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

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

### Affiliations:

<sup>1</sup>Samuel Lunenfeld Research Institute, Mount Sinai Hospital, 600 University Avenue, Room 1090, Toronto, Ontario, Canada, M5G 1X5

<sup>2</sup>MRC Protein Phosphorylation and Ubiquitylation Unit, College of Life Sciences,

University of Dundee, Dundee DD1 5EH, UK

<sup>3</sup> Present address: Nestlé Institute of Health Sciences SA, EPFL Innovation Park,

bâtiment G, 1015 Lausanne, Switzerland

<sup>4</sup>Institute for Research in Biomedicine (IRB Barcelona), Barcelona 08028, Spain

<sup>5</sup>Department of Chemistry and Chemical Biology, Cornell University, Argonne, Illinois, USA

<sup>6</sup>Department of Biochemistry and Molecular Biology, University of Barcelona, Barcelona 08028, Spain

<sup>7</sup>Institute for Research in Immunology and Cancer, University of Montreal, Montreal, Québec H3C 3J7, Canada

<sup>8</sup>Department of Medicine, University of Montreal, Montreal, Québec H3C 3J7, Canada

<sup>9</sup>Departement of Biochemistry, University of Toronto, Toronto, Canada

<sup>10</sup>Department of Molecular Genetics, University of Toronto, Toronto, Canada

<sup>\*</sup>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* DH5 $\alpha$  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<sup>-/-</sup> 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<sup>5</sup> 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 °C 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  $3 \times 1$  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  $\mu$ M 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  $\mu$ M CeGBE where indicated. Bacterially expressed CeGN (wild type or mutants) were used as substrate at 5  $\mu$ M in 100  $\mu$ l reactions. The reaction was allowed to proceed for 80 min at 25 °C and stopped by the addition of 20  $\mu$ l of 6 times concentrated SDS sample buffer followed by boiling at 95 °C 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 [<sup>14</sup>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 [<sup>14</sup>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.

Species	Shorthand protein name	Full protein name	Gene name	Uniprot ID
C. elegans	CeGN	glycogenin	вув	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 S1. Summary of genes referred to in this study and their accession numbers

Space group	P2 <sub>1</sub>
Unit cell parameters (Å; °)	
a	94.6
b	162.9
С	115.8
β	95.8
Molecules/asu	
CeGS	4
CeGN <sup>34</sup>	4
Resolution	50 - 2.60 (2.67 - 2.60)
Observed reflections	405273 (26837)
Unique reflections	106616 (7697)
Redundancy	3.8 (3.5)
Ι/σΙ	10.8 (1.5)
Completeness (%)	99.6 (97.7)
R <sub>meas</sub>	0.132 (0.145)
$CC_{1/5}(\%)$	99.6 (50.7)
$R_{\rm work}/R_{\rm free}$	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 Table S2. Summary of data collection, structure determination and refinement statistics.

### 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<sup>34</sup>) 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 2.5 $\sigma$ ) for CeGN<sup>34</sup> peptide bound to CeGS. Bottom panels: final 2*F*o-*F*c electron density maps (colored blue and contoured at 1 $\sigma$ ) for CeGN<sup>34</sup> peptide bound to CeGS. N.B. while the peptide was 35 residues in length, we refer to this as CeGN<sup>34</sup> since the last residue was disordered in the crystal structure.

**B**) Superposition of four monomers of CeGS present in the asymmetric unit. The CeGN<sup>34</sup> peptide is colored green and shown in the outset.

# Figure S3. Crystal packing interactions of CeGN<sup>34</sup> peptides found in the asymmetric unit

Crystal contacts between CeGN<sup>34</sup> (surface representation and colored green) and neighboring CeGS protomers (cartoon representation and colored gray). A contact between CeGN<sup>34</sup> and a neighboring CeGN<sup>34</sup> (cartoon representation and colored red) is also shown. The buried surface area between CeGN<sup>34</sup> 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<sup>34</sup> 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  $\alpha$ 1,4-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  $3.0\sigma$ ) 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  $1.0\sigma$ ) 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  $3.0\sigma$ ) 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  $1.0\sigma$ ) 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<sup>34</sup> 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 nonglucosylated CeGN Y194F mutant. Bottom panel: Displacement of fluorescein-labeled CeGN<sup>34</sup> 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<sup>34</sup> (residue 268) are based on the CeGS-CeGN<sup>34</sup> 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  $\pm$  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.













ScGS (without G6P) - Closed conformation

Α

ScGS (with G6P) - Open conformation





 Rossmann fold 1
 Ross

 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0



 C-ter
 V652

 V189
 V24

 R631
 V641

 N-ter
 N-ter

## **Supplementary Figure S6**

С

В

	α1'	α2'	β1	α1	β2	
CeGYS1 MmGYS2 HsGYS1 HsGYS2 DmGYS1 GgGYS2 XIGYS2 DrGYS2 ScGYS2	1     10     2     2a20       MPDHARMPRNLSSNKIAKTAGEDLDET     MLRGRSLSS     VTSLGGLPVWER	30     40       VLEMDAGQSARE     FDLE       RELPV     FDLE       RGD     IASRE       PRGD     IASRE       DE     FDRE       DE     FDRE       DE     FDRE       DE     FDRE       DE     FDRE       DE     FDRE       DE     DLPVE       DE     SLPVE       MSRDLO     MSRDLO       α3     β4	GRFVFEQAWEVAN DLLFEVAWEVAN NAVLFEVAWEVAN NRWNFEVAWEVAN DAVLFEVAWEVAN DLLFEVAWEVAN DLLFEVAWEVAN DLLFEVAWEVTN NHLLFETATEVAN β5	60       70         KVGGIYTVLLSKAOLSTE       51         KVGGIYTVLOTKAKTTAD       51         KVGGIYTVLOTKAKTAD       51         KVGGIYTVLOTKAKTAD       51         KVGGIYTVLOTKAKTAD       51         KVGGIYTVLOTKAKTAD       51         KVGGIYTVLOTKAKTAN       51         KVGGIYTVLOTKAKTAN       51         KVGGIYTVLOTKAKTAN       51         KVGGIYTVLOTKAKTAN       51         KVGGIYTVLOTKAKTAN       51         KVGGIYTVLOTKAKTAN       51	80 ELGDQYGMFGPMKD EWGENYFLIGPYFE EWGENYFLIGPYFE EMGENYFLIGPYFE EWGENYFLIGPYKE EWGENYFLIGPYFE EWGENYFLIGPYFE EWGENYFLIGPYFE EWGENYFLIGPYFE	87 71 70 71 72 70 71 71 71 51
CeGYS1 MmGYS2 HsGYS1 HsGYS2 DmGYS1 DrGYS1 GgGYS2 XIGYS2 DrGYS2 ScGYS2	90       100         GKWRLEVDPLE, DEN	110 120 IMKRFOADGFRCMYGRW VDAMNKHGCOVHFGRW ILDSMNSKGCKVYFGRW VDAMKHGCVHFGRW IDKMNSGCKVYFGRW IDKMNSGCCVFFGRW IDKNNSGCCVFFGRW IDVLOSOGCOVFFGRW IDVLOSOGCOVFFGRW IDVLOSOGCOVFFGRW IDVLOSOGCOVFFGRW IDVLOSOGCOVFFGRW IDVLOSOGCOVFFGRW IDVLOSOGCOVFFGRW IDVLOSOGCOVFFGRW IDVLOSOGCOVFFGRW IDVLOSOGCOVFFGRW IDVLOSOGCOVFFGRW IDVLOSOGCOVFFGRW IDVLOSOGCOVFFGRW IDVLOSOGCOVFFGRW IDVLOSOGCOVFFGRW IDVLOSOGCOVFFGRW IDVLOSOFCOVFFGRW IDVLO	130 14   E G Y P Y V L F D L 14   E G G P Y V V L F D I 14   E G G P Y V V L F D I 14   E G G P Y V V L F D I 14   E G S P Y V L F D I 14   E G S P Y V L F D I 14   E G N P Y V V L F D I 14   E G N P Y V L F D I 14   E N P Y V L F D I 14   E N P Y V L F D I 14   E N P Y V L F D I 14   E N P Y V L F D I 14   E N P Y V L F D I 14   E N P Y V L F D I 14   E N P Y V L F D I 14   E N P Y V L F D I 14   E N P Y V L F D I 14   E N P Y V L F D I 14   E N P Y V L F D I 14   E N P Y V L F D I 14   E N P Y V L F D I 14	140     150       GSGAVKMNEWKHELFEOC     GSGAVKMEDRWKEELFEOC       GSSAWNLDRWKGEFWDTC     GFWDTC       GYSAWNLDRWKGELWDTC     GSAAWLDRWKGLWEAC       GSAAWLDRWKGLWEAC     GSAAWLDRWKGELWEKC       GSAAWLDRWKGEFWDAT     GSAAWLDRWKGEFWDAT       GSAAWLDRWKGEFWDAT     GSAAWLDRWKGEFWDAT       GSAAWLDRWKGEFWDAT     GSAAWLDRWKGELWSAC       DSVRGYSNEWKGDLWSAC     DSVRGYSNEWKGDLWSLV	i60         KIGIPHEDIESNDA         GYGIPHEDREANDA         NIGVPWYDREANDA         HIGIPHLDHETNDA         NIGVPWFDREANDA         NIGVPWFDREANDA         NIGIPHLDNEANDA         GIPHLDNEANDA         GIPHLDNEANDA         GIGIPHLDNEANDA         GIGLPHDREANDA         GIGLPHDREANDA         GISPENDFEINDA         GIPSPENDFEINDA         α9	168 152 151 152 154 151 152 152 152 140
CeGYS1 MmGYS2 HsGYS1 HsGYS2 DmGYS1 DrGYS1 GgGYS2 XIGYS2 DrGYS2 ScGYS2	170       180         VILGEMVALF1KHFRE       SVTSY         LIEGSLTAWFLKEVTDHADGK       SE         VIFGFLTWFLGEFLA       OSE         LIFGSLTAWFLKEVTDHADGK       SE         VIFGFLTWFLGEFLA       OSE         VIFGFLTKFLGEFLA       OSE         VIFGFLTKFLGEFLA       OSE         VIFGFLTKFLGEFLA       OSE         VIFGFLTKFLGEFLA       OSE         VIFGFLTAWFLKETTD       OFO         VIFGSLTAWFLKETD       OFO         LIFGSLTAWFLKETD       OFO         LILGSLVAWFFKELTD       OFO         XULGYTVAWFLGEVAH       B8	190     200	$\begin{array}{c} 20\\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 2 \\ 3 \\ 6 \\ 7 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1$	220 230 VYTTHATLLGRHLCA G VFTTHATLLGRYLCA AN IFTTHATLLGRYLCA AN IFTTHATLLGRYLCA GN IFTTHATLLGRYLCA AN IFTTHATLLGRYLCA AN IFTTHATLLGRYLCA GN IFTTHATLLGRYLCA GN IFTTHATLLGRYLCA GN IFTTHATLLGRYLCA GN	AD LYNN LDSFDLDA IDFYNNLDSFDLDA YDFYNNLENFNYDK IDFYNNLEKFNIDK TDFYNNLDKFNIDK YDFYNNLAKFNID IDFYNNLACFNYDK YDFYNNLDCFDIDK YDFYNNLDCFDIDK ADFYNNLDCFDIDR ADFYNNLDCFDIDR	249 231 231 242 231 232 232 232 232 232
CeGYS1 MmGYS2 HsGYS1 DmGYS1 DrGYS1 GgGYS2 XiGYS2 DrGYS2 ScGYS2	250 E A G K RK I Y H O YC L E RA A C O TA H I F T T Y S E A G E R O I Y H O YC ME RA A S Y H C A H V F T T Y S E A G E R O I Y H R Y C ME RA A A H C A H V F T T Y S E A G E R O I Y H R Y C ME RA A A H C A H V F T T Y S E A G K R O I Y H R Y C L E R G A T H L A H V F T T Y S E A G D R O I Y H R Y C L E R A A C R C A H V F T T Y S E A G D R O I Y H R Y C L E R A A C R C A H V F T T Y S E A G D R O I Y H R Y C ME RA A Y H C A H V F T T Y S E A G D R O I Y H R Y C ME RA A Y H C A H V F T T Y S E A G D R O I Y H R Y C ME RA A Y H C A H V F T T Y S E A G D R O I Y H R Y C L E R A A Y H C A H V F T T Y S E A G R F G I Y H R Y C L E R A A Y H C A H V F T T Y S E A G R F G I Y H R Y C L E R A A Y H C A H V F T T Y S E A G R F G I Y H R Y C L E R A A H S A D V F T T Y S B A G R F G I Y H R Y C L E R A A H S A D V F T T Y S B A G R F G I Y H R Y C L E R A A H S A D V F T T Y S B A G R F G I Y H R Y C L E R A A H S A D V F T T Y S B A G R F G I Y H R Y C L E R A A H S A D V F T Y S B A G R F G I Y H R Y C L E R A A H S A D V F T Y S B A G R F G I Y H R Y C L E R A A H S A D V F T Y S B A G R F G I Y H R Y C L E R A A Y H C A H Y F T Y S B A G R F G I Y H R Y C L E R A A Y H C A H Y F T Y S B A G R F G I Y H R Y C L E R A A Y H C A H Y F T Y S B A G R F G I Y H R Y C L E R A A Y H C A H Y F T Y S B A G R F G I Y H R Y C L E R A A Y H C A H Y F T Y S B A G R F G I Y H R Y C L E R A A Y H C A H Y F T Y S B A G R F G I Y H R Y C L E R A A Y H C A H Y F T Y S B A G R F G I Y H R Y C L E R A A Y H C A H Y F T Y S B A G R F G I Y H R Y C H Y R Y R Y R Y R Y R Y R Y R Y R Y R Y	280 EITGLEAEHALKRKPD SOITAIEAEHMLKRKPD SOITAIEAEHMLKRKPD SOITAIEAEHMLKRKPD SOITAIEAEHMLKRKPD SOITAIEAEHLLKRKPD SOITAIEAEHMLKRKPD SOITAIEAEHMLKRKPD SOITAIEAEHMLKRKPD		320 310 A L H E F ON L H AONKEK I NO A V H E F ON L H AOYK A R I OE A V H E F ON L H AOYK A R I OE A V H E F ON L H AOYK A R I OE A I H E F ON L H AOYK A R I OE A M H E F ON L H AOYK A R I OE A M H E F ON L H AOYK A R I OE A T H E F ON L H AOYK A R I OE A T H E F ON L H AOYK A R I OE A T H E F ON L H AOYK A R I OE A T H E F ON L H AOYK A R I OE A T H E F ON L H A K K E K I NO A H E F ON L H A L K K E K I NO	330 FIRGHFHGHLDFDL FVRGHFYGHLDFDL FVRGHFYGHLDFDL FVRGHFYGHLDFDL FVRGHFYGHLDFNL FVRGHFYGHLDFSL FVRGHFYGHLDFSL FVRGHFYGHLDFSL FVRGHFHGCFDFDL	338 320 320 321 321 321 321 329
CeGYS1 MmGYS2 HsGYS1 HsGYS2 DmGYS1 DrGYS1 GgGYS2 XIGYS2 DrGYS2 ScGYS2	340 DKTLYFFTAGRYEFSNKGADIFLESLA EKTLFFFTAGRYEFSNKGADIFLESLA DKTLYFFTAGRYEFSNKGADIFLESLS DKTLYFFTAGRYEFSNKGADIFLESLS DKTLYFFTAGRYEFGNKGADIFLEALA DKTVFFFTAGRYEFSNKGADIFLEALA EKTLFFFTAGRYEFSNKGADIFLEALS EKTLFFFTAGRYEFSNKGADIFLEALS DNTLYFFTAGRYEYKGADIFLESLS DNTLYFFTAGRYEYKGADMFTEALA	370 380 R L N H Y L K T T S D P R H M G R L N F L R M H K S N R L N Y L L R Y N G S E R L N F L R M H K S D R L N F L R Y H K S D R L N F L R Y H K S D R L N F L R Y H R S D R L N F L R Y H R S D R L N F L R Y H R S D R L N F L R Y H R S D R L N F L R Y H R S D C 16	390 VTVVAFUYPAPA VTVVFFIMPART TVVAFFIMPART TVVAFFIMPAKT VTVAFFIMPAKT VTVVFFIMPAKT VTVVFFIMPAKT VTVVFFIMPAKT KTVVAFTVMPAKN β12	400 NSF NVESLKGQAVTKQLK NF NVETLKGQAVFKQLW NNF NVETLKGQAVFKQLW NNF NVETLKGQAVFKQLW NNF NVETLKGQAVFKQLW NNF NVETLKGQAVFKQLW NNF NVETLKGQAVFKQLW NNF NVETLKGQAVFKQLW NSF TVEALKGQAFVFALE α17	420 EA V D R I K E K V G O R I D T V H C K E K F G K K L D T A N T V K E K F G K K L D T A N T V K E K F G K K L D T I N N V O Q A V G K R M D T A O S V K E K F G K K L D T A O S V K E K F G K K L D T A O S V K E K F G K K L D T A O S V K E K F G K K L D T A O S V K E K F G K K L D T A O S V K E K F G K K L D T A O S V K E K F G K K L N T V H E V T T S I G K R I β13	426 404 405 404 415 404 405 405 393
CeGYS1 MmGYS2 HsGYS1 HsGYS2 DmGYS1 DrGYS1 GgGYS2 XIGYS2 DrGYS2 ScGYS2	430       440         F D I C LQ.       GHLP.         Y DG L LR.       GE IP         GE IP       DMN KMLD KED         Y DAL LR.       GE IP         FD T C LQ.       GN IP         YDAL LR.       GE IP         FD T C LQ.       GN IP         YDAL LR.       GE IP         YESL LV.       GE IP         YESL LV.       GE IP         YESL LV.       GE IP         YESL LV.       GE IP         YESL LK.       GE IP         YESL LK.       GE IP         YESL LK.       GE IP         JUNAL K.       GE IP         JUNAL K. <td< td=""><td>450 460 1 TIMK R A IFATORO 1 TIMK R A IFAT</td><td>470           SLPPVGTHNMIR.           SIPPVGTHNMIR.           SIPPVGTHNMIDD           SIPPVGTHNMIDD           SPPVGTHNMIDD           SUPPVGTHNMIDD           SUPPVGTHNMIDD           SUPPVGTHNMIDD           SUPPVGTHNMIDD           SUPPVGTHNMIDD           SUPPTTHNWIDD           SUPPTTHNWDD           SUPPTTHNWDD           SUPPTTHNWDD           SUPPTTHNWDD           SUPPTTHNWDD           SUPPTTHNWDD           SUPPTTHNWDD</td><td>ADDPVLESLRRTSLFNKP STDPILSTIRRIGLFNSR SSDPILTTIRRIGLFNSR SSDPILSTIRRIGLFNSR WNDPVLSSIRRCHLFNSR GNDPILNCIRRIGLFNSR STDPILNTIRRIGLFNSR STDPILNTIRRIGLFNSR ANDLILNKIRCVOLFNSP</td><td>500 EDRVKVVFHPEFLS ADRVKVIEHPEFLS ADRVKVIEHPEFLS HDRVKWIFHPEFLT ODRVKVIFHPEFLS TDRVKVIVHPEFLS SDRVKMIFHPEFLS SDRVKMIFHPEFLS</td><td>504 483 484 483 494 483 484 484 484 484</td></td<>	450 460 1 TIMK R A IFATORO 1 TIMK R A IFAT	470           SLPPVGTHNMIR.           SIPPVGTHNMIR.           SIPPVGTHNMIDD           SIPPVGTHNMIDD           SPPVGTHNMIDD           SUPPVGTHNMIDD           SUPPVGTHNMIDD           SUPPVGTHNMIDD           SUPPVGTHNMIDD           SUPPVGTHNMIDD           SUPPTTHNWIDD           SUPPTTHNWDD           SUPPTTHNWDD           SUPPTTHNWDD           SUPPTTHNWDD           SUPPTTHNWDD           SUPPTTHNWDD           SUPPTTHNWDD	ADDPVLESLRRTSLFNKP STDPILSTIRRIGLFNSR SSDPILTTIRRIGLFNSR SSDPILSTIRRIGLFNSR WNDPVLSSIRRCHLFNSR GNDPILNCIRRIGLFNSR STDPILNTIRRIGLFNSR STDPILNTIRRIGLFNSR ANDLILNKIRCVOLFNSP	500 EDRVKVVFHPEFLS ADRVKVIEHPEFLS ADRVKVIEHPEFLS HDRVKWIFHPEFLT ODRVKVIFHPEFLS TDRVKVIVHPEFLS SDRVKMIFHPEFLS SDRVKMIFHPEFLS	504 483 484 483 494 483 484 484 484 484
CeGYS1 MmGYS2 HsGYS1 HsGYS2 DmGYS1 DrGYS1 GgGYS2 XIGYS2 DrGYS2 ScGYS2	510 520 530 SVSPLIGLDYEDFVRGCHLGVFPSYVEF STSPLLPMDYEEFVRGCHLGVFPSYVEF STSPLLPMDYEEFVRGCHLGVFPSYVEF STNPLFGIDYEEFVRGCHLGVFPSYVEF STSPLLPMDYEEFVRGCHLGVFPSYVEF STSPLLPLDYEEFVRGCHLGVFPSYVEF STSPLLPLDYEEFVRGCHLGVFPSYVEF ANNPILGLDYDEFVRGCHLGVFPSYVEF ANNPILGLDYDEFVRGCHLGVFPSYVEF	540 PWGYTPAECTVMGIPSV PWGYTPAECTVMGIPSV PWGYTPAECTVMGIPSV PWGYTPAECTVMGIPSV PWGYTPAECTVMGIPSV PWGYTPAECTVMGIPSV PWGYTPAECTVMGIPSV PWGYTPAECTVMGIPSV PWGYTPAECTVMGVPSI PWGYTPAECTVMGVPSI PWGYTPAECTVMGVPSI	550 550 550 570 570 560 570 560 570 560 570 560 570 560 570 560 570 560 570 570 570 570 570 570 570 57	570 HVEDHEQKGIYIVDRR HVADPTAYGIYIVDRR HIADPSAYGIYILDRR HIADPSAYGIYILDRR HIADPSAYGIYIVDRR HIADPSAYGIYIVDRR HVADPAAYGIYILDRR HVADPAAYGYIYIVDRR HVADPSEYGIYIVDRR HVADPSEYGIYIVDRR LIETNQAKDYGIYIVDRR	580 580 FRSPDSCNOLTOF FRSLDDSCNOLTOF FRSLDDSCNOLTSF FRSLDDSCNOLTSF YRSVDESCNOLTSF YRSVDESCNOLTOF FRSPDESCNOLTOF FRSPDESCNOLTOF FRSADESCNOLTOF FRSADESCNOLTOF FRSADESCNOLTOF	591 570 571 570 581 570 571 571 571
CeGYS1 MmGYS2 HsGYS1 HsGYS2 DmGYS1 DrGYS1 GgGYS2 XiGYS2 DrGYS2 ScGYS2	610 MYDFCGOSRRORI ILRNSNEGLSALLDY LYGFCKOSRRORI ORNRTERLSDLLDY LYGFCKOSRRORI ORNRTERLSDLLDY LYGFCKOSRRORI ORNRTERLSDLLDY MMEEFSRLNRRORI ORNRTERLSDLLDY LFOFCOSRRORI ORNRTERLSDLLDY LYGFCOSRRORI ORNRTERLSDLLDY MEEFVKKTRRORI ORNRTERLSDLLDY MEEFVKKTRRORI ORNRTERLSDLLDY Arg cluster	620 630 VONLGV FYRDCRRLALE VRYLGRYYMSARHALS VRYLGRYYMSARHALS VRYLGRYYOMARHLTLS VRYLGRYYVSARHLALS VRYLGRYYVSARHLALS VRYLGRYYVSARHLALS VRYLGRFYMHARHLALS VRYLGRFYMHARHLALS VRYLGRFYMHARHLALS	640 R L H P D V D K I M R D N R A F P D K H L L E P T S R A F P D H H T Y E L T S R A F P D K H V E L T S A V Y P D Y . V D L L T S K A Y P D T Y I Y E P O E K A Y P D T Y I Y E P O E K A Y P D T Y I Y E P O E R S F P E K F R M P P H L N R S F P E K F R P H L N R G Y P D O E R E L V G E	650           EGK         T D G F K Y           A D A         A OG Y R Y           P P T         T EG F K Y           YGS         K NNL IF           P SS         A A G F R F           P P K         T EG F R F           L T S         T OG F R Y           E L N D S NMD A L A GGK K LKV	38         30         660         3c           P R P S         V P S R P S I S R P S I S R         S P S P S I S R         S P S P S I S R           P R P S S V P P S P S I S R         P R P S S V P P S P S I S R         P R P S S V P P S P S I S R         S P R P S S V P S P S P S I S R           P R P A S V P P S P S P S I S R         P R P S S V P P S P S P S I S R         P R P S S V P P S P S P S I S R           P R P S S V P P S P S P S P S I S R         P R P S S V P P S P S P S S I A         S I A           A R P L S V P G S P R D L R         P R L R	663 650 651 650 650 651 651 651 651
CeGYS1 MmGYS2 HsGYS1 DmGYS1 DrGYS1 GgGYS2 XIGYS2 DrGYS2 ScGYS2	4 670 HSSDGEDDE. ASSPQCSDAEDEEDPRNGPLEEDGERYDEE ASSPQSDVED. HSSPHQSEDEEDPRNGPLEEDGERYDEE ASSPQSDVED. HSSPHSVHGSDD. HSSPHSEGEDEDERY. HSSPHSEGEDEDERY. HSSPHDSEGED.	E E A E R D R L N I K SP. E E A E R D R L N I K SP. E E A A K D R RN I R A P EWP SV DE E I E K EL G I K. D E A E K D R VN I R O P. D E A E K D R VN I R O P. D E E A E R D R O N I K S T. E E A E R D R O N I K S T. E E A E R D R O N I K S T. E E A E R D R D N I K A P. V N N A D D Y F S.	RRASCTSSTSGSK	RNSVDTATSSSLSTPSEP RNSVDTATSSSLSTPSEP FSLSHVP YNLPNRN FSLGPVS FSFGPVS FSVGADT LGVNPAA	KĠKŔKĹHĠĠŶŔŔ LSPTSSLGEERN 73 HGKKKLHGEYKN KŃĊŤŶĠĹ PĖKŇ GGKKKOHGEFRN 76 RGKKKKHGENRN 76 DGKRTOPVENGN 70 DDDDDGPYADDS 70	24839042115









В



С















Not to scale



**Coomassie stain** 



Periodic-Acid Schiff (PAS) stain

### **Supplementary Figure S14**

Α

