow cell 1 and flow cell 2 in a CM5 sensor chip via standard amine-coupling procedure, using HBS-X buffer (10 mM HEPES, 150 mM NaCl, 0.05 % Tween 20) as running buffer. $EndoS_{E235Q}$(98-995) does not bind to deglycosylated IgG (18), thus, N-linked glycan of IgG1 Fc in flow cell 1 was removed using 2x10 μL of 1 mg/mL $EndoS_{WT}$(37-995) and this flow cell was used as negative control surface. Concentration series of $EndoS_{E235Q}$(37-995) (7-0.27 μM), $EndoS_{E233Q}$(98-995) (500-31.25 μM), $EndoS_{E235Q}$(98-764) (31.5-0.11 μM), EndoS(446-995) (μM), $EndoS_{E235Q/\Delta 314-323}$ (652-2.6 μM), $EndoS_{E235Q/\Delta 528-554}$ (38.2-0.07 μM),$EndoS_{E235Q/\Delta 742-750}$ (231-0.9 μM), $EndoS_{E235Q/\Delta 793-797}$ (12.2-0.05 μM), $EndoS_{W803A}$ (623-2.4 μM), $EndoS_{E833A}$ (607-2.3 μM) and $EndoS_{R908A}$ (27-0.11 μM) in running buffer were injected over flow cells 1 and 2 for 60 s per injection and allowed to dissociate for 300 s. Between binding cycles, the sensor chip surface was regenerated by washing with

2 M NaCl. Affinity constants for all the proteins were calculated using a general steady-state equilibrium model with the Biacore T100 evaluation software 2.0.4.

Molecular modeling. We used ZDOCK 3.0.2 (19) to dock the glycosylated IgG1 Fc (PDB code 4BYH) to the $EndoS_{D233Q}$(98-995) structure with 6 degree sampling, assigning ZDOCK atom type and radius parameters to the Fc glycan atoms (partial charges of these atoms were set to zero for docking). Prior to docking, we added missing loop residues to the EndoS crystal structure using Modeller (20) followed by refinement of the modeled loops in Rosetta (21), keeping the crystallographically determined coordinates fixed. The dDFIRE statistical potential (22) was used to score the 120 EndoS models with refined loops and select a structure for docking input. We performed docking refinement by adapting an algorithm recently developed for docking T cell receptors onto peptide-MHC complexes (23), iterating rigid-body and side chain movements with flexible loop minimization. EndoS residues 313-322, 742-750, and 791-798 were selected for loop minimization during docking, due to their surface exposure and proximity to IgG Fc in the docking model. Computational interface alanine scanning of the refined model to determine putative energetic hot spots was performed using the "interface" protocol of Rosetta 2.0.2 (24).

References

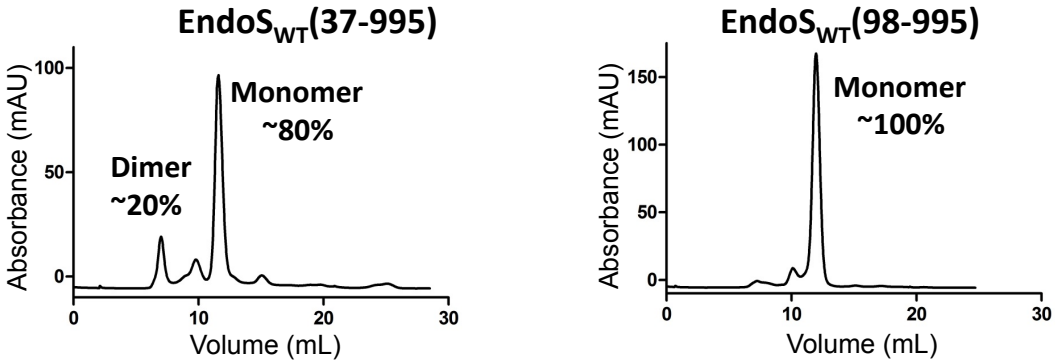
1. Huang W, Giddens J, Fan SQ, Toonstra C, & Wang LX (2012) Chemoenzymatic glycoengineering of intact IgG antibodies for gain of functions. *Journal of the American Chemical Society* 134(29):12308-12318.
2. Trastoy B, Lomino JV, Wang L-X, & Sundberg EJ (2013) Liquid-liquid diffusion crystallization improves the X-ray diffraction of EndoS, an endo-[beta]-N-acetylglucosaminidase from Streptococcus pyogenes with activity on human IgG. *Acta Crystallographica Section F* 69(12):1405-1410.
3. Studier FW (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* 41(1):207-234.
4. Kabsch W (2010) Xds. *Acta crystallographica. Section D, Biological crystallography* 66(Pt 2):125-132.
5. Kabsch W (2010) Integration, scaling, space-group assignment and post-refinement. *Acta crystallographica. Section D, Biological crystallography* 66(Pt 2):133-144.
6. Schneider TR & Sheldrick GM (2002) Substructure solution with SHELXD. *Acta Crystallographica Section D* 58(10 Part 2):1772-1779.
7. Sheldrick G (2008) A short history of SHELX. *Acta Crystallographica Section A* 64(1):112-122.
8. Sheldrick GM (2002) Macromolecular phasing with SHELXE. in *Zeitschrift für Kristallographie/International journal for structural, physical, and chemical aspects of crystalline materials*, p 644.
9. Terwilliger TC & Berendzen J (1999) Automated MAD and MIR structure solution. *Acta Crystallographica Section D* 55(4):849-861.
10. Terwilliger TC (2000) Maximum-likelihood density modification. *Acta Crystallogr D Biol Crystallogr* 56(Pt 8):965-972.
11. McCoy AJ*, et al.* (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(Pt 4):658-674.
12. Adams PD*, et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography* 66(Pt 2):213-221.
13. Painter J & Merritt EA (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta crystallographica. Section D, Biological crystallography* 62(Pt 4):439-450.
14. Krissinel E & Henrick K (2007) Inference of Macromolecular Assemblies from Crystalline State. *J Mol Biol* 372(3):774-797.
15. Smolsky IL*, et al.* (2007) Biological small-angle x-ray scattering facility at the Stanford synchrotron radiation laboratory. *Journal of Applied Crystallography* 40.
16. McPhillips TM*, et al.* (2002) Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines. *Journal of Synchrotron Radiation* 9.
17. Martel A, Liu P, Weiss TM, Niebuhr M, & Tsuruta H (2012) An integrated high-throughput data acquisition system for biological solution X-ray scattering studies. *Journal of Synchrotron Radiation* 19.
18. Allhorn M, Olin AI, Nimmerjahn F, & Collin M (2008) Human IgG/Fc gamma R interactions are modulated by streptococcal IgG glycan hydrolysis. *PLoS One* 3(1):e1413.
19. Pierce BG, Hourai Y, & Weng Z (2011) Accelerating protein docking in ZDOCK using an advanced 3D convolution library. *PLoS One* 6(9):e24657.
20. Fiser A & Sali A (2003) Modeller: generation and refinement of homology-based protein structure models. *Methods in enzymology* 374:461-491.
21. Mandell DJ, Coutsias EA, & Kortemme T (2009) Sub-angstrom accuracy in protein loop reconstruction by robotics-inspired conformational sampling. *Nat Methods* 6(8):551-552.
22. Yang Y & Zhou Y (2008) Specific interactions for ab initio folding of protein terminal regions with secondary structures. *Proteins* 72(2):793-803.
23. Pierce BG & Weng Z (2013) A flexible docking approach for prediction of T cell receptor-peptide-MHC complexes. *Protein science : a publication of the Protein Society* 22(1):35-46.
24. Kortemme T & Baker D (2002) A simple physical model for binding energy hot spots in protein-protein complexes. *Proceedings of the National Academy of Sciences of the United States of America* 99(22):14116-14121.

Table S1. Data collection and refinement statistics

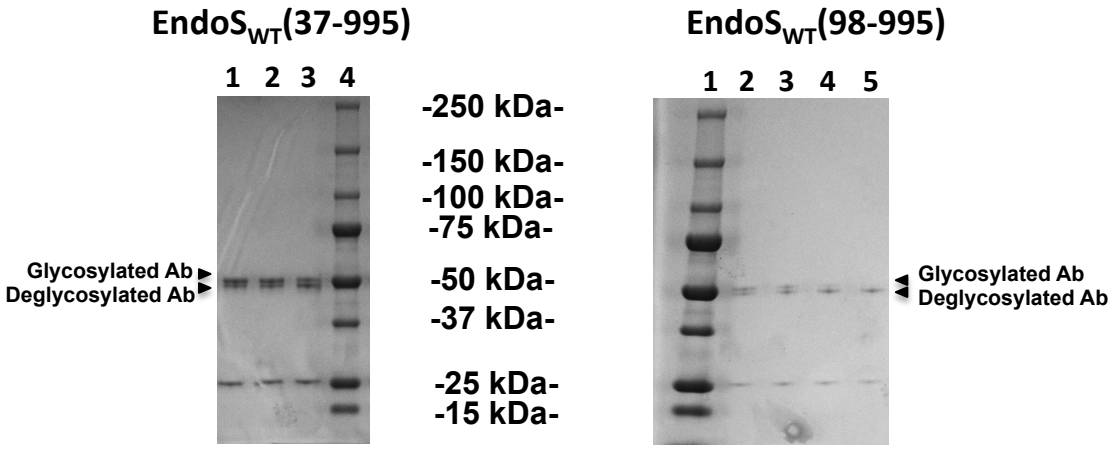
	SeMetEndoS$_{D233Q}$			EndoS$_{D233Q}$	EndoS$_{WT}$
Data collection					
Space group		$P2_12_12_1$		$P2_12_12_1$	$P2_12_12_1$
Cell dimensions					
a, b, c (Å)	86.3, 93.0, 137.8	86.7, 93.5, 138.5	86.2, 92.9, 137.8	92.6, 96.1, 141.2	92.3, 94.5, 142.8
α, β, γ (º)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
	Peak	Inflection	Remote		
Wavelength	0.979	0.978	0.918	0.979	0.979
Resolution (Å)	40.0-3.18 (3.37-3.18)	40.0-3.17 (3.36-3.17)	40-3.18 (3.37-318)	30-1.91 (2.02-1.91)	30-2.63 (2.77-2.63)
R_{sym}	11.1 (43.4)	10.0 (44.5)	12.2 (59.4)	4.6 (47.9)	9.4 (63.7)
I/σI	17.2 (4.6)	9.6 (2.5)	15.9 (3.6)	10.6 (1.6)	9.8 (1.6)
Completeness (%)	98.1 (89.5)	94.7 (84.1)	98.0 (89.0)	96.8 (95.2)	97.8 (90.6)
Redundancy	8.0 (7.5)	3.8 (3.8)	8.1 (7.5)	2.0 (1.8)	2.6 (2.5)
Refinement					
Resolution (Å)				29.7-1.9	29.3-2.6
No. reflections				98436	38280
R_{work}/R_{free}				19.2/23.5	21.3/26.2
No. atoms					
Protein				7047	6978
Ligand				1	1
Water				848	159
B-factors					
Protein				38.4	60.4
Ligand				42.4	84.0
Water				40.3	49.9
RMS deviations					
Bond lengths (Å)				0.016	0.002
Bond angles (º)				1.51	0.61
Ramachandran					
Most favored (%)				97	97
Additional allowed (%)				2.9	3
Disallowed (%)				0.1	0
PDB code				4NUZ	4NUY

Number of crystals for each structure is one.
*Values in parentheses are for highest-resolution shell

A



B



C

Construct	IgG1	IgG2	IgG3	IgG4
EndoS$_{WT}$(37-995)	++++	++	++	++
EndoS$_{WT}$(98-995)	++	+	+	+
EndoS$_{WT}$(98-764)	++	+	+	+
EndoS$_{WT}\Delta_{314-323}$	-	-	-	-
EndoS$_{WT}\Delta_{528-554}$	-	-	-	-
EndoS$_{WT}\Delta_{742-750}$	+	-	-	-
EndoS$_{WT}\Delta_{793-797}$	+++	++	++	++
EndoS$_{WT/W803A}$	-	-	-	-
EndoS$_{WT/E833A}$	++++	++	++	+++
EndoS$_{WT/R908A}$	+++	++	++	++

Figure S1. Oligomerization and hydrolytic activity analysis of EndoS$_{WT}$(37-995) and EndoS$_{WT}$(98-995). **(A)** Size exclusion chromatographic analysis of EndoS$_{WT}$(37-995) (*left panel*) and EndoS$_{WT}$(98-995) (*right panel*)**. (B)** *Left panel,* EndoS$_{WT}$(37-995) (0.5 nM) digest of Rituximab (3.3 μM): Lane 1, at 40 min; Lane 2, at 80 min; Lane 3, at 160 min; Lane 4, molecular weight ladder. *Right panel,* EndoS$_{WT}$(98-995) (100 nM) digest of Rituximab(3.3 μM): Lane 1, molecular weight ladder; Lane 2, at 20 min; Lane 3, at 40 min; Lane 4, at 80 min; Lane 5, at 160 min. 70% hydrolysis of an identical amount of glycosylated IgG antibody was achieved by 0.5 nM EndoS$_{WT}$(37-995) in 160 minutes and by 100 nM EndoS$_{WT}$(98-995) in 40 minutes. Thus, the catalytic rate of EndoS$_{WT}$(37-995) is 50-fold faster than that of EndoS$_{WT}$(98-995). **(C)** Hydrolytic activity for EndoS proteins for all human IgG subclasses. ++++, 100% hydrolysis at 1 hour; +++, 75-90% hydrolysis at 1 hour; ++, 100% hydrolysis at 3 hours; + <100% hydrolysis at 3 hours; -, no hydrolysis at 3 hours.

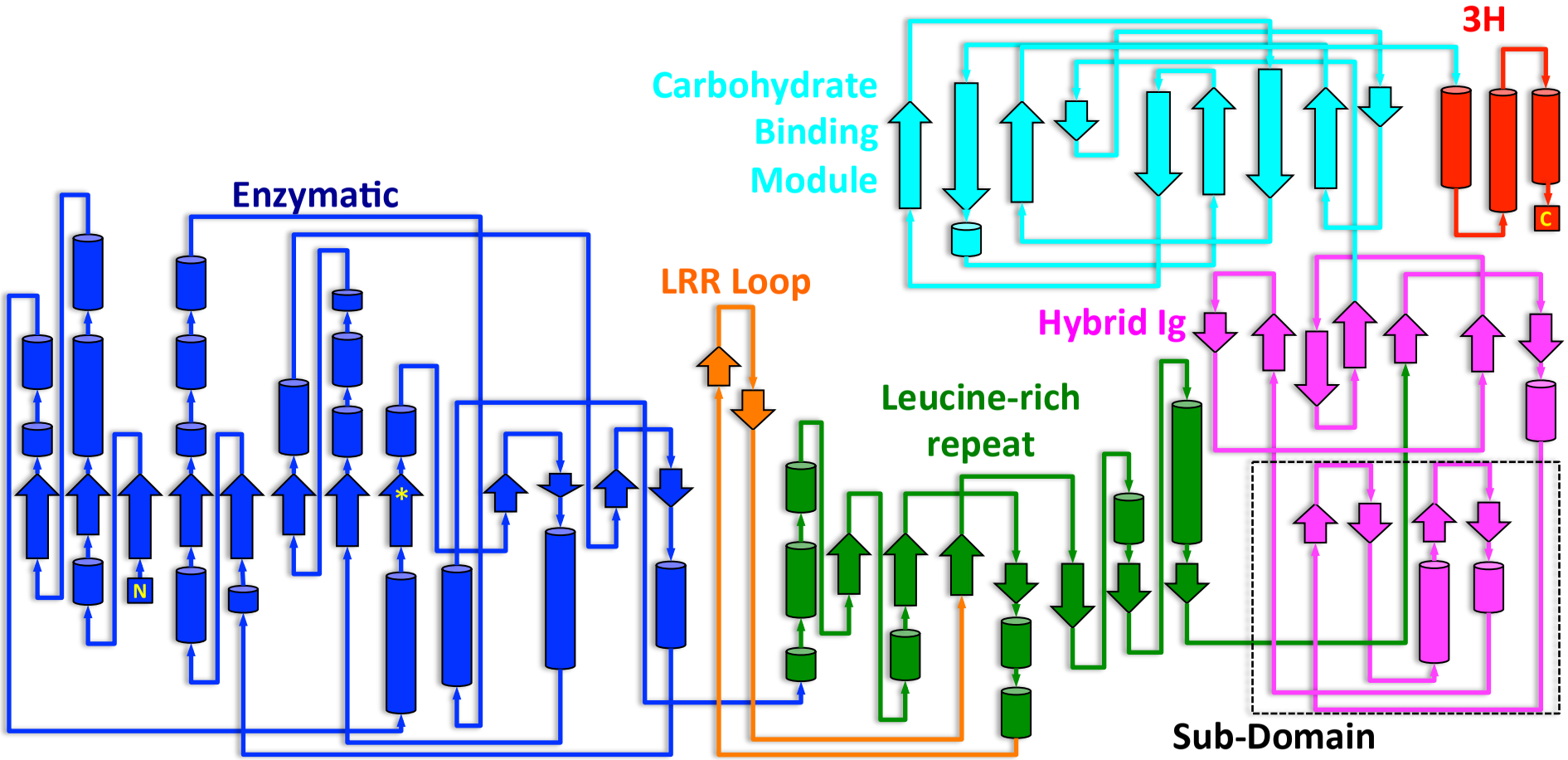


Figure S2. Secondary structure topology of the EndoS(98-995) structure. The N- and C-termini are marked by "N" and "C," respectively; the active site position is marked by a yellow asterisk; α helices are represented by cylinders; β strands are represented by arrows.

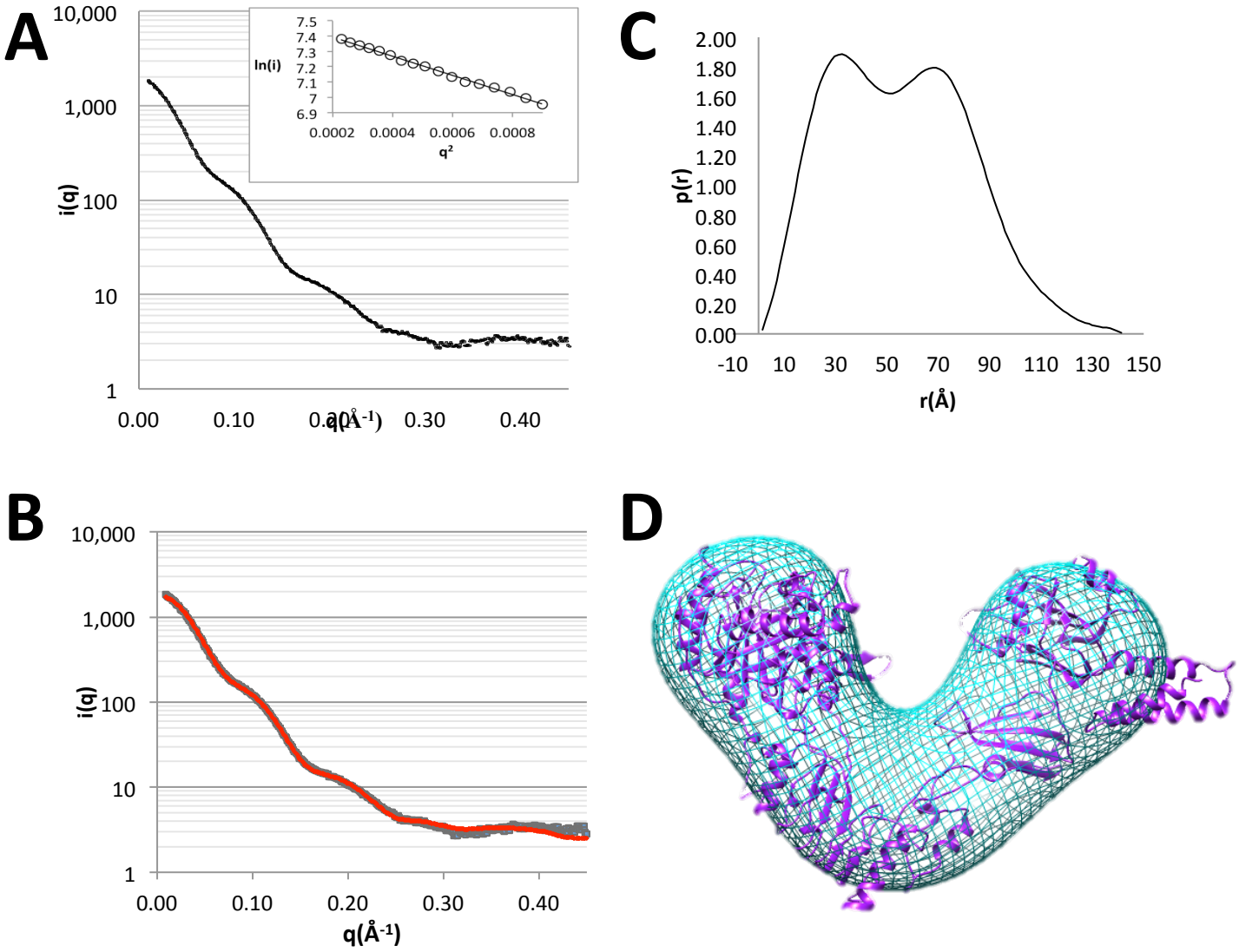


Figure S3. Small angle X-ray scattering (SAXS) analysis of EndoS. **(A)** Data were collected from a 1, 2, 5 and 10 mg/ml concentration series in order to detect concentration-dependent intermolecular interactions. Fifteen 1 s images were averaged using SasTool (http://ssrl.slac.stanford.edu/~saxs/analysis/sastool.htm). The averaged buffer curve was then subtracted from the averaged protein curves. The curves were examined using PRIMUS and a final curve for further analysis was produced by merging the low-q region of the 1 mg/ml curve with high q region of the 10 mg/ml curve. **(B)** Analysis of the Guinier region between 0.0002 and 0.0008 $Å^2$ (consisting of 16 data points, and a q x R_g max of 1.295) showed no significant evidence of aggregation or inter-particle effects, and gave an R_g of 43.2 +/- 0.2 Å, which is similar to the R_g of the crystal structure calculated from CRYSOL of 41.67 Å. Comparison of the computed scattering curve of the solved structure to the experimental data by CRYSOL gave a Chi of 3.35, showing the protein in solution adopts a similar confirmation to the crystal structure **(C)** Data from q=0.015 to 0.185 $Å^{-1}$ as suggested by AUTOGNOM were used to produce the pr function, giving a dmax of 141.3 Å , a reciprocal space R_g of 42.89 Å, and a real space R_g of 42.96 +/- 0.072 Å. The pair-distribution function was shown to be bimodal, consistent with the V shape conformation of the EndoS crystal structure. **(D)** The pr function was used as input for *ab initio* modeling with the Shapeup shape construction module of SASTBX. Superimposition of $EndoS_{D233Q}$(98-995) x-ray crystal structure (purple ribbon) into the envelope (blue mesh) is shown.

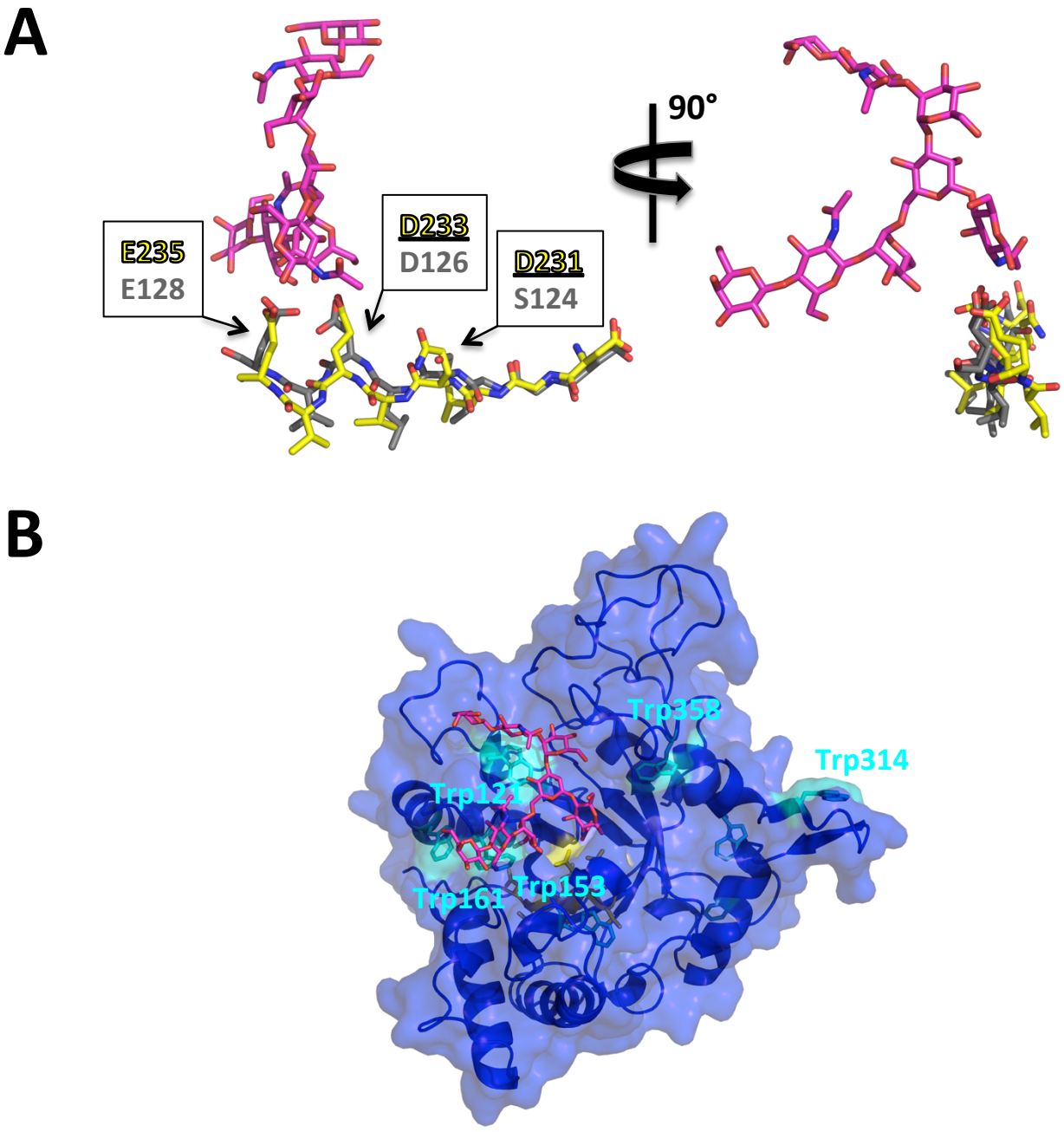


Figure S4. Structural details of the EndoS glycosidase domain. **(A)** Active site residues of EndoS (yellow) align with those of $EndoF_3$ (grey) when their respective glycosidase domains are superimposed. The glycan from the $EndoF_3$-glycan structure is in magenta. **(B)** Cartoon and surface representation of the EndoS glycosidase domain. Surface-exposed Trp residues (cyan) that line the molecular grooves that accept the glycan and protein components of the glycoprotein substrate are highlighted. The glycan from from the $EndoF_3$-glycan structure is in magenta.

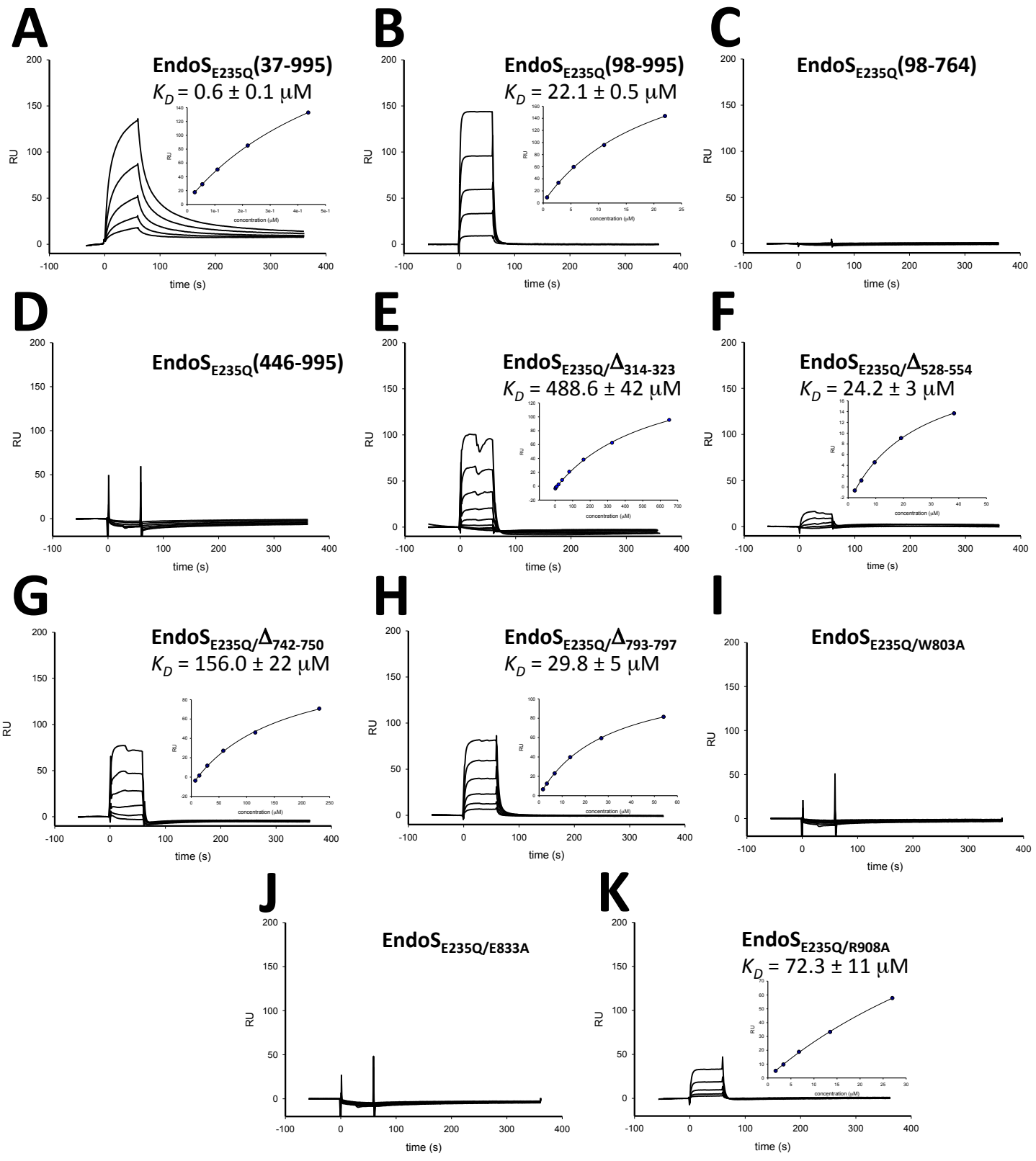


Figure S5. SPR sensograms of (**A**) EndoS$_{E235Q}$(37-995), (**B**) EndoS$_{E235Q}$(98-995), (**C**) EndoS$_{E235Q}$(98-764), (**D**) EndoS(446-995), (**E**) EndoS$_{E235Q/\Delta 314-323}$, (**F**) EndoS$_{E235Q/\Delta 528-554}$, (**G**) EndoS$_{E235Q/\Delta 742-750}$, (**H**) EndoS$_{E235Q/\Delta 793-797}$, (**I**) EndoS$_{E235Q/W803A}$, (**J**) EndoS$_{E235Q/E833A}$, (**K**) EndoS$_{E235Q/R908A}$ at different concentrations binding to immobilized IgG1 Fc. ***Inserts***: equilibrium responses as a function of EndoS mutant concentration.

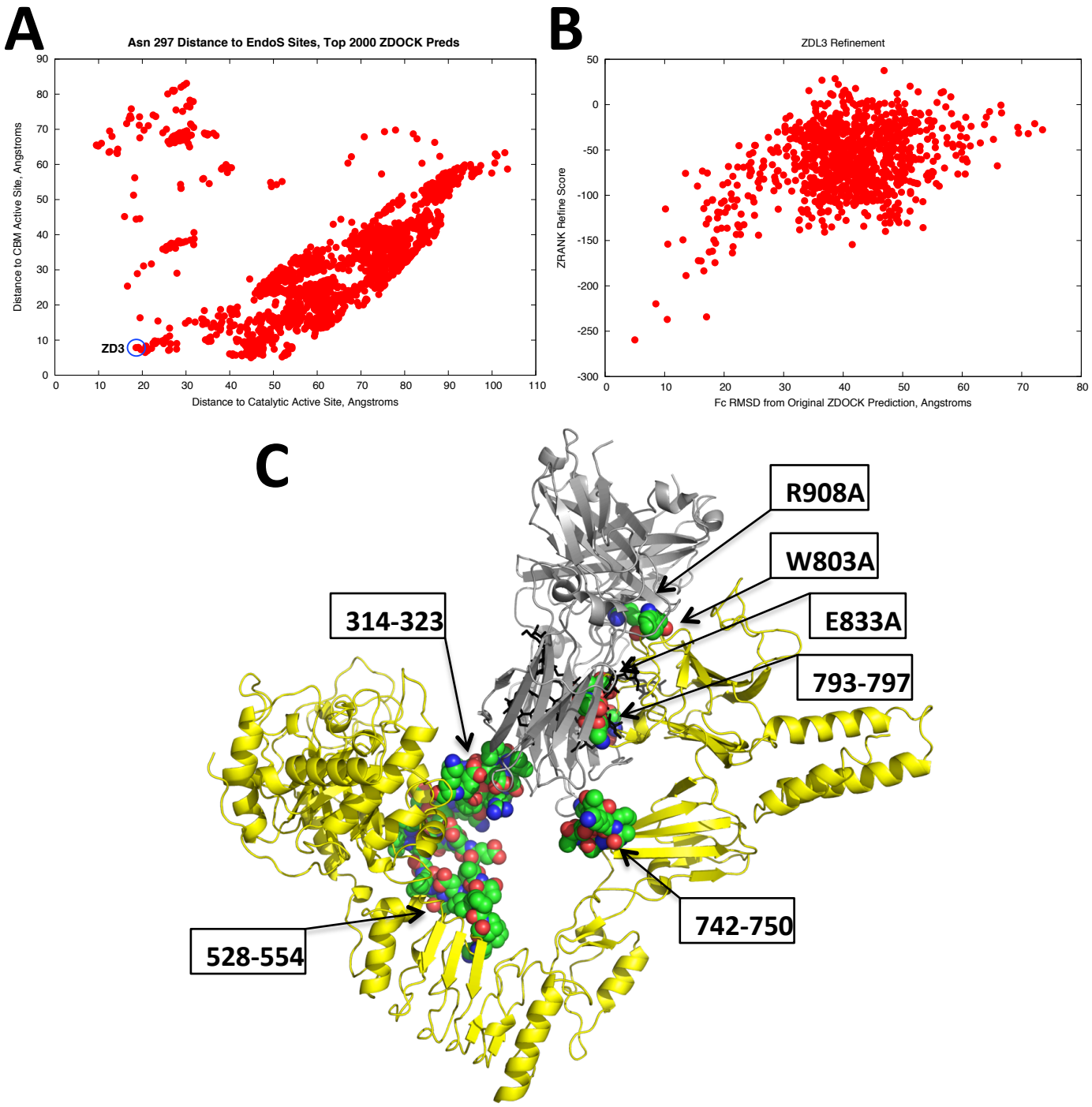


Figure S6. Molecular modeling of the EndoS-IgG1 Fc encounter complex and validation of the model. **(A)** The top 2000 ranked models from the initial modeling plotted in terms of distance of the average Asn297-linked glycan position to EndoS residues likely to bind glycan in the glycosidase (horizontal axis) and carbohydrate binding module (vertical axis) domains. **(B)** Subsequent refinement of model ZDL3 in terms of the position of IgG1 Fc relative to its pre-refinement position (horizontal axis) and its ZRANK refinement score (vertical axis). **(C)** EndoS is shown as cartoon in yellow, IgG1 Fc as cartoon in grey and Asn297-linked glycans as sticks in black. EndoS truncated loops (D314-323, D528-554, D742-750, D793-797) and alanine point mutations (W803A, E833A, and R908A) are shown as van der Waals spheres with carbon atoms in green, oxygen atoms in red and nitrogen atoms in blue.

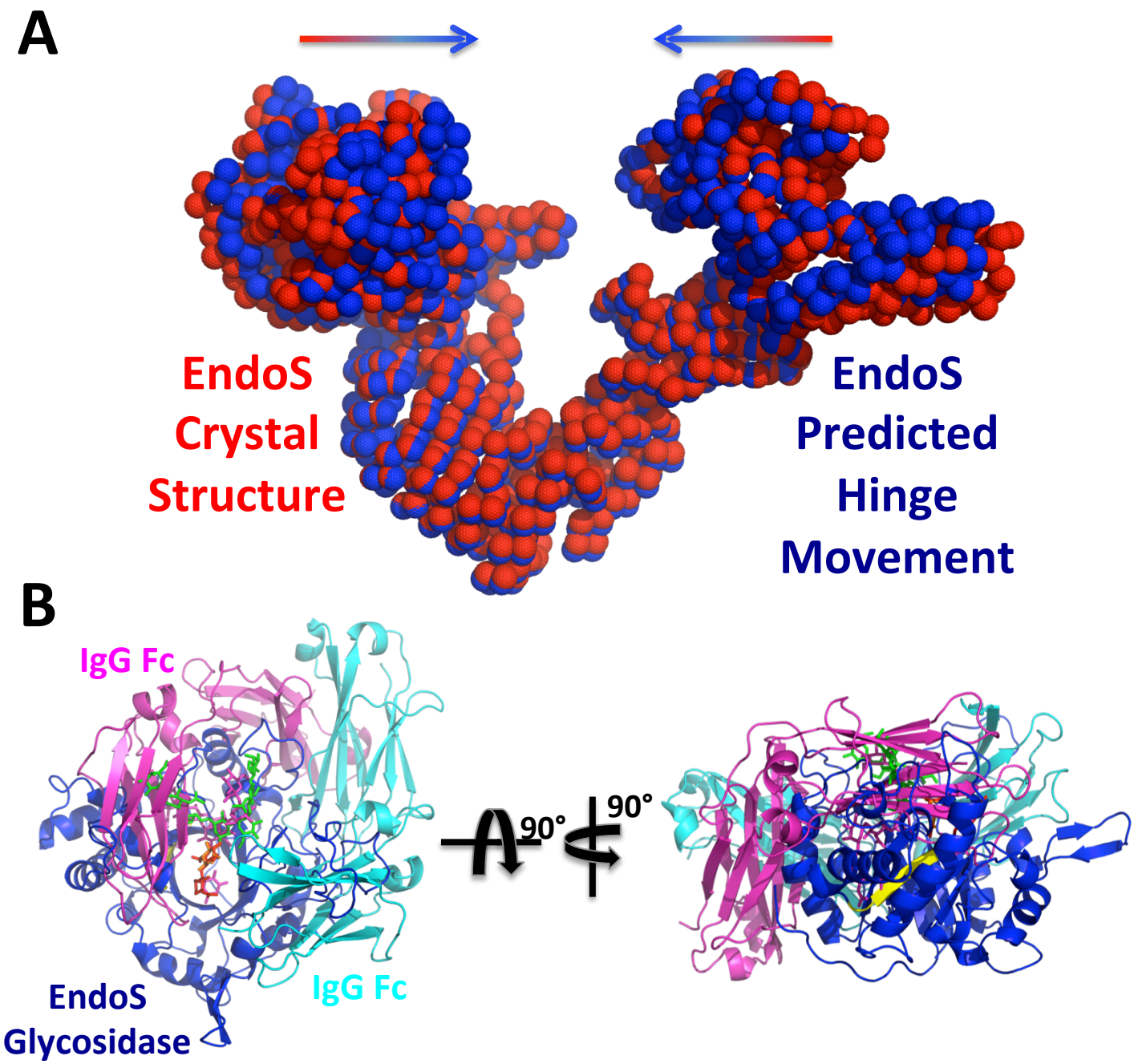


Figure S7. EndoS conformational changes required for enzymatic activity. **(A)** HingeProt analysis of the EndoSD233Q(98-995) crystal structure (red) compared to the most extreme predicted conformation (blue), representative of the predicted movement of the glycosidase domain (top left of molecule) and carbohydrate binding module (top right of molecule) domains toward one another. **(B)** Steric clashes between the EndoS glycosidase domain and IgG1 Fc when the former is aligned to the glycosidase domain of $EndoF_3$ in the $EndoF_3$-glycan and the Asn297-linked glycan of IgG1 Fc is aligned to glycan of $EndoF_3$ in the $EndoF3_3$-glycan complex structure.