SUPPLEMENTARY INFORMATION TO: Structural Basis for the Recruitment of Glycogen Synthase by Glycogenin

Elton Zeqiraj[1], Jean Tang[1], Roger W. Hunter[2,3], Mar García-Rocha[4], Andrew Judd[1], Maria Deak[2,3], Alexander von Wilamowitz-Moellendorff[2], Igor Kurinov[5], Joan J. Guinovart[4,6], Mike Tyers[7,8*], Kei Sakamoto[2,3*] and Frank Sicheri[1,9,10*]

Affiliations:

[1]Samuel Lunenfeld Research Institute, Mount Sinai Hospital, 600 University Avenue, Room 1090, Toronto, Ontario, Canada, M5G 1X5

[2]MRC Protein Phosphorylation and Ubiquitylation Unit, College of Life Sciences, University of Dundee, Dundee DD1 5EH, UK

[3] Present address: Nestlé Institute of Health Sciences SA, EPFL Innovation Park, bâtiment G, 1015 Lausanne, Switzerland

[4]Institute for Research in Biomedicine (IRB Barcelona), Barcelona 08028, Spain

[5]Department of Chemistry and Chemical Biology, Cornell University, Argonne, Illinois, USA

[6]Department of Biochemistry and Molecular Biology, University of Barcelona, Barcelona 08028, Spain

[7]Institute for Research in Immunology and Cancer, University of Montreal, Montreal, Québec H3C 3J7, Canada

[8]Department of Medicine, University of Montreal, Montreal, Québec H3C 3J7, Canada

[9]Departement of Biochemistry, University of Toronto, Toronto, Canada

[10]Department of Molecular Genetics, University of Toronto, Toronto, Canada

[*]To whom correspondence should be addressed.

E-mail: sicheri@lunenfeld.ca

E-mail: Kei.Sakamoto@rd.nestle.com

E-mail: md.tyers@umontreal.ca

Supplementary Materials & Methods

General methods and buffers

Restriction enzyme digests, DNA ligations, and other recombinant DNA procedures were performed using standard protocols. All mutagenesis were performed using the QuikChange site-directed mutagenesis method (Stratagene) with the KOD polymerase (Novagen). DNA constructs used for transfection were purified from *E. coli* DH5α using Qiagen Plasmid kits according to the manufacturer's protocol. All DNA constructs were verified by DNA sequencing. Lysis buffer used for *E. coli* BL21(DE3) cells was composed of 50 mM Tris-HCl (pH 7.6), 300 mM NaCl, 20 mM imidazole, 5% (v/v) glycerol, 0.075% (v/v) 2-mercaptoethanol, 1 mM benzamidine, 1 mM PMSF and 0.3 mg/ml lysozyme. Wash buffer A was the same as lysis buffer without PMSF and lysozyme. High salt wash buffer B was the same as buffer A with 500 mM NaCl. All protein concentrations were calculated using the Beer-Lambert law by measuring the absorbance at 280 nm. Theoretical extinction coefficients were calculated using the ProtParam tool at the ExPASy proteomics server (http://www.expasy.org).

Protein expression and purification of *C. elegans* GN

GN wild type and mutants were cloned, expressed and purified using the same method used for GS with the omission of the ion exchange step. Final sizing chromatography was performed using a Superdex S75 10/300 GL column pre-equilibrated in 50 mM HEPES-NaOH (pH 7.5), 150 mM NaCl, 5% glycerol and 2 mM TCEP. Peak fractions containing >95% pure GN were pooled, aliquoted into PCR tubes, snap-frozen in liquid nitrogen and stored at −80 °C.

Protein expression and purification of mouse GST-GN

GST-MmGN1 was cloned into a pGEX-6P vector and expressed in BL21-(DE3)-RIL *E. coli* cells cultured in LB broth and induced with 0.2 mM IPTG for 16 h at 26 °C. Cells were disrupted by sonication in 50 mM Tris-HCl (pH 7.5), 0.3 M NaCl, 1 mM EDTA and 5% (w/v) glycerol and clarified at 30,000g for 30 min. GST-MmGN1 was captured with glutathione Sepharose and eluted with 50 mM Tris-HCl (pH 8), 20 mM reduced glutathione and 0.03 % Brij-35. Fractions containing GST-MmGN1 were further purified by gel-filtration as described for *C. elegans* GN and stored at -80 °C.

Mouse hepatocyte isolation and culture

Collagenase perfusion by inferior cava vein was used to isolate hepatocytes from MmGS2$^{-/-}$ randomly fed, 12 weeks old male mice. Isolated hepatocytes were suspended in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10 mM glucose, 10% fetal bovine serum (FBS), 100 nM insulin and 100 nM dexamethasone and then seeded onto 6-well plates treated with 0.01% (w/v) collagen solution (Sigma) at a final density of 4.8 x 10^5 cells/well. After 3 h media were replaced with fresh DMEM containing 25 mM glucose and 0.5% FBS and cells were treated for 2 h with adenoviruses at a similar multiplicity of infection. Media were changed and cells were overnight incubated with serum and glucose-free DMEM. Cells were then incubated in the absence or presence of 25 mM glucose for 6 h. At the end of each experiment, cell monolayers were frozen in liquid nitrogen and stored at -80 ºC for biochemical analysis.

Immunoprecipitation of GS-GN from yeast

For protein extraction and immunoprecipitation, cells were lysed by bead beating in buffer containing 50 mM HEPES-NaOH, pH 7.5, 150 mM NaCl, 5 mM EDTA, 5mM NaF, 0.1% NP-40, 10% glycerol and supplemented with Complete protease inhibitor cocktail (Roche). Cleared cell lysates were immunoprecipitated with anti-HA (Santa cruz) or anti-FLAG (Sigma) conjugated beads. Resulting proteins were analyzed on SDS-PAGE followed by immunoblot.

Immunoprecipitation of GS-GN from COS1 cells

COS1 cells were maintained in DMEM containing 2 mM L-glutamine, 100 U/ml penicillin G, 100 µg/ml streptomycin and 10% serum. For analysis of MmGN1 mutants, cells were transfected separately with pEBG6P-MmGN1 (GST-tagged) and pCMV5-MmGS2 (untagged) constructs using polyethylenimine and harvested 36 h later in 50 mM tris-HCl pH 7.5, 1 mM EDTA, 1 mM EGTA, 0.27 M sucrose, 1% (w/v) Triton X-100, 50 mM NaF, 20 mM sodium glycerol-2-phosphate, 5 mM sodium pyrophosphate, 1 mM DTT, 0.5 mM PMSF, 5 µg/ml leupeptin, 1 mM benzamidine and 1 µM microcystin-LR. Expression of GST-tagged MmGN1 was determined by quantitative immunoblot using a LiCOR Odyssey infra-red imaging instrument and ~50 µg normalized extracts were mixed with 50 µg MmGS2 bearing lysate and 5 µl glutathione Sepharose for 2 h at 4 °C.

Complexes were pelleted at 500 g for 2 min, washed 3×1 ml lysis buffer and eluted with Laemmli buffer before analysis by immunoblot for MmGS2 (Cell Signalling Technology [CST] #3886), pS8 GS2 (1), pS641 GS (CST #3891) and GST (sheep polyclonal antibody raised against full-length *S. japonicum* GST by the Division of Signal Transduction Therapy (DSTT), University of Dundee). For analysis of MmGS2 mutants, COS1 were transfected with pCMV5-MmGS2 (untagged) constructs and MmGS2 was precipitated from ~50 µg normalised extracts using 1 µg GST-MmGN1 (produced in *E. coli*) as described above.

Glycogen synthase gel-based activity assays

Bacterially expressed CeGS was assayed at 0.5 μM in 50 mM HEPES-NaOH pH 7.5, 150 mM NaCl, 1 mM TCEP, 5% glycerol and 3 mM EDTA with the addition of 3 mM G6P and/or 0.5 μM CeGBE where indicated. Bacterially expressed CeGN (wild type or mutants) were used as substrate at 5 μM in 100 μl reactions. The reaction was allowed to proceed for 80 min at 25 ºC and stopped by the addition of 20 μl of 6 times concentrated SDS sample buffer followed by boiling at 95 ºC for 5 min. Samples were then separated by SDS-PAGE electrophoreses using 10% Bis-Tris NOVEX gels (Life technologies, USA) in MOPS running buffer for 70 min at 200 V. Proteins were visualized by standard Coomassie blue staining and/or glucosylated species were visualized by staining using the periodic acid-Schiff (PAS) method (Glycoprotein staining kit, Thermo Scientific, USA).

Glycogen Synthase assay from hepatocytes

Kinetic properties of MmGS2 in Ad-infected hepatocytes was determined using the filter method of Thomas *et al.,* (2). Reactions for estimating K_m (UDP-glucose) contained 0.5 mU lysate (determined under saturating conditions), 25 mM Tris-HCl (pH 7.8), 50 mM NaF, 1 mM EDTA, 0.9 % (w/v) glycogen, 10 mM G6P and increasing concentrations of UDP-glucose (0.05 µCi [^{14}C]-UDP-glucose) for 20 min at 30 °C. To determine K_a (G6P), 0.5 mU lysate was assayed in reactions containing [UDP-glucose] = K_m, 1 mCi.mmol [^{14}C]-UDP-glucose and increasing concentrations of G6P. Kinetic parameters were determined by non-linear regression using GraphPad Prism and the standard Michaelis-Menten model or a modified equation incorporating the Hill coefficient.

Supplementary Table S1. Summary of genes referred to in this study and their accession numbers

Species	Shorthand protein name	Full protein name	Gene name	Uniprot ID
C. elegans	CeGN	glycogenin	*gyg*	H2KYQ6
C. elegans	CeGS	glycogen synthase	*gsy-1*	Q9U2D9
M. musculus	MmGN1	glycogenin-1	*Gyg1*	Q9R062
M. musculus	MmGS2	glycogen synthase 2	*Gys2*	Q8VCB3
S. cerevisiae	ScGS2	glycogen synthase 2	*GSY2*	C7GQY1

Supplementary Table S2. Summary of data collection, structure determination and refinement statistics.

Space group	$P2_1$
Unit cell parameters (Å; °)	
a	94.6
b	162.9
c	115.8
β	95.8
Molecules/asu	
CeGS	4
CeGN[34]	4
Resolution	50 – 2.60 (2.67 – 2.60)
Observed reflections	405273 (26837)
Unique reflections	106616 (7697)
Redundancy	3.8 (3.5)
$I/\sigma I$	10.8 (1.5)
Completeness (%)	99.6 (97.7)
R_{meas}	0.132 (0.145)
$CC_{1/5}$ (%)	99.6 (50.7)
R_{work}/R_{free}	0.179/0.224
RMSD from ideal geometry	
Bonds (Å)	0.002
Angles (°)	0.58
Ramachandran plot statistics	
Residues in favored region	96.6%
Residues in allowed region	3.3%
Residues in outlier region	0.1%

Supplementary figure legends

Figure S1. Sequence conservation of glycogenin (GN)

Multiple sequence alignment (black = conserved, white = not conserved) of GN from the indicated species. Alignments were performed with MUSCLE (3) and edited and displayed using ALINE (4). Relative contact area for residues involved in the CeGS—CeGN interaction are shown above the alignment as green bars. GS-interacting residues mutated in this study are boxed. The secondary structure (analyzed by DSSP) is coloured red, the variable linker region is colored orange and the C-terminal GS-interacting motif ($CeGN^{34}$) is colored green. The *C. elegans* and murine GN sequences are numbered at the top and bottom of the alignment respectively. *Ce, C. elegans; Oc, O. cuniculus, Hs, H. sapiens; Dr, D. rerio; Xt, X. tropicalis, Mm, M. musculus, Sc, S. cerevisiae.*

Figure S2. Representative electron density maps and contents of the asymmetric unit

A) Top panels: stereo figure of unbiased *F*o-*F*c electron density maps (colored green and contoured at 2.5σ) for $CeGN^{34}$ peptide bound to CeGS. Bottom panels: final 2*F*o-*F*c electron density maps (colored blue and contoured at 1σ) for $CeGN^{34}$ peptide bound to CeGS. N.B. while the peptide was 35 residues in length, we refer to this as $CeGN^{34}$ since the last residue was disordered in the crystal structure.

B) Superposition of four monomers of CeGS present in the asymmetric unit. The $CeGN^{34}$ peptide is colored green and shown in the outset.

Figure S3. Crystal packing interactions of $CeGN^{34}$ peptides found in the asymmetric unit

Crystal contacts between $CeGN^{34}$ (surface representation and colored green) and neighboring CeGS protomers (cartoon representation and colored gray). A contact between $CeGN^{34}$ and a neighboring $CeGN^{34}$ (cartoon representation and colored red) is also shown. The buried surface area between $CeGN^{34}$ and each labeled interacting partner is shown in brackets.

Figure S4. Comparison of dimer pairs within *C. elegans* and yeast GS tetramer structures

A) Closed conformation of the CeGS dimer obtained without G6P.

B) Closed conformation of a yeast ScGS dimer obtained without G6P (PDBID 3NAZ)**.**

C) Open conformation of a yeast ScGS dimer obtained with G6P (PDBID 3NBO).

Figure S5. Conservation of the CeGN[34] interaction surface and inferred ligand binding sites

A) Ribbons representation of the *C. elegans* CeGS structure with inferred ligands discerned from yeast GS structures. Ligands were oriented by superposition of GS structures with G6P (PDBID = 3NBO), UDP (PDBID = 3O3C), and α1,4-linked glucose chains (PDBID = 3RSZ and 3RT1).

B) Surface representation of *C. elegans* GS structure shown in the same orientation as **(A)** and colored in shades of blue reflecting the sequence conservation across metazoan species. N- and C-terminal phosphoregulatory regions were not included in the calculation of sequence conservation, and are shown as ribbons colored in green and red respectively.

Figure S6. N- and C-terminal phosphoregulatory regions of CeGS

A) Structure of N- and C-terminal CeGS phosphoregulatory regions. The regulatory phosphosites are shown as ball and stick models with violet-colored carbon atoms.

B) Zoom in stereo view of N-terminal phosphorregulatory region.

C) Zoom in stereo view of C-terminal phosphoregulatory region.

Figure S7. Sequence conservation of glycogen synthase

Multiple sequence alignment (black = conserved, white = not conserved) of GS from the indicated species. Alignments were performed with MUSCLE (3) and edited and displayed using ALINE (4). Residues involved in GS—GN interaction (atom pairs closer than 3.9 Å, analyzed by CONTACT from the CCP4 package) are indicated with magenta-colored bars. Key GS-interacting residues, which were mutated in this study are boxed. The secondary structure (analyzed by DSSP) is at the top of the alignment. Phospho-regulatory sites are highlighted in green and residues that form the Arg cluster

are highlighted in violet. *Ce, C. elegans; Hs, H. sapiens; Dr, D. rerio; Xl, X. laevis, Mm, M. musculus, Sc, S. cerevisiea, Dm, D. melanogaster, Gg, G. gallus.*

Figure S8. Representative electron density maps of N- and C-terminal regulatory regions

A) Top panels: stereo figure of *F*o-*F*c electron density omit maps after simulated annealing (colored green and contoured at 3.0σ) for CeGS N-terminal residues 7-42. Bottom panels: stereo figure of 2*F*o-*F*c electron density omit maps after simulated annealing (colored blue and contoured at 1.0σ) for CeGS N-terminal residues 7-42.

B) Top panels: stereo figure of *F*o-*F*c electron density omit maps after simulated annealing (colored green and contoured at 3.0σ) for CeGS C-terminal residues 639-664. Bottom panels: stereo figure of 2*F*o-*F*c electron density omit maps after simulated annealing (colored blue and contoured at 1.0σ) for CeGS C-terminal residues 639-664.

Figure S9. Crystal packing interactions of CeGS N- and C-terminal regulatory regions

A) Crystal contacts between CeGS C-terminal tail (surface representation and colored red) and neighboring CeGS protomers (cartoon representation and colored gray). N- and C-terminal regulatory regions of the reference CeGS protomer are colored green and red respectively. Residues from the Arg cluster from the crystal contacting CeGS protomer are shown in ball and stick representation and colored purple.

B) Crystal contacts between CeGS N-terminal tail (surface representation and colored green) and neighboring CeGS protomers (cartoon representation and colored gray). N- and C-terminal regulatory regions of the reference CeGS protomer are colored green and red respectively. Residues from the Arg cluster from the crystal contacting CeGS protomer are shown in ball and stick representation and colored purple.

Figure S10. Interaction of dimer pairs within the GS tetramer

Chain pairs A and B (top panel), A and C (middle panel), and A and D (bottom panel) are shown. The N-terminal phosphoregulatory region of GS is colored magenta in each protomer of the dimer. Calculated buried surface area (BSA) and BSA in the absence of the N-terminal phosphoregulatory region (N-terminus) are shown to the right.

Figure S11. CeGN linker region dictates glycogen particle size *in vitro*

A) Schematic of CeGN linker mutants generated by the insertion (ins) of a Gly-Ala linker or deletion of the variable linker region.

B) Displacement of the fluorescein-labeled $CeGN^{34}$ peptide tracer by wild type and mutant CeGN proteins and unlabeled $CeGN^{34}$ peptide.

C) Glucosylation activity of *C. elegans* GN constructs shown in **(A)**. Reactions were performed in the absence (left panel) and presence (right panel) of GBE. *De novo* glycogen particles were separated by SDS-PAGE on a 3-8% Tris-acetate gel and visualized by periodic acid—Schiff stain. Results are representative of three independent experiments.

Figure S12. Auto-glucosylated CeGN and non-glucosylated CeGN (Y194F) bind to CeGS with the same affinity

A) Top panel: Cartoon representation of auto-glucosylated wild type CeGN and the non-glucosylated CeGN Y194F mutant. Bottom panel: Displacement of fluorescein-labeled $CeGN^{34}$ tracer peptide by wild type CeGN and the Y194F CeGN mutant.

B) SDS-PAGE analysis of bacterially expressed wild type CeGN and the indicated CeGN mutants. Proteins were visualized by Coomassie stain.

Figure S13. Model of GS-GN interaction

Model for GS recruitment of GN. Distances between the C-termini (residue 265) of the globular domain of glycogenin are based on the rabbit GN structure (PDBID = 1ZCU). Distances between the N-termini of $CeGN^{34}$ (residue 268) are based on the CeGS-$CeGN^{34}$ structure reported here. GN = glycogenin, GS = glycogen synthase.

Figure S14. Glucosylation of CeGN by CeGS and CeGBE

C. elegans GS was incubated with the indicated wild type and mutant forms of GN, GBE, UDP-glucose and G6P. Reactions were incubated for 80 min and terminated by the addition of SDS sample buffer and boiling. Glucosylated species were resolved by SDS-PAGE on a 10% Bis-Tris gel for 60 min and stained by Coomassie blue **(A)** or periodic acid-Schiff stain **(B)**.

Figure S15. Kinetic properties of MmGS2 from Ad-infected hepatocytes

A & B) Kinetic properties of wild type and mutant forms of mouse GS2 in hepatocyte lysates using the filter paper method of Thomas *et al.,* (2) and purified glycogen as substrate. Results represent mean ± SD of three determinations from two pools of Ad-infected hepatocytes.

C) Ad-infected hepatocyte lysates were analysed by immunoblot with the indicated antibodies. Results are representative of two independent experiments.

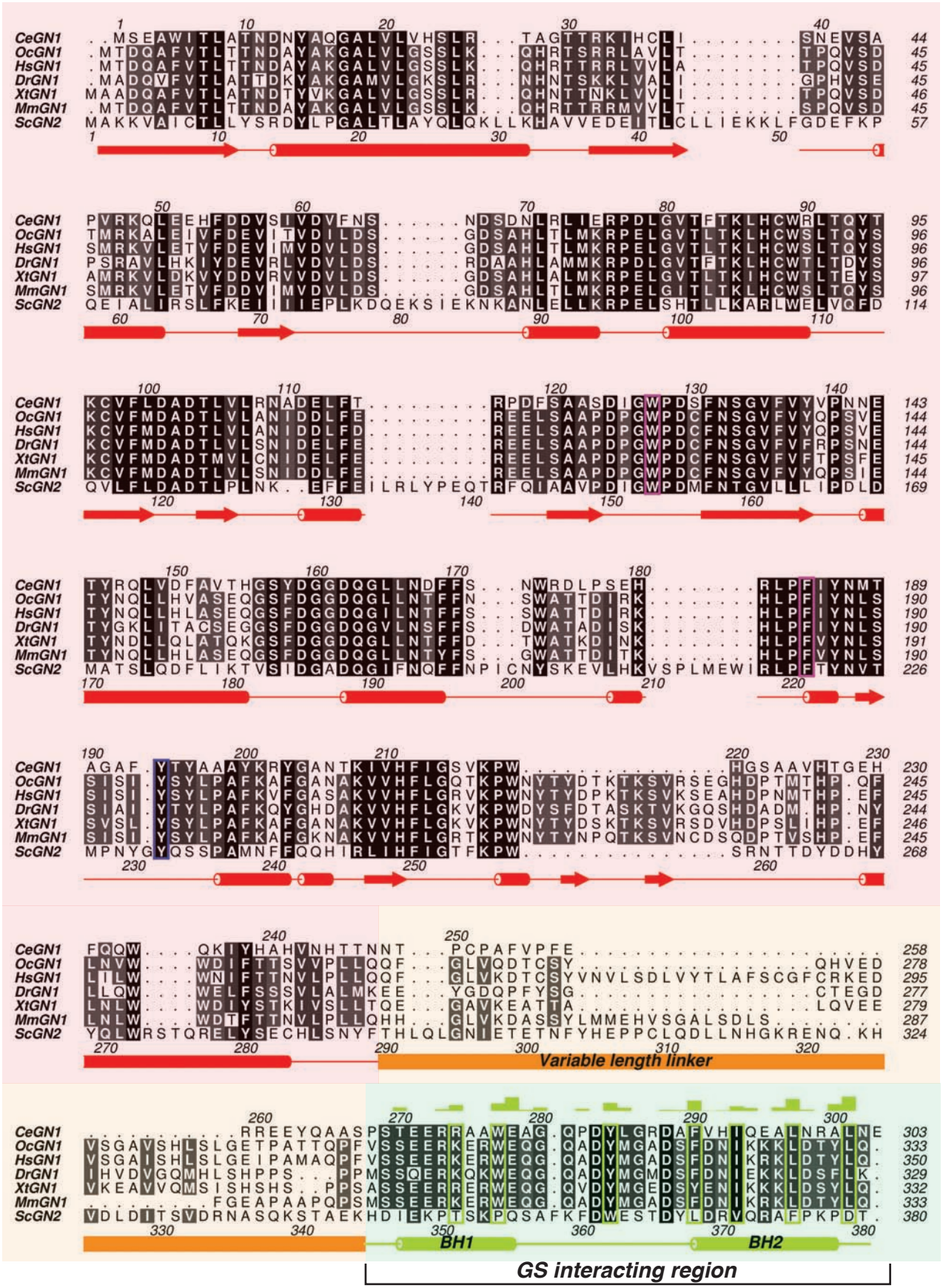
Supplementary references

1. Wilamowitz-Moellendorff von A et al. (2013) Glucose-6-phosphate-mediated activation of liver glycogen synthase plays a key role in hepatic glycogen synthesis. *Diabetes* 62:4070–4082.

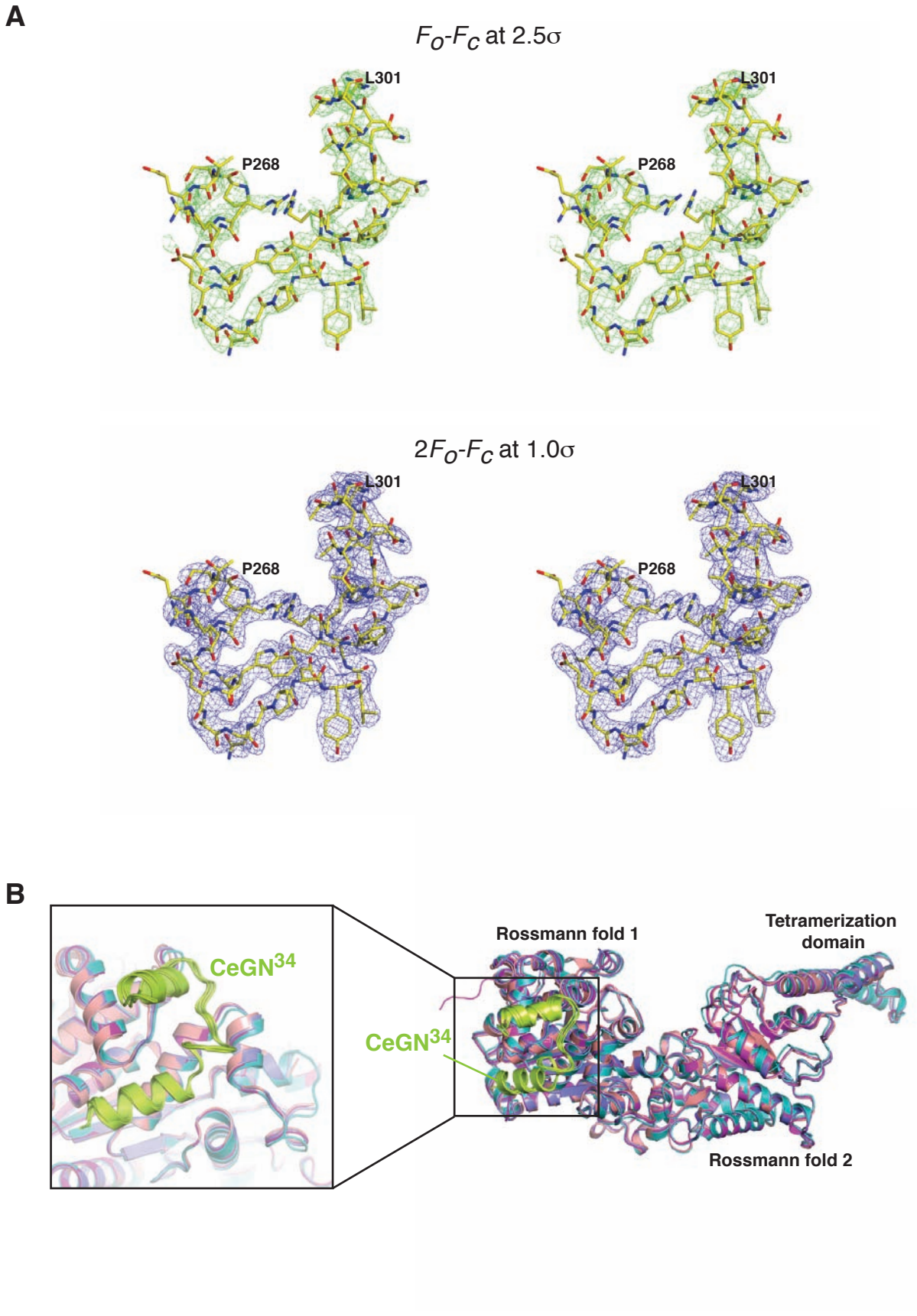
2. Thomas JA, Schlender KK, Larner J (1968) A rapid filter paper assay for UDPglucose-glycogen glucosyltransferase, including an improved biosynthesis of UDP-14C-glucose. *Analytical Biochemistry* 25:486–499.

3. Edgar R (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*.

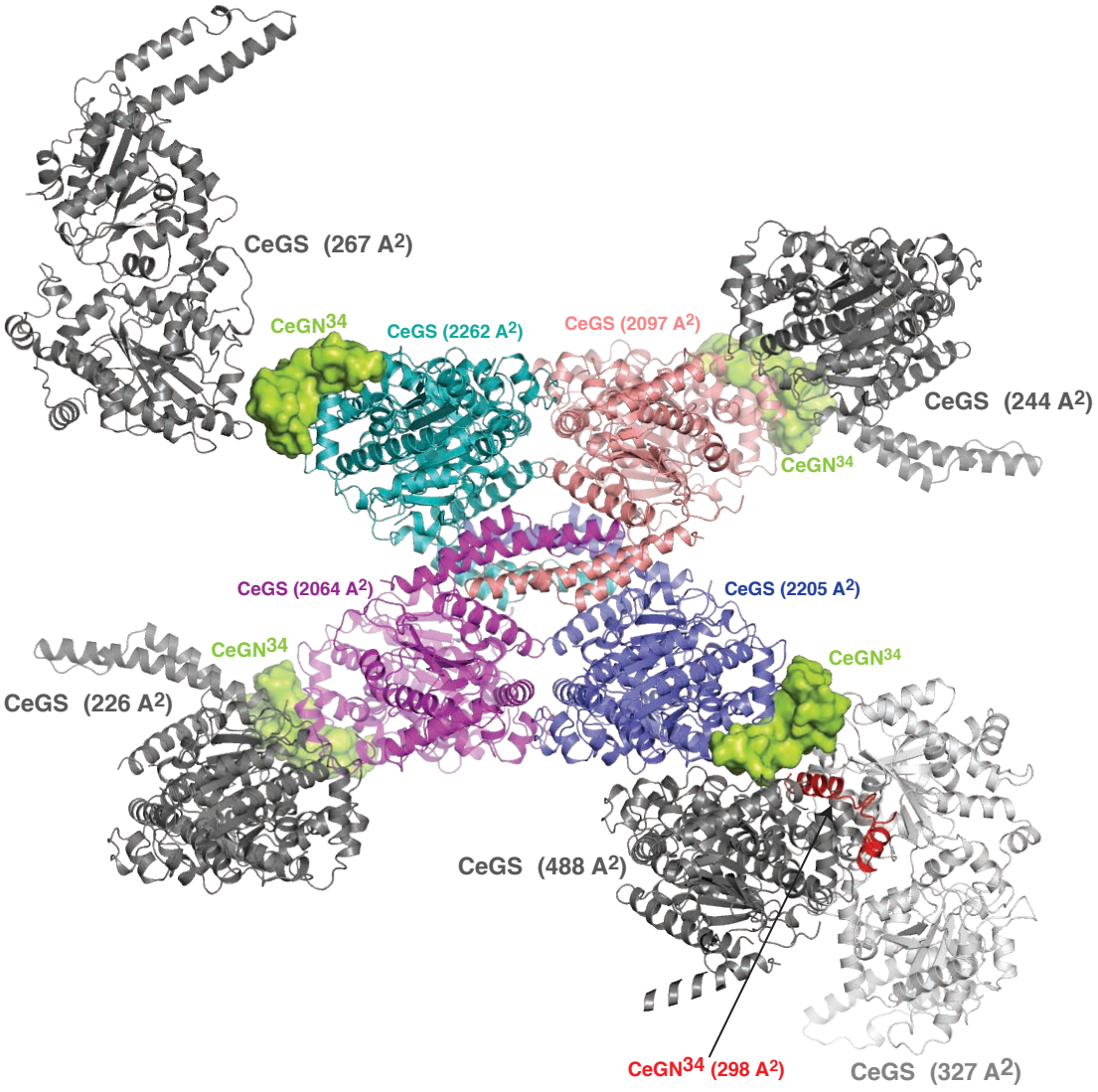
4. Bond CS, Schüttelkopf AW (2009) ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. *Acta Crystallogr D Biol Crystallogr* 65:510–512.



Supplementary Figure S1

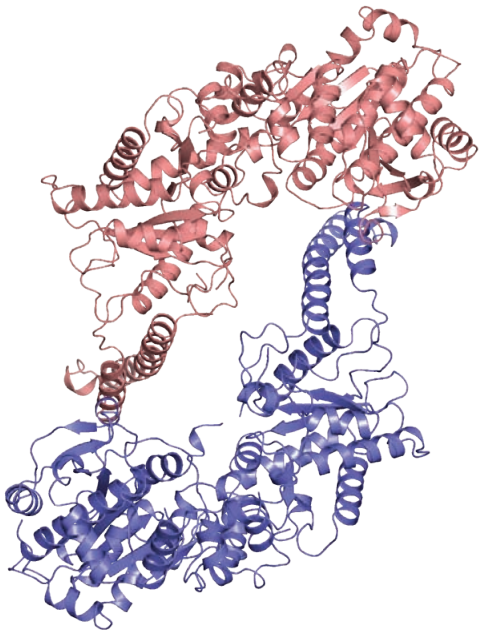


Supplementary Figure S2



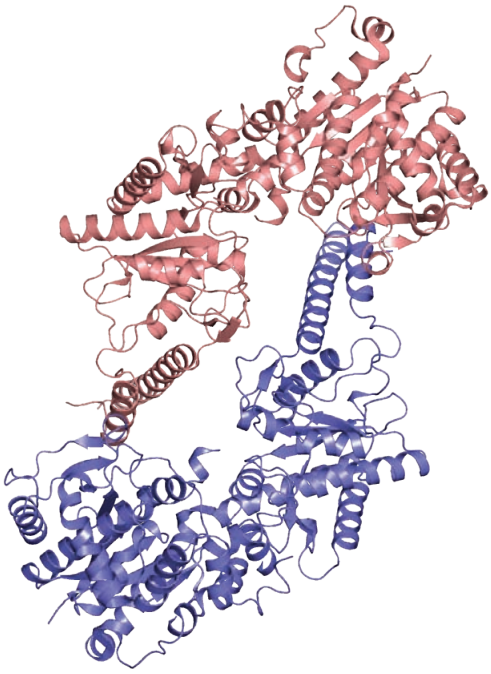
Supplementary Figure S3

A



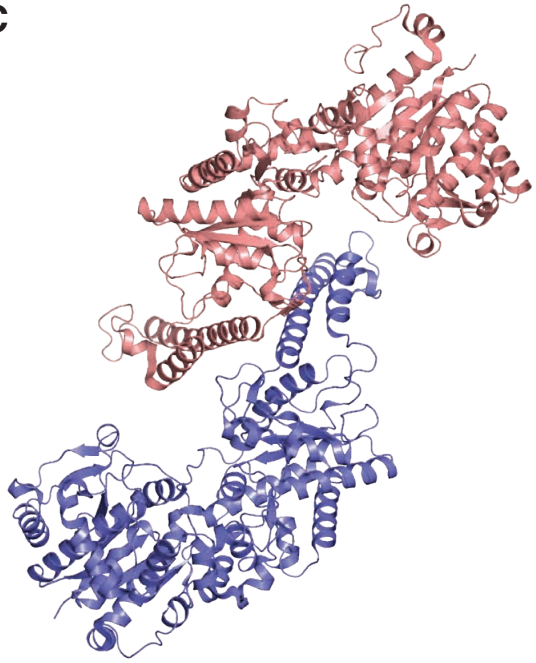
CeGS (without G6P) - Closed conformation

B



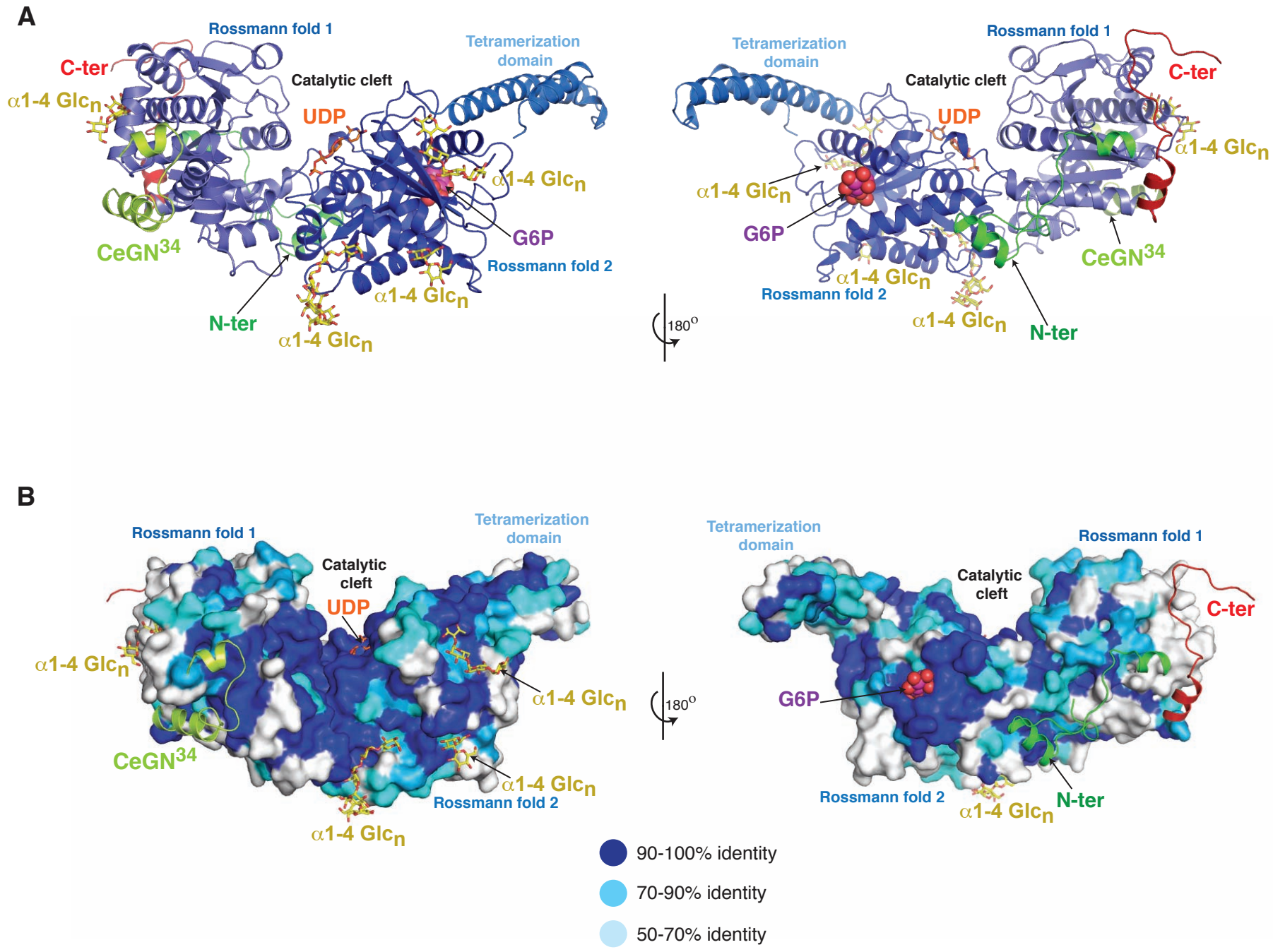
ScGS (without G6P) - Closed conformation

C

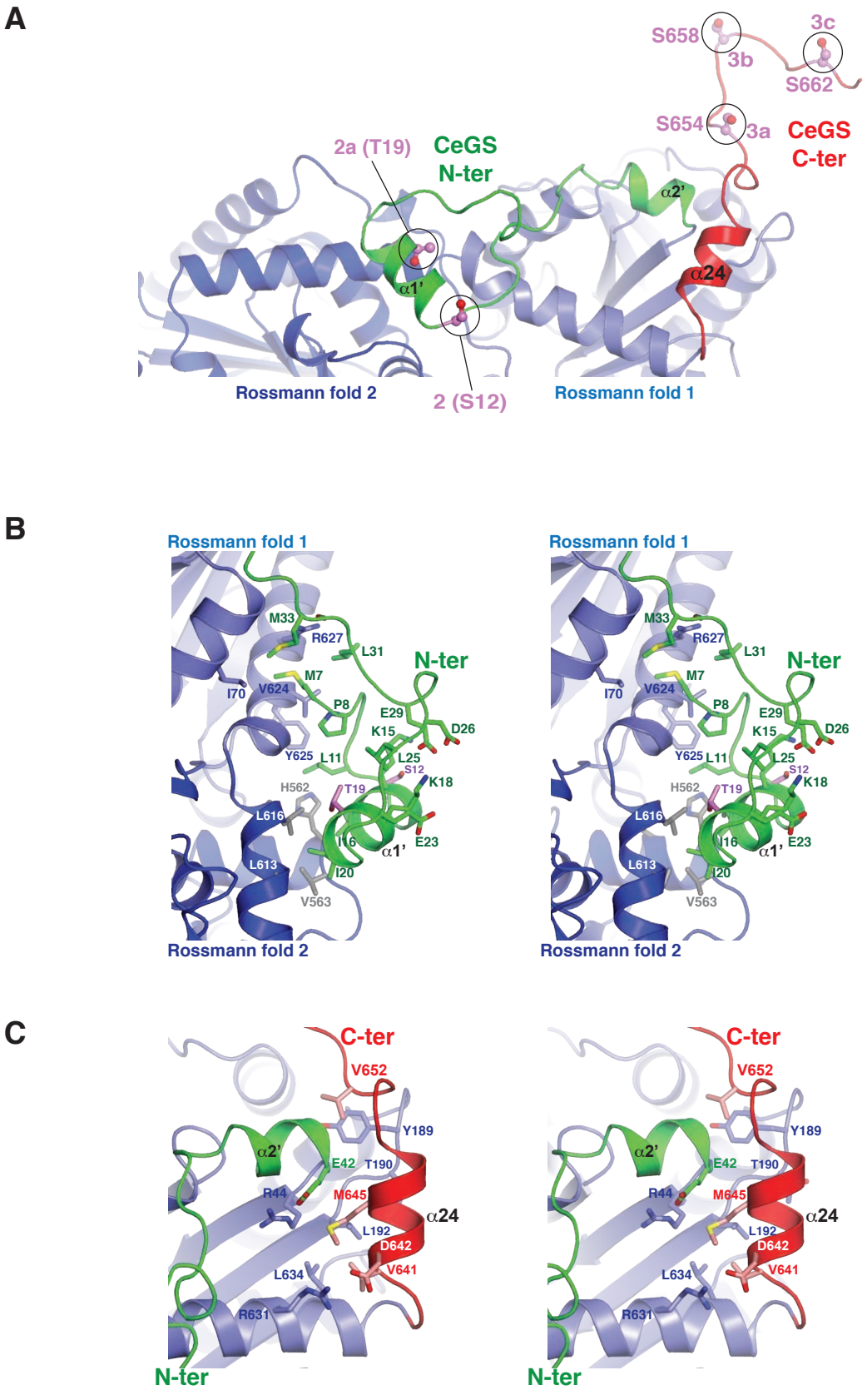


ScGS (with G6P) - Open conformation

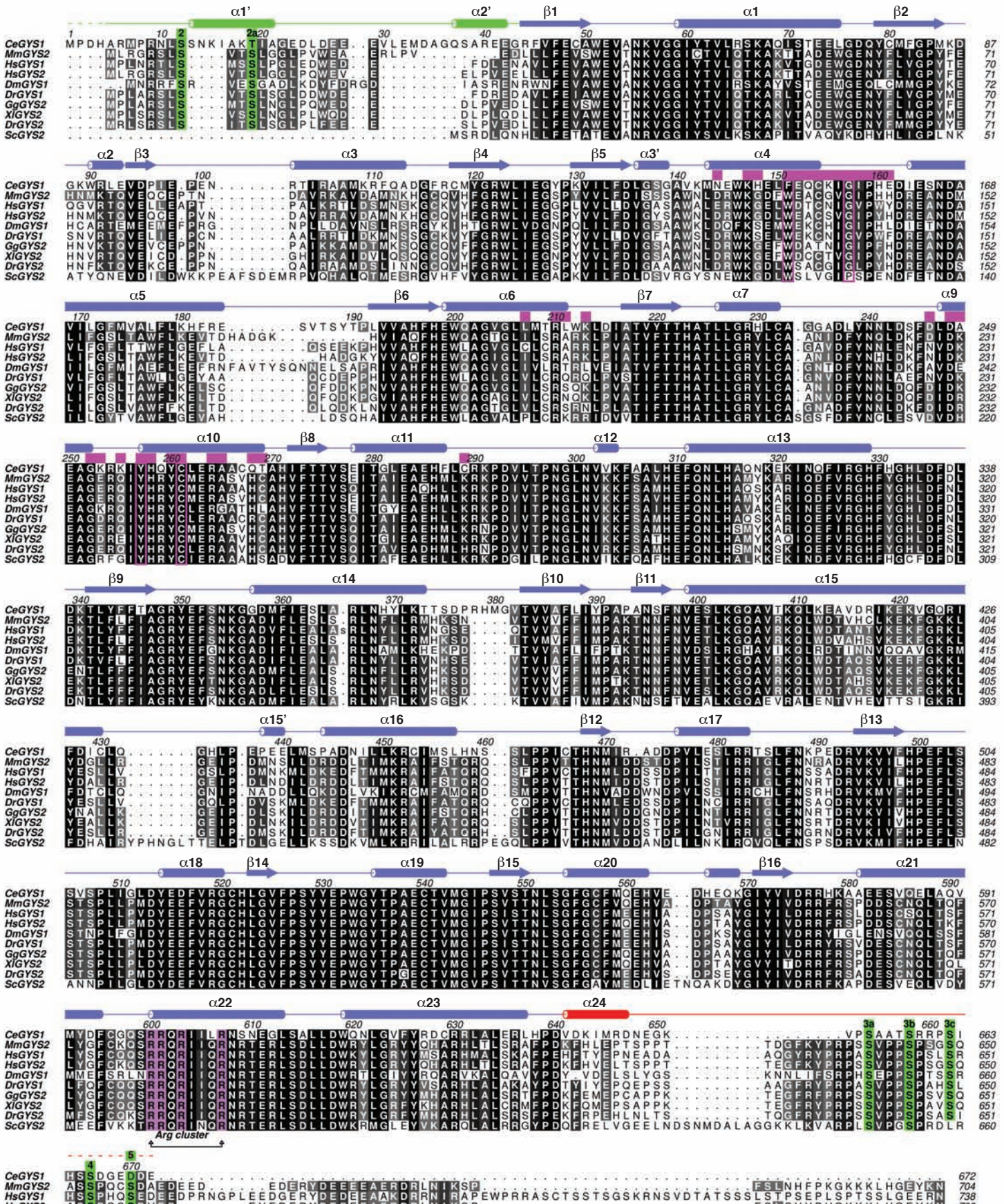
Supplementary Figure S4



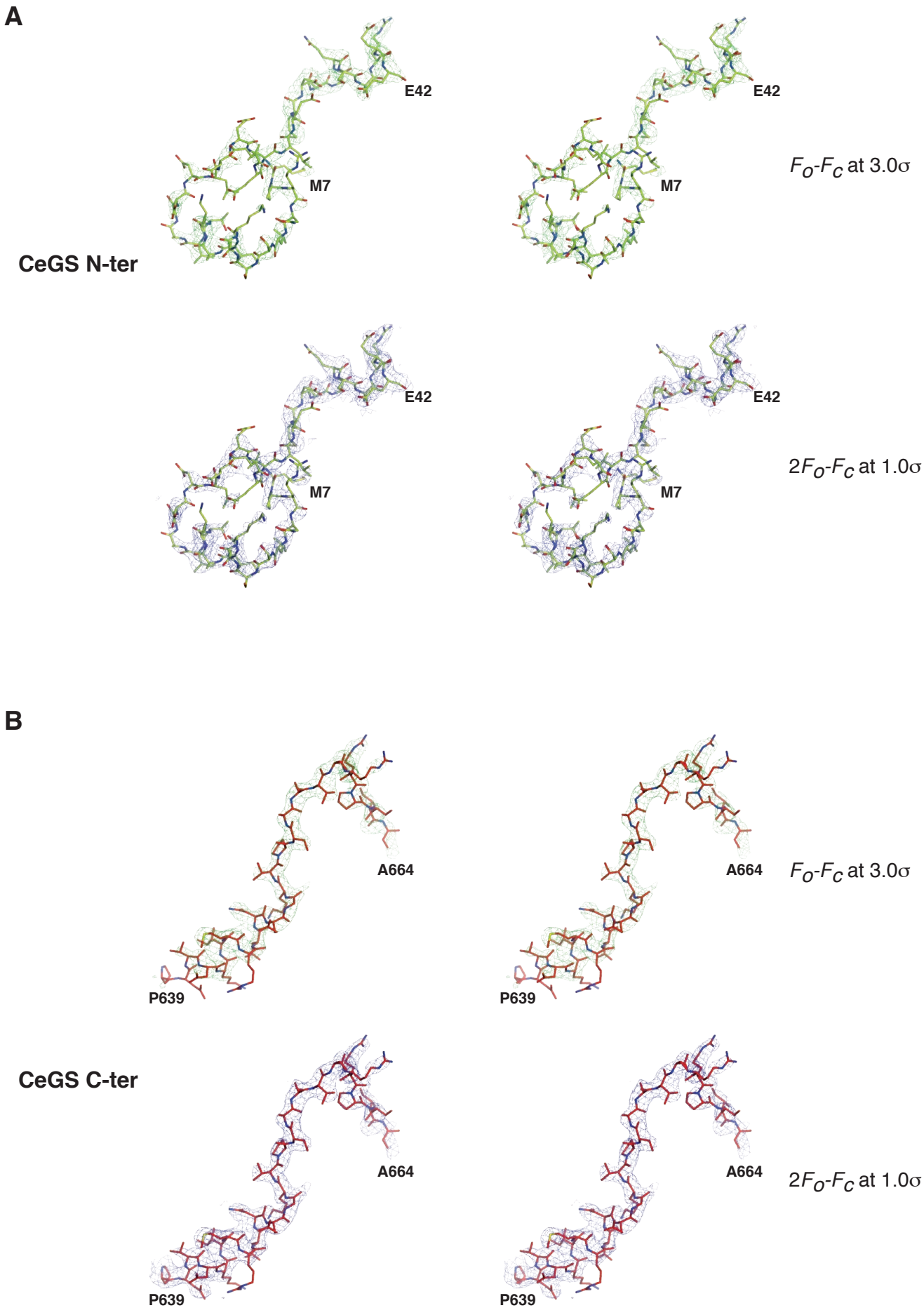
Supplementary Figure S5



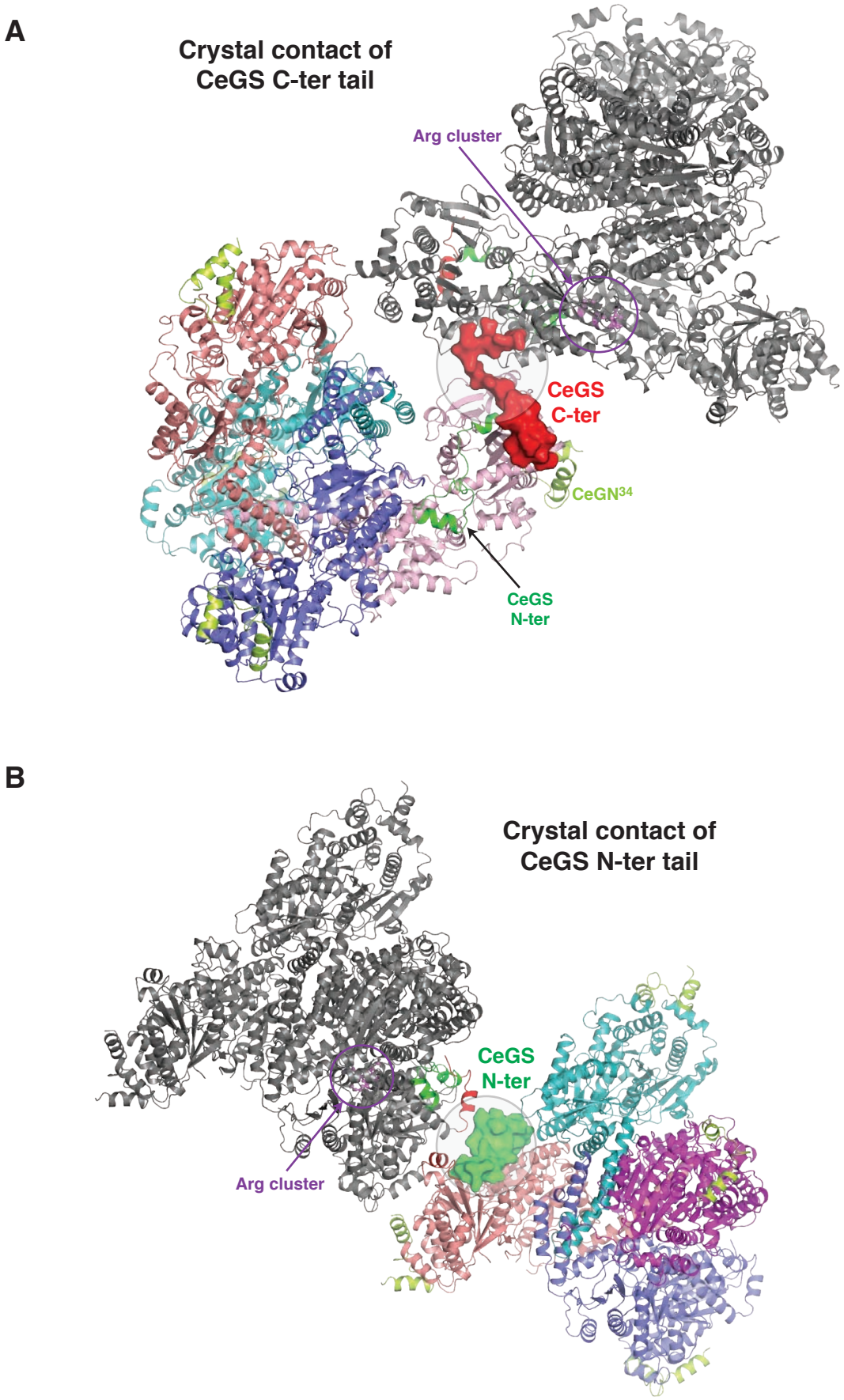
Supplementary Figure S6



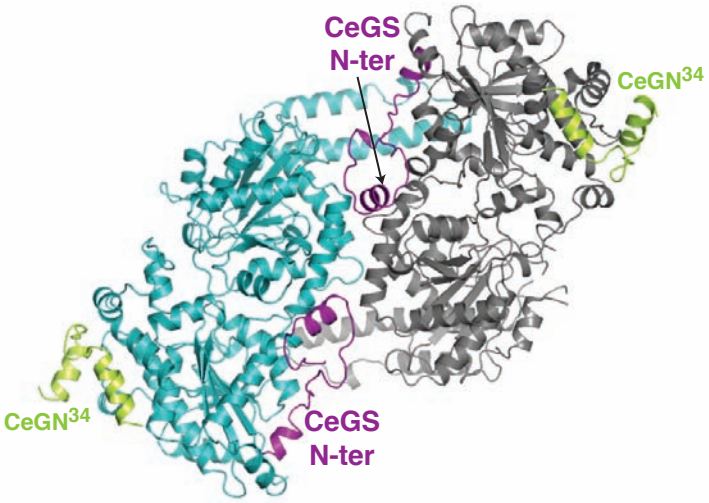
Supplementary Figure S7



Supplementary Figure S8

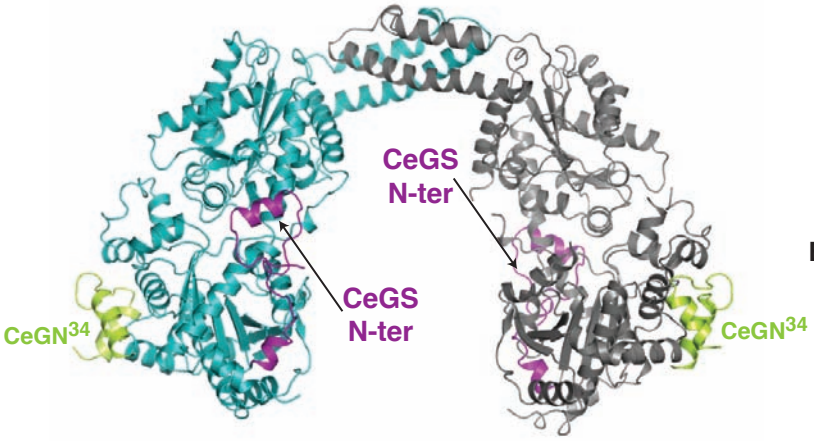


Supplementary Figure S9



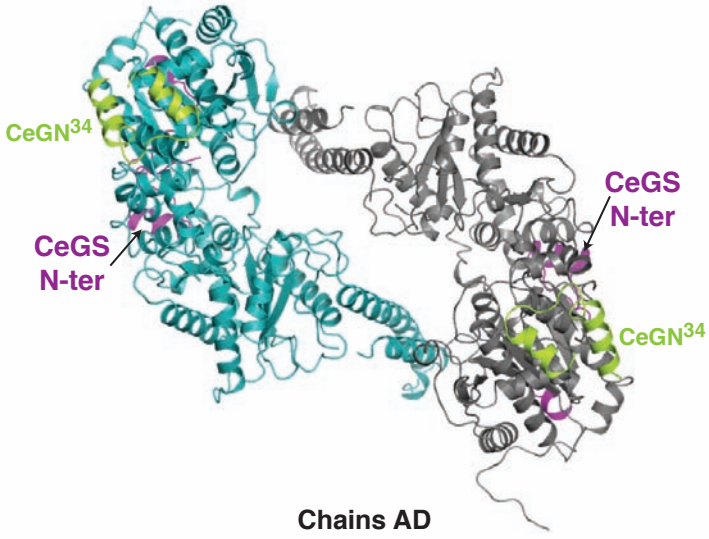
Chains AB

BSA* = 1572 Å^2
BSA (-N-terminus) = 966 Å^2



Chains AC

BSA = 2778 Å^2
BSA (-N-terminus) = 2778 Å^2



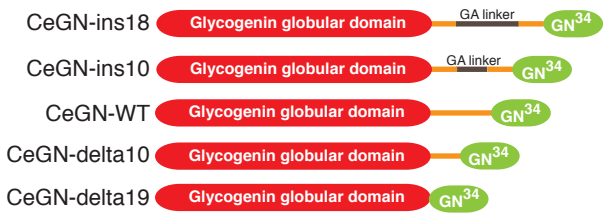
Chains AD

BSA = 1556 Å^2
BSA (-N-terminus) = 1556 Å^2

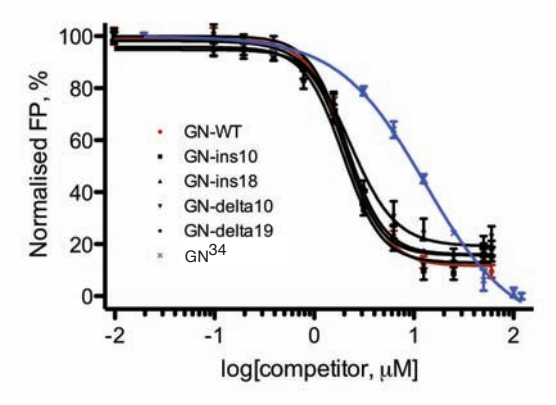
*BSA = Buried surface area

Suplementary Figure S10

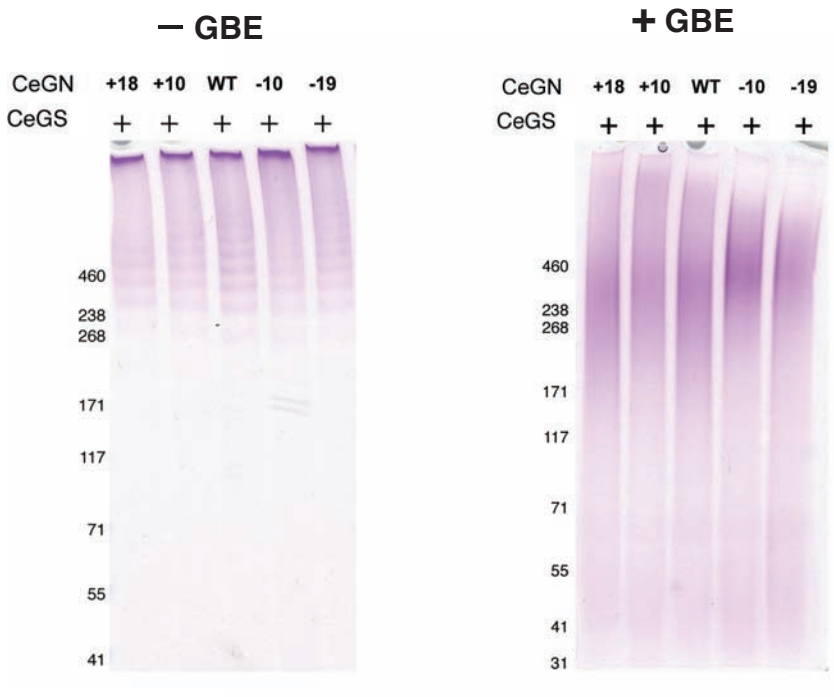
A



B

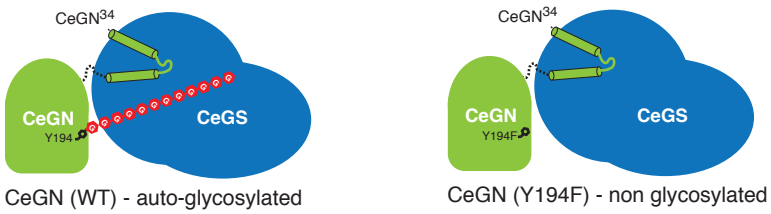


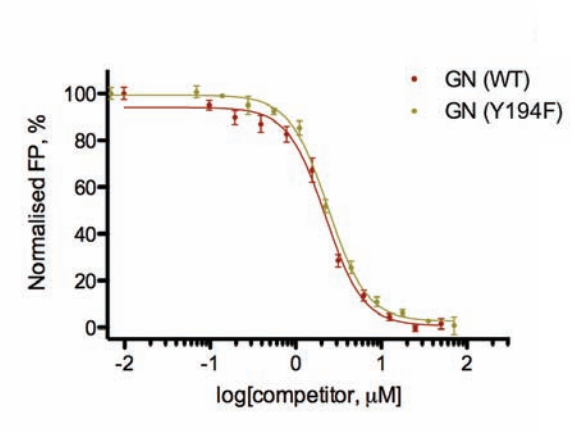
C



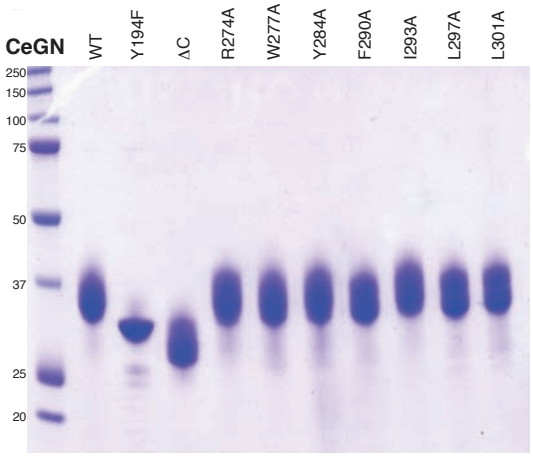
Supplementary Figure S11

A

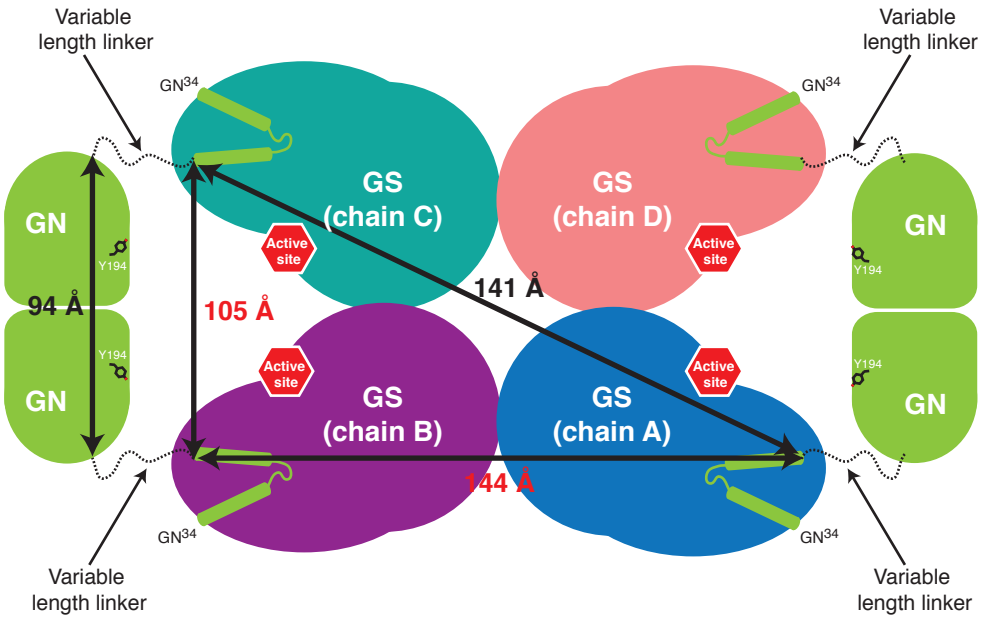




B



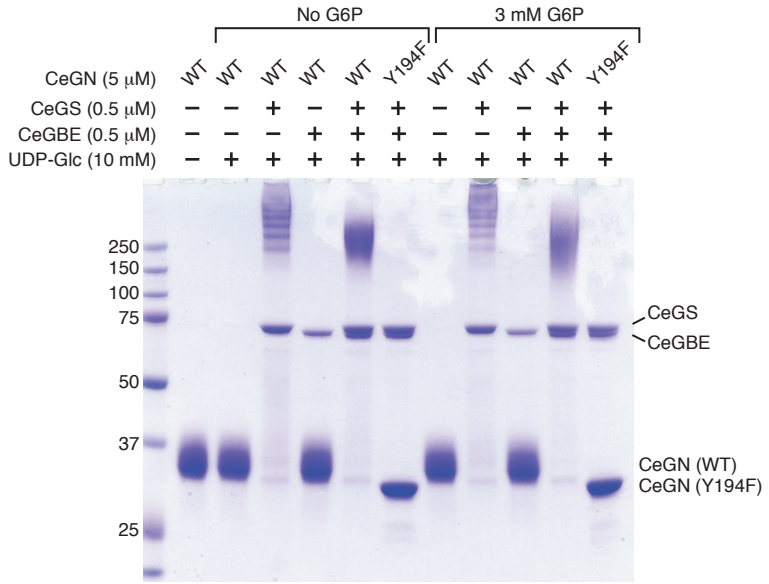
Supplementary Figure S12



Not to scale

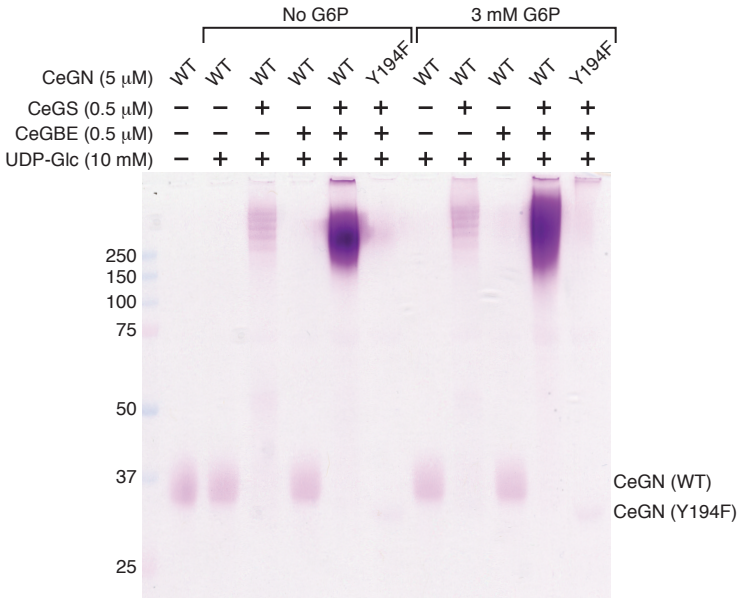
Supplementary Figure S13

A



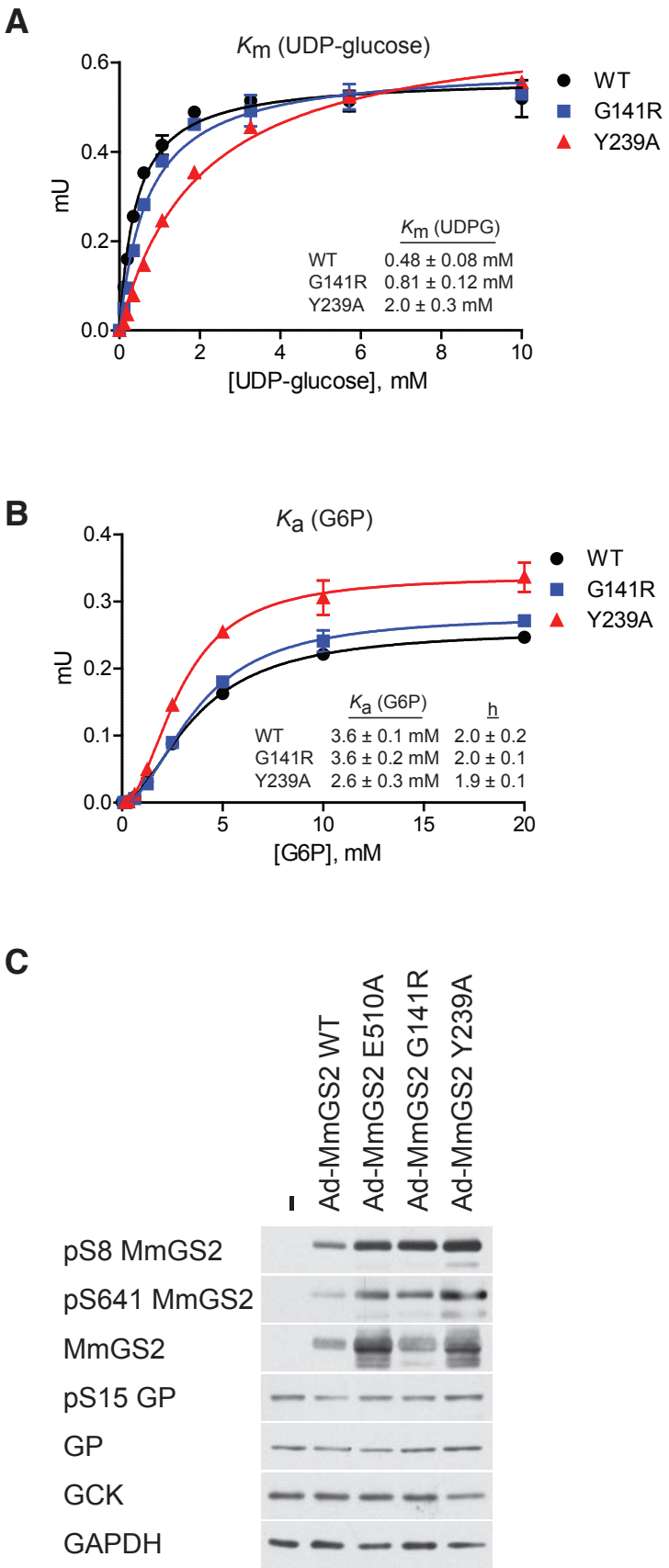
Coomassie stain

B



Periodic-Acid Schiff (PAS) stain

Supplementary Figure S14



Supplementary Figure S15