Crystal structure of *Streptococcus pyogenes* EndoS, an immunomodulatory endoglycosidase specific for human IgG antibodies

Beatriz Trastoy[a], Joseph V. Lomino[a,b], Brian G. Pierce[c], Lester G. Carter[d], Sebastian Günther[a], John P. Giddens[a,b], Greg A. Snyder[a,e], Thomas M. Weiss[d], Zhiping Weng[c], Lai-Xi Wang[a,b], and Eric J. Sundberg[a,e,f,1]

[a]Institute of Human Virology and Departments of [b]Biochemistry and Molecular Biology, [e]Medicine, and [f]Microbiology and Immunology, University of Maryland School of Medicine, Baltimore, MD 21201; [c]Program in Bioinformatics and Integrative Biology and Department of Biochemistry and Molecular Pharmacology, University of Massachusetts School of Medicine, Worcester, MA 01655; and [d]Stanford Synchrotron Radiation Lightsource, Stanford Linear Accelerator Center National Laboratory, Menlo Park, CA 94025

Edited by Jeffrey V. Ravetch, The Rockefeller University, New York, NY, and approved March 27, 2014 (received for review December 11, 2013)

To evade host immune mechanisms, many bacteria secrete immunomodulatory enzymes. *Streptococcus pyogenes*, one of the most common human pathogens, secretes a large endoglycosidase, EndoS, which removes carbohydrates in a highly specific manner from IgG antibodies. This modification renders antibodies incapable of eliciting host effector functions through either complement or Fc γ receptors, providing the bacteria with a survival advantage. On account of this antibody-specific modifying activity, EndoS is being developed as a promising injectable therapeutic for autoimmune diseases that rely on autoantibodies. Additionally, EndoS is a key enzyme used in the chemoenzymatic synthesis of homogenously glycosylated antibodies with tailored Fc γ receptor-mediated effector functions. Despite the tremendous utility of this enzyme, the molecular basis of EndoS specificity for, and processing of, IgG antibodies has remained poorly understood. Here, we report the X-ray crystal structure of EndoS and provide a model of its encounter complex with its substrate, the IgG1 Fc domain. We show that EndoS is composed of five distinct protein domains, including glycosidase, leucine-rich repeat, hybrid Ig, carbohydrate binding module, and three-helix bundle domains, arranged in a distinctive V-shaped conformation. Our data suggest that the substrate enters the concave interior of the enzyme structure, is held in place by the carbohydrate binding module, and that concerted conformational changes in both enzyme and substrate are required for subsequent antibody deglycosylation. The EndoS structure presented here provides a framework from which novel endoglycosidases could be engineered for additional clinical and biotechnological applications.

Successful infection and colonization by microbes depends on their abilities to evade host immunity. One of the primary routes by which microorganisms escape host immune responses is through the production of enzymes that modify the immune system. *Streptococcus pyogenes*, a Gram-positive bacterium that is one of the most common human pathogens and the cause of group A streptococcal infections, expresses numerous extracellular enzymes that modulate immune mechanisms, including those that proteolyze antibodies and complement factors, detoxify oxygen free radicals, inhibit T-cell proliferation, and remodel glycans on host proteins (1). This last activity is typically carried out by endo-β-*N*-acetylglucosaminidases (endoglycosidases), which release N-linked oligosaccharides from glycoproteins by cleaving the β (1–4) glycosidic bond between two *N*-acetyl glucosamine (GlcNAc) residues of the *N,N′*-diacetylchitobiose core. X-ray crystal structures of numerous bacterial endoglycosidases (2–7) have shown that these enzymes adopt a common $(\beta/\alpha)_8$ barrel conformation—a cyclic eightfold repeat comprised of β-strand/loop/α helix motifs in which the parallel β-strands form a central barrel with active site residues located within the open barrel structure. Diversity in the loops connecting the β-strands and α-helices define their specificities for both glycan and protein components of a given substrate.

S. pyogenes secretes a 108-kDa endoglycosidase, EndoS, that specifically hydrolyzes core glycans on human IgG antibodies (8). EndoS has enzymatic activity on natively folded IgG, but not on denatured IgG (9). This activity contributes to increased survival of *S. pyogenes* in human blood ex vivo, on account of reduced IgG binding to Fc γ receptors and impaired complement pathway activation (10). Injection of EndoS into mice results in the efficient removal of IgG-associated carbohydrate, with a murine IgG subclass specificity of IgG1 = IgG2b > IgG2a (11). EndoS releases the glycan linked to residue Asn297 of the human Fc region CH_2 domain, which affects the local structure of IgG (12, 13) and its ability to bind complement factor C1q (14) and Fc γ receptors (15). These binding events regulate two key effector functions induced by IgG antibodies.

The same properties of EndoS that benefit the bacteria can be leveraged for the treatment of autoimmune diseases. When used as an in vivo modulator of IgG glycosylation and effector function activity, EndoS has successfully treated numerous autoimmune conditions in animal models (11, 16–22). Because EndoS is specific for IgG bearing complex-type versus high-mannose carbohydrates, it can also potentially be used to enhance the in vivo efficacy of IgG monoclonal antibodies, when produced

Significance

Because bacteria colonize hostile environments they have evolved immune evasion mechanisms, including the expression of enzymes that specifically modify host immune system proteins. *Streptococcus pyogenes* secretes an enzyme called EndoS that removes carbohydrates specifically from human antibodies, impairing their ability to activate immune defenses. Because of its high substrate specificity, EndoS is also being developed as a treatment for autoimmune diseases and is a key enzyme used in the production of antibodies bearing customized carbohydrates. We have determined the three-dimensional structure of EndoS and present a molecular model depicting how EndoS engages antibodies with high specificity. Our data provide a roadmap for engineering EndoS variants with unique activities for clinical and biotechnological applications.

Author contributions: B.T., J.V.L., B.G.P., Z.W., L.-X.W., and E.J.S. designed research; B.T., J.V.L., B.G.P., L.G.C., and T.M.W. performed research; B.T., J.V.L., B.G.P., L.G.C., J.P.G., T.M.W., and L.-X.W. contributed new reagents/analytic tools; B.T., J.V.L., B.G.P., L.G.C., S.G., J.P.G., G.A.S., T.M.W., Z.W., L.-X.W., and E.J.S. analyzed data; and B.T., J.V.L., B.G.P., L.G.C., S.G., G.A.S., T.M.W., Z.W., L.-X.W., and E.J.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 4NUY and 4NUZ).

[1]To whom correspondence should be addressed. E-mail: esundberg@ihv.umaryland.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1322908111/-/DCSupplemental.

recombinantly with high-mannose glycans, by reducing competition for Fc γ receptor binding from serum antibodies (23).

Endoglycosidases, including EndoS, have also been used extensively for in vitro glycan remodeling to modulate the properties of glycoproteins (24). Natural glycoproteins exist as mixtures of glycoforms, of which only one or a few typically exhibit maximal activity. However, individual glycoforms are difficult to purify (25) and recombinant expression of glycoproteins yields heterogeneous glycoforms, even when glycotransferase mutant cell lines are used (26, 27). To circumvent these limitations on glycoprotein homogeneity, recombinant protein expression combined with chemoenzymatic glycan remodeling (28) has been developed. IgG monoclonal antibodies are used extensively as therapeutics and their activities, as mediated by effector functions, depend on the chemistry of their core glycans. EndoS deglycosylates antibody glycoforms that are refractory to processing by other endoglycosidases (29) and glycosynthase mutants of EndoS efficiently transfer predefined N-glycans to intact IgG (30). Together, these catalytic properties of EndoS enzymes allow for customization of IgG glycoforms that can enhance the therapeutic capacities of monoclonal antibodies.

Here, we report the X-ray crystal structure of EndoS and a model of its encounter complex with IgG1 Fc. These findings reveal the structural determinants of EndoS specificity for IgG antibodies, provide a molecular mechanism for its enzymatic activity, and suggest ways in which novel endoglycosidases could be engineered for clinical and biotechnological applications.

Results

Oligomerization and Hydrolytic Activity of EndoS. Using analytical size exclusion chromatography, we determined the oligomerization state of $EndoS_{WT}$(37–995), missing the N-terminal signal peptide (residues 1–36) but containing a putative coiled coil (residues 37–97), as ∼20:80 dimer:monomer, whereas $EndoS_{WT}$(98–995) is entirely monomeric in solution (*SI Appendix*, Fig. S1*A*). We examined the Fc glycan hydrolysis activity of these EndoS proteins by SDS/PAGE analysis using the IgG1 monoclonal antibody Rituximab as a substrate. For $EndoS_{WT}$(37–995), both dimer and monomer fractions exhibited approximately the same hydrolytic activity, whereas the relative activity of $EndoS_{WT}$(98–995) was reduced at least 50-fold (*SI Appendix*, Fig. S1*B*). Thus, whereas the truncated enzyme can process glycosylated IgG1 completely, the putative coiled coil increases the rate of hydrolysis.

Overall Structure of EndoS. We crystallized SeMet-$EndoS_{D233Q}$(98–995), a glycosynthase mutant, and determined its structure at 3.2-Å resolution by multiwavelength anomalous dispersion (MAD). We also crystallized native $EndoS_{D233Q}$(98–995) and $EndoS_{WT}$(98–995) and solved their structures by molecular replacement at 1.9- and 2.6-Å resolution, respectively, using the partially refined MAD-phased $EndoS_{D233Q}$(98–995) structure as a search model (*SI Appendix*, Table S1). With no significant differences in the two structures, we refer to the higher resolution structure throughout the manuscript. The overall morphology of EndoS is that of a letter "V" with a small extension from one end (Fig. 1*A*). The V shape of the protein measures ∼130 Å across and ∼83 Å high, with a tapered cleft measuring ∼42 Å across its opening. EndoS is composed of five distinct protein domains including, from N to C terminus: (*i*) an endoglycosidase enzymatic domain, residues 98–445; (*ii*) a leucine-rich repeat (LRR) domain, residues 446–631; (*iii*) a hybrid Ig domain, residues 632–764; (*iv*) a carbohydrate-binding module (CBM), residues 765–923; and (*v*) a three-helix bundle (3H) domain, residues 924–995 (*SI Appendix*, Fig. S2). The EndoS structure is not fully globular but, instead, has individual domains arranged akin to beads on a string. Using small angle X-ray scattering (SAXS), we confirmed that the observed V shape of the crystallized EndoS was maintained in solution (*SI Appendix*, Fig. S3). The SAXS radial distribution, or p(r), function is bimodal, consistent with the overall V shape of the crystal structure. The maximal diameter is 149 ± 7 Å with two peaks centered ∼40 Å apart, values highly similar to the dimensions of

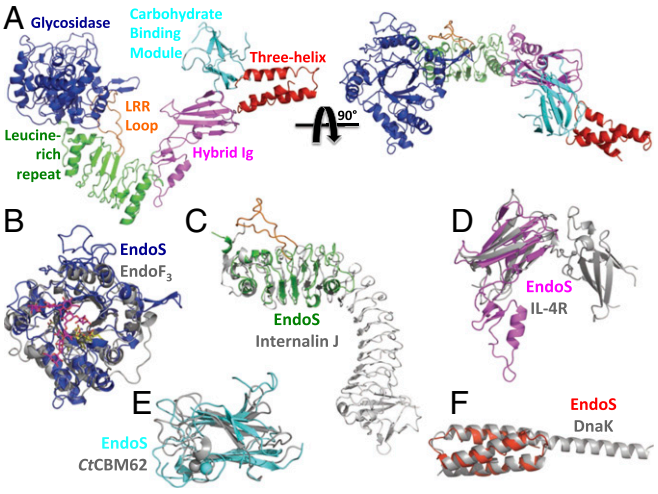


Fig. 1. EndoS structure, domain organization, and comparison with known protein structures. (*A*) Structure of $EndoS_{D233Q}$(98–995) viewed in two orientations, inclusive of the glycosidase (blue), leucine-rich repeat (green), hybrid Ig (magenta), carbohydrate binding module (cyan), and three-helix bundle (red) domains. The loop extending from the middle of the leucine-rich repeat is in orange. Superposition of the EndoS glycosidase domain to that of $EndoF_3$ (*B*), leucine-rich repeat domain to InternalinJ (*C*), hybrid Ig domain to the IL-4 receptor (*D*), carbohydrate binding module to *Ct*CBM62 (*E*), and three-helix bundle domain to DnaK (*F*).

crystallized EndoS. Using CRYSOL (31), we observed a nearly perfect fit of the SAXS data to the solution scattering profile calculated from the crystal structure. We also generated ab initio models of the protein envelope, all of which exhibited an open shape, and we fit the EndoS crystal structure into the top-ranked model with a correlation coefficient of 0.87.

The Glycosidase Domain Exhibits a Common Fold with Unique Structural Features. The EndoS glycosidase domain adopts the $(\beta/\alpha)_8$ barrel conformation typical of bacterial endoglycosidases. Structurally, it is most similar (32) to the analogous domain of $EndoF_3$ (PDB ID code 1E0M; Dali Server Z score = 21.9; Fig. 1*B*). EndoS cleaves specifically biantennary complex oligosaccharides (29), whereas $EndoF_3$ is specific for both biantennary and triantennary complex oligosaccharides (6); neither have activity on high mannose oligosaccharides. The basic structure of the $(\beta/\alpha)_8$ barrel fold and the positions of the active site residues are nearly identical in these two enzymes (*SI Appendix*, Fig. S4*A*). The loops surrounding the active site barrel opening, however, differ substantially between EndoS and $EndoF_3$, resulting in distinct molecular surfaces that could accommodate divergent glycoprotein substrates (Fig. 2*A*). In EndoS, tryptophan residues are important for both substrate specificity and enzymatic activity (33), several of which are exposed on the surface surrounding the active site barrel opening (*SI Appendix*, Fig. S4*B*), including: Trp314 and Trp358, both of which line the groove that likely accommodates the protein portion of the substrate, and Trp121, Trp153, and Trp161, which line the two grooves in which the oligosaccharide chains in the biantennary complex carbohydrate are likely positioned during hydrolysis.

The LRR Domain Extends Additional Domains Away from the Catalytic Site. The LRR domain of EndoS extends from the bottom of the glycosidase domain, opposite the $(\beta/\alpha)_8$ barrel opening. This domain forms the base of the V-shape structure, keeping the remaining C-terminal domains from folding back onto the glycosidase domain. The LRR domain contains three typical LRR motifs followed by two modified LRR motifs, all of which are capped by several short α-helices. It is structurally most similar to internalins, such as InternalinJ (PDB ID code 3BZ5;

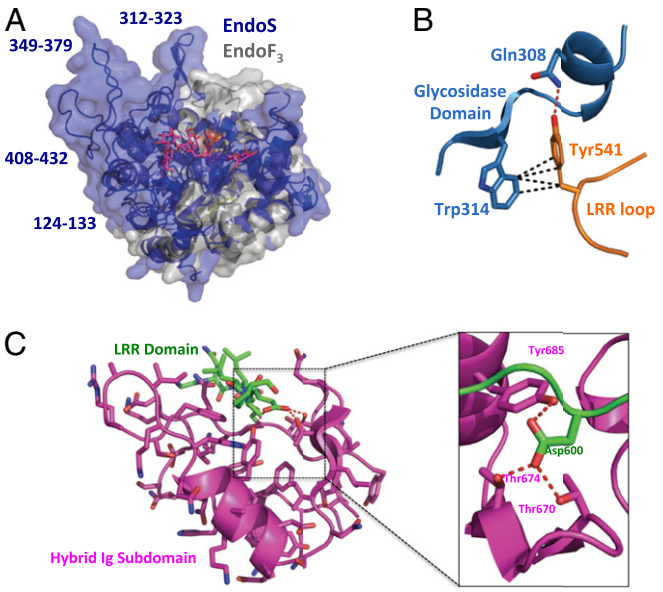


Fig. 2. Structural features of the EndoS glycosidase, leucine-rich repeat, and hybrid Ig domains. (*A*) Superposition of the EndoS glycosidase domain (blue) with that of $EndoF_3$ (gray) with surfaces rendered. The glycan structure as resolved in the $EndoF_3$–glycan complex structure is in magenta. EndoS glycosidase domain loops that diverge most in structure from those in $EndoF_3$ are labeled. (*B*) Intramolecular contacts between residues in the EndoS glycosidase domain (blue) and the apex of the loop extending from the leucine-rich repeat domain (orange). Van der Waals interactions are shown as black dashed lines; hydrogen bond as a dashed red line. (*C*) Intramolecular engagement of residues from the leucine-rich repeat (green) and the inserted hybrid Ig subdomain (magenta). (*C*, *Left*) Interaction showing the entire hybrid Ig subdomain. (*C*, *Right*) Close-up view of contacts made by leucine-rich repeat residue Asp600 (green) with hybrid Ig subdomain residues. Hydrogen bonds are shown as dashed red lines.

Z score = 9.5; Fig. 1*C*). Despite its limited number of repeat motifs, the EndoS LRR adopts the typical curvature of longer LRR domains.

A Unique LRR Domain Loop Packs Against the Glycosidase Domain. Within the third LRR motif, a long loop (residues 528–554) extends away from the LRR domain and packs against the side of the glycosidase domain. Residues 534–541 bury 1,112 $Å^2$ of surface area between this loop and the glycosidase domain, approximately one-half of the total 2,173 $Å^2$ of buried surface area between the glycosidase and LRR domains. Known LRR motifs do not typically include such loop insertions. Tyr541, at the tip of the LRR loop, forms a hydrogen bond with Gln308 and several van der Waals interactions with Trp314 of the glycosidase domain (Fig. 2*B*). This latter residue, as described above, forms the edge of the surface groove on the top of the glycosidase domain that leads to the $(\beta/\alpha)_8$ barrel opening and the active site. The LRR domain loop effectively extends this groove from the active site and, thus, could potentially play a role in glycoprotein substrate specificity and/or the stability of the protein.

EndoS Exhibits a Topologically Unique Hybrid Ig Domain. Extending from the LRR domain at the base of the V-shaped EndoS structure is a hybrid Ig domain. This domain is composed of two subdomains that are topologically entwined—the smaller of the two subdomains is inserted within the loop that connects the second and third β-strands of the larger subdomain, which is a typical Ig domain structurally similar to the interleukin-4 receptor (PDB ID code 1IAR; Z score = 5.2; Fig. 1*D*) and other Ig proteins, including antibodies. The smaller subdomain, conversely, is structurally unique compared with all previously determined protein structures. It acts as a molecular spacer between the LRR domain and the Ig subdomain, the interface between which consists primarily of a loop from the former (residues 598–603) positioned into a deep cleft on the surface of the latter (Fig. 2*C*), burying 685 $Å^2$ of the 1,562 $Å^2$ of surface area between these two domains. Although most of this interface is comprised of van der Waals interactions, the two terminal oxygen atoms of Asp600 make hydrogen bonds to side-chain oxygen atoms from Thr670, Thr674, and Tyr685 (Fig. 2*C*).

EndoS Contains a Putative Carbohydrate-Binding Module. C-terminal to the hybrid Ig domain is a domain that exhibits a high degree of structural homology to noncatalytic CBMs from *Clostridium thermocellum*, including *Ct*CBM62 (PDB ID code 2YFU; Z score = 7.4; Fig. 1*E*), which binds galactose-containing polysaccharides. Complex carbohydrates, such as those attached to IgG1 Fc and processed by EndoS, typically contain galactose molecules (12). Like *Ct*CBM62, the EndoS CBM coordinates a single Ca^{2+} ion by using main-chain oxygen atoms of Lys786, Gly790, Gln791, and Pro915, and a side-chain oxygen atom from Glu916. Neither the oligomerization state nor the hydrolytic activity of EndoS changed in the presence of EDTA. The sole substrate for EndoS is the glycosylated Fc region of IgG (8), an obligate homodimer with Asn-linked glycans attached to each of the two monomer subunits. In the EndoS structure, the two domains, glycosidase and CBM, that most likely bind IgG glycans are located at opposite ends of the V-shaped molecule (Fig. 1*A*).

The C-Terminal Domain Is a Three-Helix Bundle Pointing Away from All Other Domains. The final EndoS protein domain, 3H, is a three-helix bundle motif most structurally similar to the substrate binding domain of the Hsp70 chaperone DnaK (PDB ID code 4JNE; Z score = 8.6; Fig. 1*E*). The 3H domain extends away from the opening of the V-shaped structure (Fig. 1*A*). However, this domain is not packed extensively against its neighboring EndoS domains, the CBM and hybrid Ig domains, exhibiting 841 and 365 $Å^2$ of buried surface area with each domain, respectively, suggesting that it could potentially rotate and translate in a rigid-body fashion when substrate is bound.

EndoS Requires Multiple Domains for Substrate Specificity. We expressed catalytically inactive versions of the enzyme, $EndoS_{E235Q}$(98–995), which includes all five domains, EndoS (446–995), which contains only domains C-terminal to the glycosidase domain, and $EndoS_{E235Q}$(98–764), which excludes the CBM and 3H domains, and measured their binding affinities to immobilized IgG1 Fc by surface plasmon resonance (SPR) analysis (*SI Appendix*, Fig. S5). $EndoS_{E235Q}$(98-995) bound Fc with an affinity (K_D) of 22 μM, ∼50-fold lower than the oligomerization-dependent avidity of $EndoS_{D235Q}$(37–995) (34), essentially equivalent to its relative reduction in hydrolytic activity (*SI Appendix*, Fig. S1*B*). Neither EndoS(446–995) nor $EndoS_{E235Q}$(98–764) exhibited detectable binding to IgG1 Fc. Thus, there exist at least two distinct domains important for the EndoS specificity for IgG1 Fc: (*i*) between residues 98 and 446, corresponding to the glycosidase domain alone; and (*ii*) and between residues 766 and 995, inclusive of the CBM and 3H domains.

A Model of the EndoS/IgG1 Fc Encounter Complex. We modeled the EndoS/IgG1 Fc encounter complex by docking an IgG1 Fc homodimer structure (PDB ID code 4BYH) onto the $EndoS_{D233Q}$(98–995) crystal structure, modified by adding any missing loop residues using ZDOCK 3.0.2 (35). Postprocessing of the docking output was performed in two stages. First, we analyzed the top 2,000 ZDOCK predictions for proximity of the two $Asn297_{IgG1\ Fc}$ residues to EndoS glycosidase and CBM domain residues likely involved in glycan binding based on homology to other glycosidase and CBM structures (e.g., residues $Leu213_{EndoS}$, $Gln233_{EndoS}$, $Glu235_{EndoS}$, and $Asp827_{EndoS}$, $Arg830_{EndoS}$, $Asn836_{EndoS}$, respectively). One of the top-ranked ZDOCK models (ZD3, ranked number 3; *SI Appendix*, Fig. S6*A*) featured

the two $Asn297_{IgG1\ Fc}$ residues with relatively small distances to both of these sets of EndoS CBM residues (<10 Å) and glycosidase domain residues (<20 Å). Overall, the vast majority of the initial-stage docking predictions featured $Asn297_{IgG1\ Fc}$ residues positioned near the EndoS CBM (Fig. 3*A*). Second, we performed structural refinement of this model, generating an additional 1,000 refined models. Comparing the refined models with the original model from ZDOCK, we observed an energy funnel-like landscape (*SI Appendix*, Fig. S6*B*). In this final model of the EndoS/IgG1 Fc encounter complex (Fig. 3*B*), the IgG1 Fc is bound predominantly through residues from its CH_2 domains and primarily to the EndoS CBM in an orientation in which the plane of the V shape of EndoS is nearly perpendicular to the plane that runs through the two monomer subunits of the Fc homodimer. The EndoS glycosidase domain loop comprised of residues 312–323, unique among all known endoglycosidase structures, bisects the two monomer subunits of the Fc homodimer (Fig. 3*C*). Positioned as such, there is direct access to the Asn297-linked glycans for the EndoS active site and adequate space between the Fc monomer N-termini and the LRR domain for the remainder of the IgG1 antibody, the Fab domains, to exit. Computational alanine scanning mutagenesis identified CBM domain residues $Trp803_{EndoS}$, $Arg908_{EndoS}$, and $Glu833_{EndoS}$ as hot spots for binding.

Validating the Encounter Complex Model. Our model of the EndoS/IgG1 Fc encounter complex indicated several EndoS loops that likely interact with the glycoprotein substrate, including the aforementioned glycosidase domain loop (residues 312–323) and loops from the hybrid Ig (residues 742–750) and CBM (residues 793–797) domains (*SI Appendix*, Fig. S6*C*). To determine the energetic contributions to binding for each of these loops, and the LRR domain loop (residues 528–554), we produced versions of $EndoS_{E235Q}$ with each loop removed and measured their binding affinities to IgG1 Fc by SPR, all of which exhibited reduced affinities for Fc relative to $EndoS_{E235Q}$ (*SI Appendix*, Fig. S5). We observed the largest reduction in binding affinity, K_D = 489 μM (~20-fold), for $EndoS_{E235Q/\Delta 314–323}$. $EndoS_{E235Q/\Delta 742–750}$ bound Fc with a K_D of 156 μM (~10-fold reduction). $EndoS_{E235Q/\Delta 528–554}$ and $EndoS_{E235Q/\Delta 793–797}$ bound Fc with K_D values of 24 and 30 μM, respectively (approximately equivalent to $EndoS_{E235Q}$). For $EndoS_{E235Q/\Delta 528–554}$, we obtained markedly lower protein expression and observed significantly reduced R_{max} values for SPR binding isotherms, suggesting that the protein is unstable without this loop. Taken together, these data indicate that loops from the glycosidase and hybrid Ig domains are major determinants of Fc specificity, whereas loops from the LRR domain, which makes no contacts but may be important for the stability of the protein, and the CBM, which makes predominantly glycan interactions, in the encounter complex model are nearly dispensable for Fc binding. We also tested mutants identified by computational alanine scanning mutagenesis as hot spots for binding, including $EndoS_{E235Q/W803A}$, $EndoS_{E235Q/R908A}$, and $EndoS_{E235Q/E833A}$ (*SI Appendix*, Fig. S5), all located in the CBM. $EndoS_{E235Q/W803A}$ and $EndoS_{E235Q/E833A}$ each exhibited no detectable binding to IgG1 Fc at this concentration. $EndoS_{E235QR908A}$ exhibited threefold weaker binding than did $EndoS_{E235Q}$. The binding capacities of these EndoS mutants largely correlated with their hydrolytic activities (*SI Appendix*, Fig. S1*C*). Each mutant that exhibited substantially reduced or no detectable binding also exhibited reduced or no hydrolytic activity, except $EndoS_{E235Q/E833A}$. In particular, the glycosidase domain loop (residues 314–323) and a point mutation in the CBM domain, $EndoS_{E235Q/E803A}$, were absolutely required for hydrolytic activity.

Conformation Changes Required for Enzymatic Activity. According to our modeled EndoS/IgG1 Fc encounter complex, which is transient by definition, there is a clear requirement for conformational changes to carry out the enzymatic reaction. In the encounter complex model, the active site residues are located

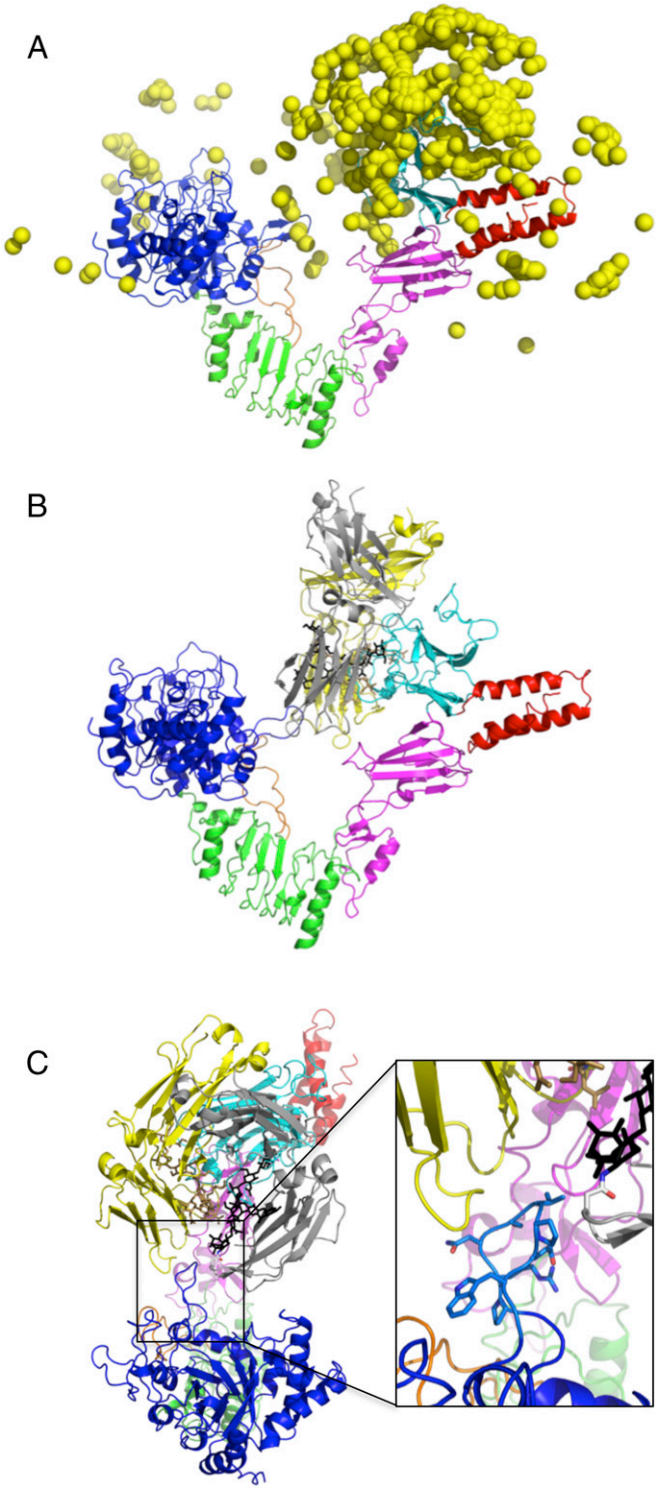


Fig. 3. Model of the EndoS/IgG1 Fc encounter complex. (*A*) Distribution of location of the IgG1 Fc on the EndoS structure from the top 500 initial models of the encounter complex. Colors of EndoS domains are as in previous figures. Each yellow ball represents the center of mass of a single IgG1 Fc dimer. (*B*) Side view of the final refined EndoS/IgG1 Fc encounter complex model. Monomers of the Fc homodimer are in gray (with corresponding glycan in black) and in yellow (with corresponding glycan in tan). (*C*) Top view of the encounter complex model with close-up of the EndoS glycosidase loop bisecting the Fc monomers in shown in the box on the right.

some 15 Å from the bond between the first and second GlcNAc moieties at which enzymatic cleavage occurs. Thus, rotation and translation of the glycosidase domain relative to the likely more fixed position of the IgG1 Fc bound to the CBM must occur for catalysis to take place. Using HingeProt (36), we found several EndoS regions predicted to act as hinges, including: (*i*) the residues linking the glycosidase and LRR domains; (*ii*) the N-terminal portion of the LRR domain; and (*iii*) the juncture between the two subdomains of the hybrid Ig domain. Concerted movements of EndoS domains around these hinge points would result in a narrowing of its V-shape opening, moving the glycosidase domain closer in space to the CBM (*SI Appendix*, Fig. S7*A*). Regardless of the structure of the encounter complex and ensuing conformational changes in EndoS, further conformational changes must take place in the IgG substrate structure to allow the glycan to be properly positioned in the active site for cleavage. Superpositions of IgG1 Fc with the $EndoF_3$–glycan complex by aligning their respective glycans, and of EndoS with the $EndoF_3$–glycan complex by aligning their respective glycosidase domains, indicates that substantial steric clashes would prohibit such a complex (*SI Appendix*, Fig. S7*B*). Barring remarkable structural deformation of the EndoS glycosidase domain, protein and/or glycan conformation changes in the antibody would be required for enzymatic activity.

Human IgG Subclass Specificity. Because EndoS is known to have distinct specificities for individual murine IgG subclasses (11), we determined whether there existed any comparable human IgG subclass specificity. We found that both EndoS(37-995) and EndoS(98-995) exhibited slightly decreased binding to IgG2, IgG3, and IgG4, and reduced hydrolytic activity with these substrates (*SI Appendix*, Fig. S1*C*), compared with IgG1. We also tested the hydrolytic activity of our entire panel of EndoS(98-995) mutants and found that, nearly universally, mutants capable of hydrolyzing IgG1 were also capable of hydrolyzing IgG2, IgG3, and IgG4 (*SI Appendix*, Fig. S1*C*), and those that could not hydrolyze IgG1 could also not hydrolyze any other IgG subclass. However, each of the hydrolytic EndoS mutants exhibited higher activity on IgG1 relative to IgG2, IgG3, and IgG4. These data indicate that EndoS has a human IgG subclass specificity of IgG1 > IgG2 = IgG3 = IgG4.

Discussion

EndoS is an enzyme secreted by *S. pyogenes* that removes carbohydrates highly specifically from human IgG antibodies. Because antibodies are central players in many human immune responses and bridge the innate and adaptive arms of immunity, the analysis and manipulation of the enzymatic activity of EndoS impacts diverse fields in biomedicine. Clinically, EndoS contributes to the abilities of *S. pyogenes* to evade the human immune response (8, 10, 34); the development of specific inhibitors of EndoS activity could improve clinical outcomes of patients suffering a range of inflammatory conditions. Therapeutically, EndoS is already showing great promise in animal models as a treatment for diverse autoimmune diseases that rely on autoantibodies (11, 16–22); fine-tuning the specificity and activity of EndoS will be an important aspect of its further development as a protein therapeutic for use in humans. Biotechnologically, EndoS is a unique glycoprotein-modifying enzyme with the capacity to both remove glycans from and, as a glycosynthase variant, attach glycans to antibodies (29, 30); expanding the repertoire of homogeneous glycosylated antibodies that can be produced with newly designed EndoS variants will be critical for realizing the full potential of engineered antibodies. Our X-ray crystal structure of EndoS provides a platform for future clinical, therapeutic, and biotechnological progress.

EndoS is comprised of five distinct protein domains that adopt a striking V-shaped structure in which the four domains C-terminal to the glycosidase domain extend from the face opposite that used for glycan binding and catalytic activity. Instead of forming a typical globular protein, the additional domains extend away from the glycosidase domain, resulting in the CBM forming one point of the V shape opposite the point formed by the glycosidase domain. The relative positions of these two domains in EndoS known to engage carbohydrate in other proteins suggests that glycosylated antibodies enter the V shape and are held within the concave surface of the enzyme during processing.

Indeed, our molecular modeling predicts that an encounter complex is formed between EndoS and human IgG1 Fc in which the substrate enters the V and is trapped by the CBM, with EndoS making contacts mostly to residues from the IgG1 CH_2 domain, in which the N-linked glycan resides. A notable feature of this complex is that the tip of the glycosidase domain loop (residues 312–323) that extends away from the active site barrel opening bisects the two monomer subunits of the IgG1 Fc homodimer. When we measured the binding affinity of the glycosynthase mutant with this loop removed, we observed an ∼20-fold weaker affinity relative to the same protein with this loop intact and complete abolition of hydrolytic activity. Additional components of the EndoS structure implicated by the model in substrate binding and validated by mutagenesis include a loop from the hybrid Ig domain (residues 742–750) and several hot spot residues from the CBM.

Evident from both our structure of the enzyme in the absence of substrate and our model of the EndoS/IgG1 Fc encounter complex, enzymatic activity would require a number of conformational changes, likely in both the enzyme and the substrate. Accordingly, we found, by computational analysis, several hinge points between EndoS domains that predict concerted movements that effectively squeeze the ends of the V shape together, bringing the glycosidase and CBM domains in closer proximity to one another. In the encounter complex model, in which IgG1 Fc is bound predominantly by the CBM, this movement would result in the substrate approaching the active site. Additional conformational changes likely to accompany transition from the encounter complex to a processing–competent complex structure include movement of the glycosidase loop (residues 312–323) that bisects the two IgG monomers and rotation and further translation of the glycosidase domain toward the bound substrate. Regardless of these movements, conformational changes in the substrate are certainly also required for enzymatic activity. Assuming that the N-linked glycan of IgG must eventually occupy a position similar to that observed for the $EndoF_3$–glycan complex (6), the IgG1 Fc homodimer would have to separate and/or the glycans would have to move outside of the space between the CH_2 domains that they normally occupy. Whether such concerted enzyme and substrate structural changes would allow the processing of both glycans subsequent to the formation of a single encounter complex or necessitate a unique encounter complex event for the processing of each of the two glycans remains unclear at this point.

EndoS exhibits high specificity for glycosylated IgG antibodies and has no, or only marginal, hydrolytic activity toward N-glycans in the context of other glycoproteins or glycopeptides. Our studies indicate that several structural features scattered throughout the EndoS sequence are important for this high substrate specificity. First, the glycosidase loop (residues 312–323) extending from the barrel opening that engages IgG1 in our encounter complex model has a significant effect on substrate binding and activity. The nature of this loop suggests that simply swapping out glycosidase domains from other endoglycosidases to engineer EndoS variants that maintain IgG specificity but exhibit diverse glycan specificities may not be achievable without incorporating a similar loop in the engineered enzyme. Second, EndoS domains beyond the glycosidase domain are important for substrate specificity. The importance of these domains has previously been implicated by the deactivation of EndoS by SpeB, which releases the glycosidase domain from the remainder of the enzyme (33). That these noncatalytic domains are important for substrate specificity suggests that modifying them might produce EndoS variants with unique substrate specificities, such as those that exhibit distinct human IgG subclass specificity or altered antibody isotype binding.

In summary, we present the X-ray crystal structure of EndoS and a model of the encounter complex formed with IgG1 Fc. Our studies provide a molecular basis for substrate specificity and enzymatic activity and suggest ways in which EndoS variants could be engineered for novel therapeutic and biotechnological purposes.

Materials and Methods

For details, see *SI Appendix, SI Materials and Methods*.

Protein Production, Oligomerization State, and Hydrolytic Activity Analysis. All EndoS proteins were expressed in *Escherichia coli* and purified as described (30, 37). IgG1 Fc was obtained from papain digestion and chemoenzymatic glycoengineering of Rituximab as described (30). IgG2, IgG3, and IgG4 were purchased (Sigma-Aldrich). Oligomerization state was assessed by size exclusion chromatography in comparison with molecular weight standards. Hydrolytic activity was calculated by measuring densitometric changes in high molecular weight glycosylated IgG antibodies over time by using SDS/PAGE.

Protein Structure Analysis. Crystallization of native EndoS by liquid-liquid diffusion using Crystal Formers (Microlytic) has been described in detail (37). SeMet-EndoS$_{D233Q}$(98–995) crystals were obtained by microseeding with native crystals in the previously determined crystallization condition. All crystals were flash cooled at 100 K in mother liquor containing 20% ethylene glycol before synchrotron X-ray diffraction data collection. The structure of SeMet-EndoS$_{D233Q}$(98–995) was solved by MAD at 3.2 Å, whereas those of 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 the SeMet-EndoS$_{D233Q}$(98–995) structure as a model. SAXS data were measured for EndoS(98–995) at concentrations of 1, 2, 5, and 10 mg/mL.

Binding Analysis. All SPR experiments were performed by using a Biacore T100 instrument (GE Healthcare) with IgG1 Fc or whole IgG2, IgG3, or IgG4 antibodies immobilized to two flow cell surfaces with one (i.e., negative control) deglycosylated by EndoS pretreatment. Concentration series of all EndoS proteins were injected and affinity constants were calculated by using a general steady-state equilibrium model.

Molecular Modeling. We used ZDOCK 3.0.2 (35) to dock the glycosylated IgG1 Fc (PDB ID code 4BYH) to the EndoS$_{D233Q}$(98–995) structure. We performed computational interface alanine scanning of the refined model to determine putative energetic hot spots by using the "interface" protocol of Rosetta 2.0.2 (38).

ACKNOWLEDGMENTS. We thank the beamline scientists at 23-ID-D and 23-ID-B, Advanced Photon Source (APS) and at 11-1, Stanford Synchrotron Radiation Lightsource (SSRL). These studies were supported by National Institutes of Health Grants R01AI090866 (to E.J.S.) and R01 GM080374 (to L.-X.W.). Use of APS is supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract DE-AC02-06CH11357. Portions of this research were carried out at SSRL, a national user facility operated by Stanford University on behalf of the US Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research, and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program.

1. Collin M, Olsén A (2003) Extracellular enzymes with immunomodulating activities: Variations on a theme in Streptococcus pyogenes. *Infect Immun* 71(6):2983–2992.
2. Ling Z, et al. (2009) The X-ray crystal structure of an Arthrobacter protophormiae endo-beta-N-acetylglucosaminidase reveals a (beta/alpha)(8) catalytic domain, two ancillary domains and active site residues key for transglycosylation activity. *J Mol Biol* 389(1):1–9.
3. Yin J, et al. (2009) Structural basis and catalytic mechanism for the dual functional endo-beta-N-acetylglucosaminidase A. *PLoS ONE* 4(3):e4658.
4. Abbott DW, Macauley MS, Vocadlo DJ, Boraston AB (2009) Streptococcus pneumoniae endohexosaminidase D, structural and mechanistic insight into substrate-assisted catalysis in family 85 glycoside hydrolases. *J Biol Chem* 284(17):11676–11689.
5. Van Roey P, Rao V, Plummer TH, Jr., Tarentino AL (1994) Crystal structure of endo-beta-N-acetylglucosaminidase F1, an alpha/beta-barrel enzyme adapted for a complex substrate. *Biochemistry* 33(47):13989–13996.
6. Waddling CA, Plummer TH, Jr., Tarentino AL, Van Roey P (2000) Structural basis for the substrate specificity of endo-beta-N-acetylglucosaminidase F(3). *Biochemistry* 39(27):7878–7885.
7. Rao V, Guan C, Van Roey P (1995) Crystal structure of endo-beta-N-acetylglucosaminidase H at 1.9 A resolution: Active-site geometry and substrate recognition. *Structure* 3(5):449–457.
8. Collin M, Olsén A (2001) EndoS, a novel secreted protein from Streptococcus pyogenes with endoglycosidase activity on human IgG. *EMBO J* 20(12):3046–3055.
9. Collin M, Olsén A (2001) Effect of SpeB and EndoS from Streptococcus pyogenes on human immunoglobulins. *Infect Immun* 69(11):7187–7189.
10. Collin M, et al. (2002) EndoS and SpeB from Streptococcus pyogenes inhibit immunoglobulin-mediated opsonophagocytosis. *Infect Immun* 70(12):6646–6651.
11. Albert H, Collin M, Dudziak D, Ravetch JV, Nimmerjahn F (2008) In vivo enzymatic modulation of IgG glycosylation inhibits autoimmune disease in an IgG subclass-dependent manner. *Proc Natl Acad Sci USA* 105(39):15005–15009.
12. Arnold JN, Wormald MR, Sim RB, Rudd PM, Dwek RA (2007) The impact of glycosylation on the biological function and structure of human immunoglobulins. *Annu Rev Immunol* 25:21–50.
13. Lund J, Takahashi N, Pound JD, Goodall M, Jefferis R (1996) Multiple interactions of IgG with its core oligosaccharide can modulate recognition by complement and human Fc gamma receptor I and influence the synthesis of its oligosaccharide chains. *J Immunol* 157(11):4963–4969.
14. Krapp S, Mimura Y, Jefferis R, Huber R, Sondermann P (2003) Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity. *J Mol Biol* 325(5):979–989.
15. Nimmerjahn F, Ravetch JV (2008) Fcgamma receptors as regulators of immune responses. *Nat Rev Immunol* 8(1):34–47.
16. Nandakumar KS, et al. (2007) Endoglycosidase treatment abrogates IgG arthritogenicity: Importance of IgG glycosylation in arthritis. *Eur J Immunol* 37(10):2973–2982.
17. Collin M, Shannon O, Björck L (2008) IgG glycan hydrolysis by a bacterial enzyme as a therapy against autoimmune conditions. *Proc Natl Acad Sci USA* 105(11):4265–4270.
18. van Timmeren MM, et al. (2010) IgG glycan hydrolysis attenuates ANCA-mediated glomerulonephritis. *J Am Soc Nephrol* 21(7):1103–1114.
19. Allhorn M, et al. (2010) The IgG-specific endoglycosidase EndoS inhibits both cellular and complement-mediated autoimmune hemolysis. *Blood* 115(24):5080–5088.
20. Tradtrantip L, Ratelade J, Zhang H, Verkman AS (2013) Enzymatic deglycosylation converts pathogenic neuromyelitis optica anti-aquaporin-4 immunoglobulin G into therapeutic antibody. *Ann Neurol* 73(1):77–85.
21. Hirose M, et al. (2012) Enzymatic autoantibody glycan hydrolysis alleviates autoimmunity against type VII collagen. *J Autoimmun* 39(4):304–314.
22. Lood C, et al. (2012) IgG glycan hydrolysis by endoglycosidase S diminishes the proinflammatory properties of immune complexes from patients with systemic lupus erythematosus: A possible new treatment? *Arthritis Rheum* 64(8):2698–2706.
23. Baruah K, et al. (2012) Selective deactivation of serum IgG: A general strategy for the enhancement of monoclonal antibody receptor interactions. *J Mol Biol* 420(1-2):1–7.
24. Wang L-X (2008) Chemoenzymatic synthesis of glycopeptides and glycoproteins through endoglycosidase-catalyzed transglycosylation. *Carbohydr Res* 343(10-11):1509–1522.
25. Rudd PM, et al. (1994) Glycoforms modify the dynamic stability and functional activity of an enzyme. *Biochemistry* 33(1):17–22.
26. Umaña P, Jean-Mairet J, Moudry R, Amstutz H, Bailey JE (1999) Engineered glycoforms of an antineuroblastoma IgG1 with optimized antibody-dependent cellular cytotoxic activity. *Nat Biotechnol* 17(2):176–180.
27. Yamane-Ohnuki N, et al. (2004) Establishment of FUT8 knockout Chinese hamster ovary cells: An ideal host cell line for producing completely defucosylated antibodies with enhanced antibody-dependent cellular cytotoxicity. *Biotechnol Bioeng* 87(5):614–622.
28. Wang LX, Lomino JV (2012) Emerging technologies for making glycan-defined glycoproteins. *ACS Chem Biol* 7(1):110–122.
29. Goodfellow JJ, et al. (2012) An endoglycosidase with alternative glycan specificity allows broadened glycoprotein remodelling. *J Am Chem Soc* 134(19):8030–8033.
30. Huang W, Giddens J, Fan SQ, Toonstra C, Wang LX (2012) Chemoenzymatic glycoengineering of intact IgG antibodies for gain of functions. *J Am Chem Soc* 134(29):12308–12318.
31. Svergun D, et al. (1995) CRYSOL - Evaluation of the solution scattering from macromolecules with known atomic srtructure and fitting to experimental data. *J Appl Cryst* 28:768–773.
32. Holm L, Rosenström P (2010) Dali server: Conservation mapping in 3D. *Nucleic Acids Res* 38(Web Server issue):W545–W549.
33. Allhorn M, Olsén A, Collin M (2008) EndoS from Streptococcus pyogenes is hydrolyzed by the cysteine proteinase SpeB and requires glutamic acid 235 and tryptophans for IgG glycan-hydrolyzing activity. *BMC Microbiol* 8:3.
34. 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.
35. Pierce BG, Hourai Y, Weng Z (2011) Accelerating protein docking in ZDOCK using an advanced 3D convolution library. *PLoS ONE* 6(9):e24657.
36. Emekli U, Schneidman-Duhovny D, Wolfson HJ, Nussinov R, Haliloglu T (2008) HingeProt: Automated prediction of hinges in protein structures. *Proteins* 70(4):1219–1227.
37. Trastoy B, Lomino JV, Wang L-X, Sundberg EJ (2013) Liquid-liquid diffusion crystallization improves the X-ray diffraction of EndoS, an endo-β-N-acetylglucosaminidase from Streptococcus pyogenes with activity on human IgG. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 69(Pt 12):1405–1410.
38. Kortemme T, Baker D (2002) A simple physical model for binding energy hot spots in protein-protein complexes. *Proc Natl Acad Sci USA* 99(22):14116–14121.

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

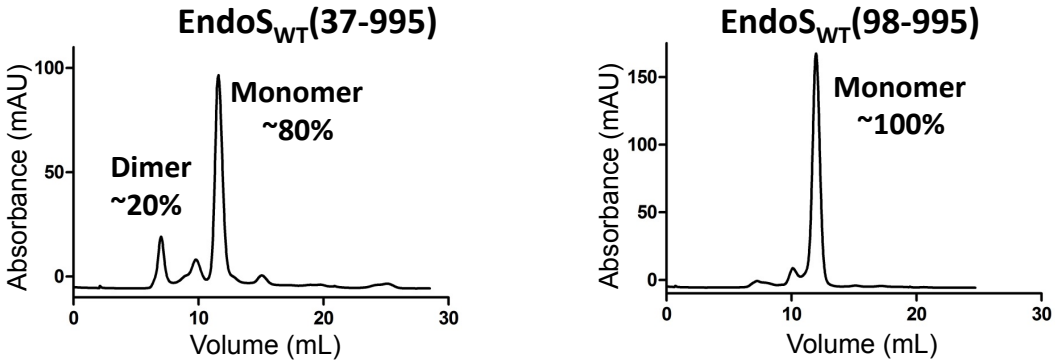
1. Huang W, Giddens J, Fan SQ, Toonstra C, & Wang LX (2012) Chemoenzymatic glycoengineering of intact IgG antibodies for gain of functions. *Journal of the American Chemical Society* 134(29):12308-12318.
2. Trastoy B, Lomino JV, Wang L-X, & Sundberg EJ (2013) Liquid-liquid diffusion crystallization improves 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

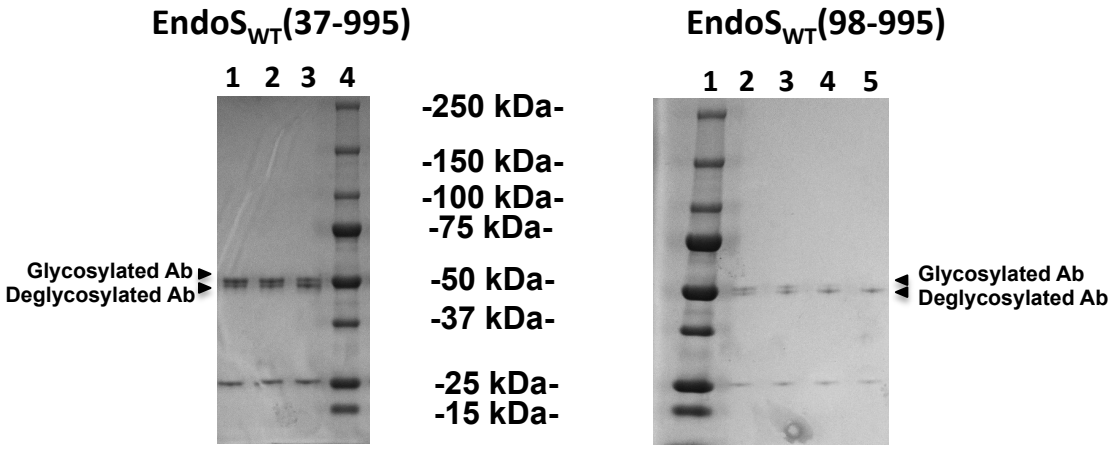
	SeMetEndoS$_{D233Q}$			EndoS$_{D233Q}$	EndoS$_{WT}$
Data collection					
Space group		$P2_12_12_1$		$P2_12_12_1$	$P2_12_12_1$
Cell dimensions					
a, b, c (Å)	86.3, 93.0, 137.8	86.7, 93.5, 138.5	86.2, 92.9, 137.8	92.6, 96.1, 141.2	92.3, 94.5, 142.8
α, β, γ (º)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 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.

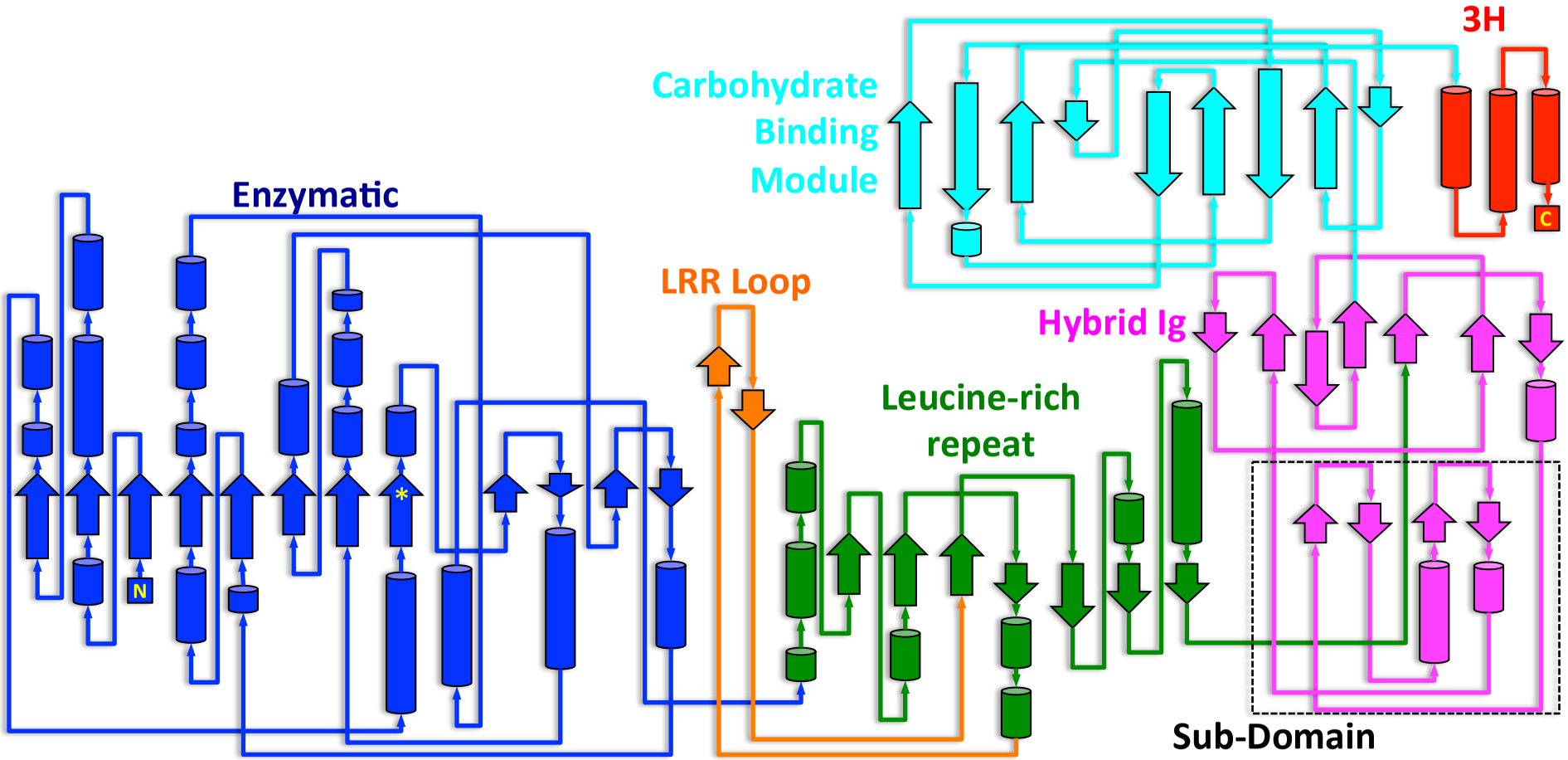


Figure S2. Secondary structure topology of the EndoS(98-995) structure. The N- and C-termini are marked by "N" and "C," respectively; the active site position is marked by a yellow asterisk; α helices are represented by cylinders; β strands are represented by arrows.

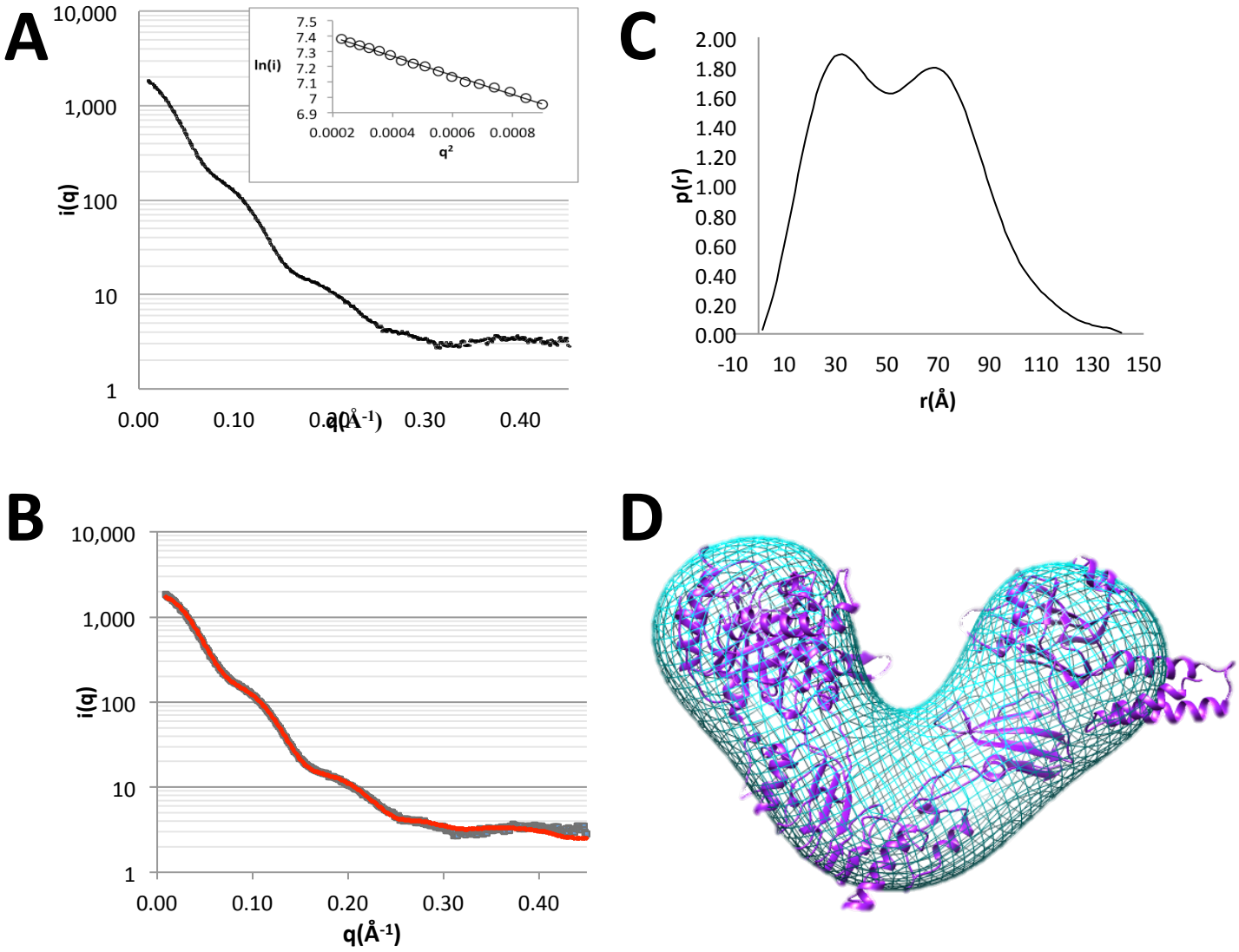


Figure S3. Small angle X-ray scattering (SAXS) analysis of EndoS. **(A)** Data were collected from a 1, 2, 5 and 10 mg/ml concentration series in order to detect concentration-dependent intermolecular interactions. Fifteen 1 s images were averaged using SasTool (http://ssrl.slac.stanford.edu/~saxs/analysis/sastool.htm). 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.

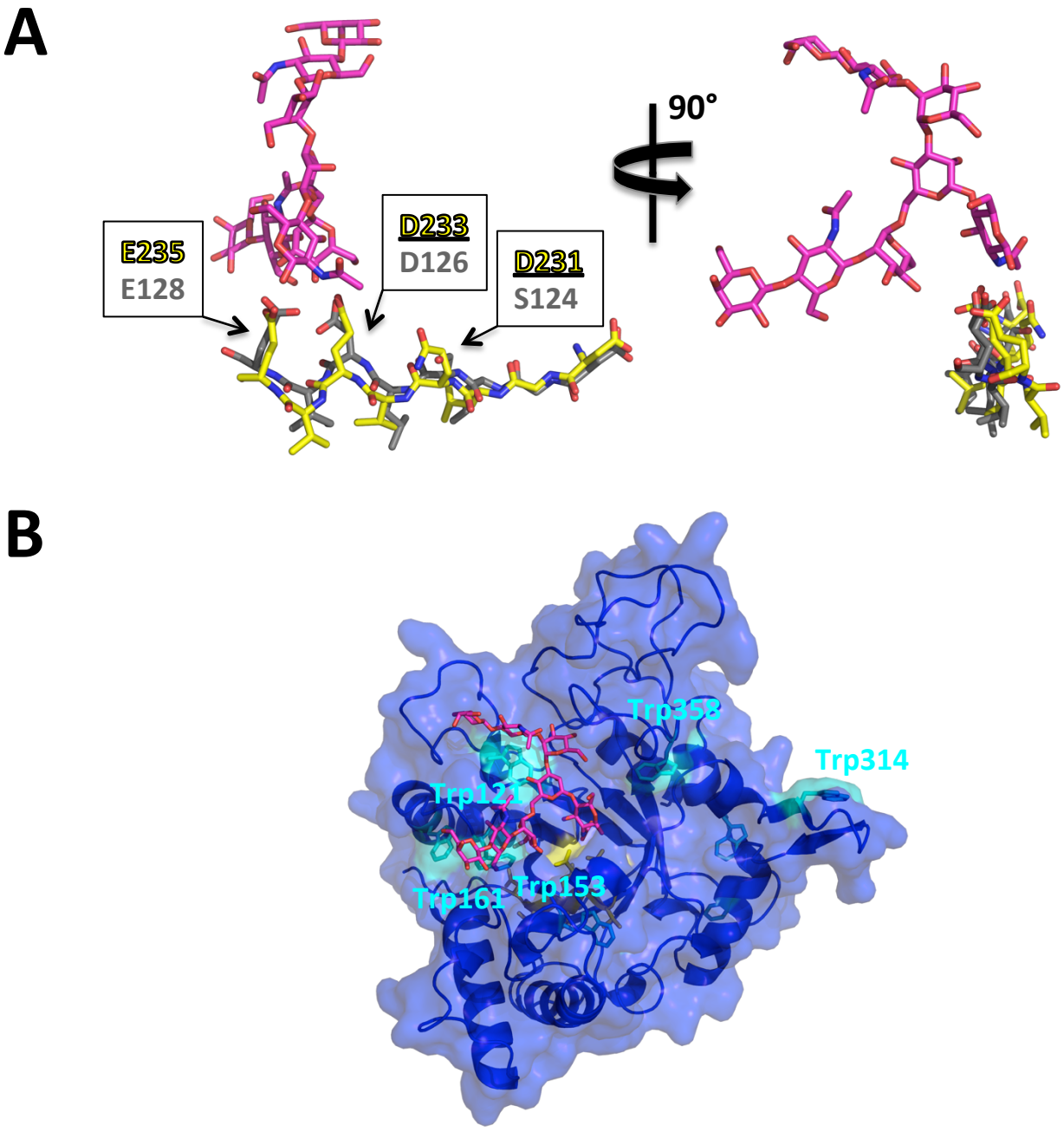


Figure S4. Structural details of the EndoS glycosidase domain. **(A)** Active site residues of EndoS (yellow) align with those of $EndoF_3$ (grey) when their respective glycosidase domains are superimposed. The glycan from the $EndoF_3$-glycan structure is in magenta. **(B)** Cartoon and surface representation 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.

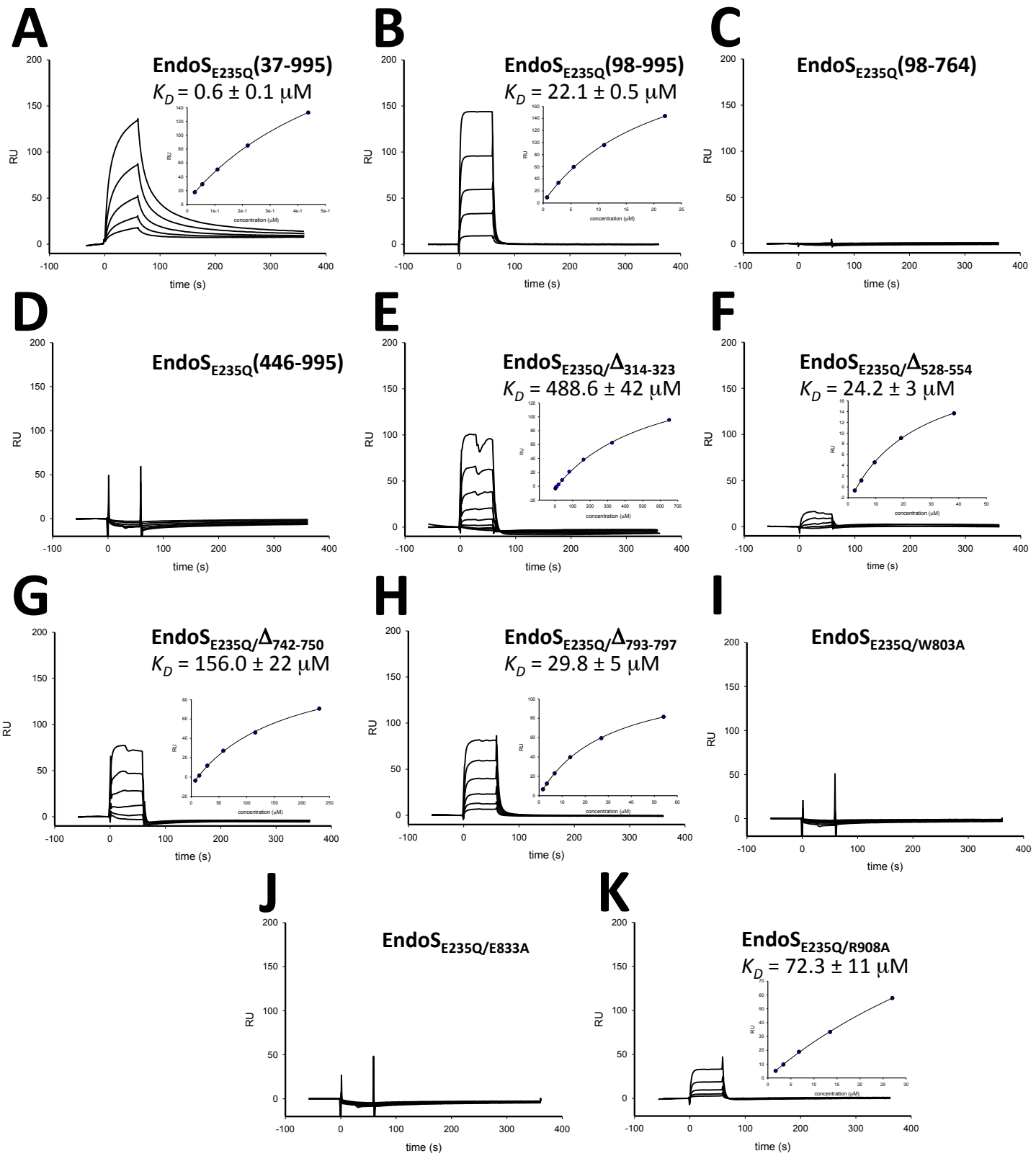


Figure S5. SPR sensograms of (**A**) EndoS$_{E235Q}$(37-995), (**B**) EndoS$_{E235Q}$(98-995), (**C**) EndoS$_{E235Q}$(98-764), (**D**) EndoS(446-995), (**E**) EndoS$_{E235Q/\Delta 314-323}$, (**F**) EndoS$_{E235Q/\Delta 528-554}$, (**G**) EndoS$_{E235Q/\Delta 742-750}$, (**H**) EndoS$_{E235Q/\Delta 793-797}$, (**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.

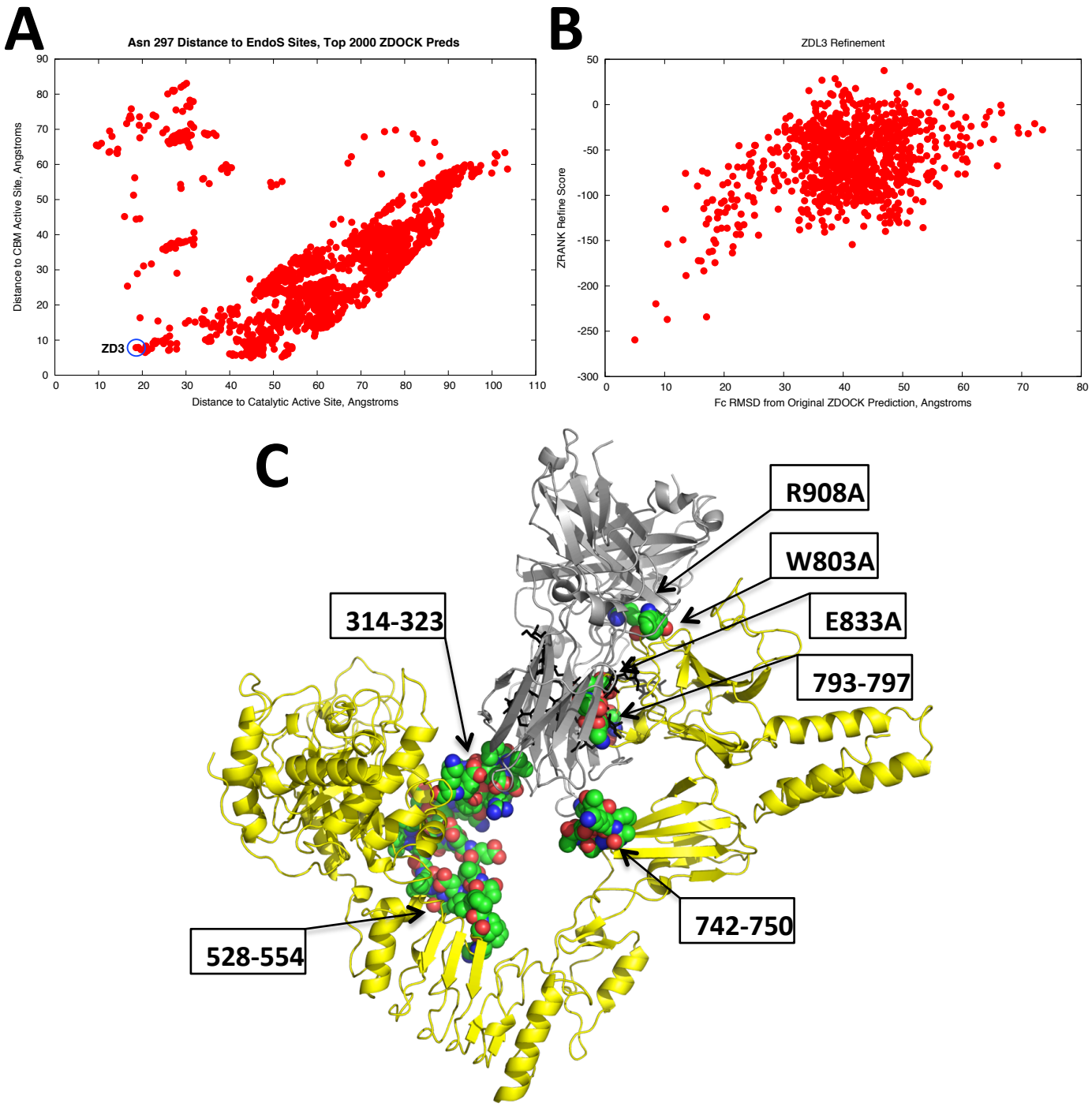


Figure S6. Molecular modeling of the EndoS-IgG1 Fc encounter complex and validation of the model. **(A)** The top 2000 ranked models from the initial modeling plotted in terms of distance of the average Asn297-linked glycan position to EndoS residues likely to bind glycan in the glycosidase (horizontal 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.

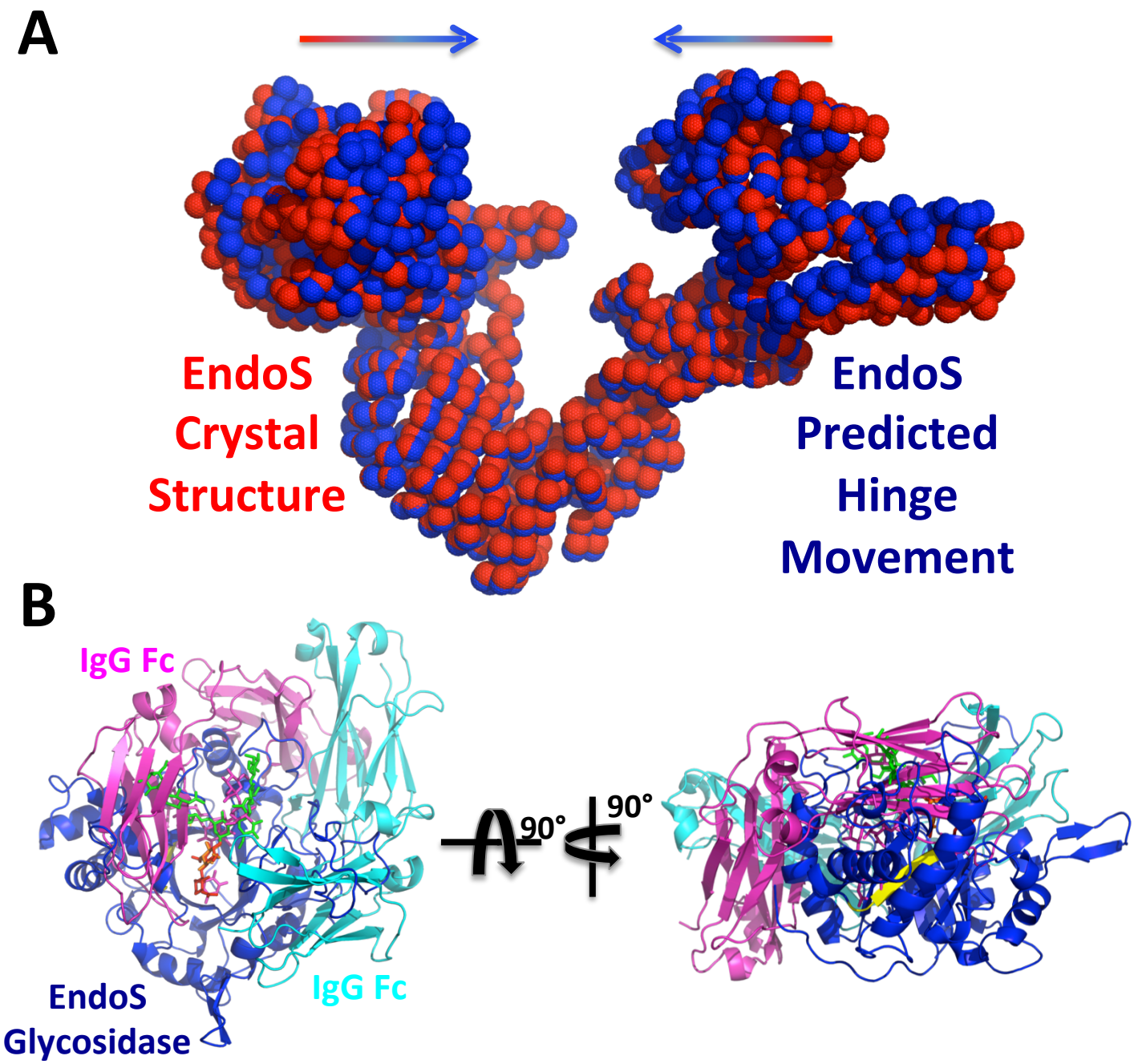


Figure S7. EndoS conformational changes required for enzymatic activity. **(A)** HingeProt analysis of the EndoSD233Q(98-995) crystal structure (red) compared to the most extreme predicted conformation (blue), representative of the predicted movement of the glycosidase domain (top left of molecule) and 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.