Supporting Information

Baskaran et al. 10.1073/pnas.1006340107

SI Text

Intein-Mediated Semisynthetic Enzyme Methods. The harvested cells (3 L culture) were resuspended in a buffer containing 20 mM Tris-HCl, 500 mM NaCl pH 8.5, 2 mM EDTA, 2 mM benzamidine, 1 mM PMSF and 0.1% Triton X-100. Cells were lysed by French press and the clarified lysate was loaded onto chitin resin (3 mL) preequilibrated with lysis buffer and washed to baseline with the lysis buffer without Triton X-100 (equilibration buffer). For the truncated core enzyme, the column was treated with 10 mL equilibration buffer containing 50 mM DTT. The semisynthetic peptide-ligated proteins were produced by treating the column-bound protein with 2 mL 2% (v/v) thiophenol in equilibration buffer followed immediately by 1 mL of 2 mM synthetic peptide and 2% thiophenol in equilibration buffer. After incubating at 25 °C for 24 hr, Gsy2p was eluted with 20 mL equilibration buffer and 10 mL equilibration buffer containing 50 mM maltose. Gsy2p showed some specific binding with the chitin bead and hence 50 mM maltose was included in the buffer to facilitate more complete elution. Both elution pools were combined and dialyzed against 20 mM Tris pH 8.0 and 1 mM 2-mercaptoethanol. Typical protein yields from 3 L bacterial cultures were 3–5 mg of peptide-ligated Gsy2p. The phosphopeptides and control nonphosphopeptides were synthesized either by the Peptide Synthesis Core facility of Indiana University or Antagene, Inc. The fusion peptides require an N-terminal Cys residue. The wild-type enzyme has a nonconserved Lys at position 642 and this residue was exchanged for Cys in the designed fusion peptides leaving the remainder of the peptide sequences unchanged. The phosphorylated peptide sequences used are as follows (nonphosphorylated controls are identical with the exception of the phosphate attached to Thr668):

642-CKLKVARPLSVPGSPRDLRSNSTVYMT(PO_3)
PGDLGTLQ-676
642-CKLKVARPLSVPGSPRDLRSNSTVYMT(PO_3)
PGDLGTLQEV NNADDYFSLGVN-690

Phasing Methods. The program package SOLVE (1) located a common set of four tantalum cluster sites in each of the two derivatives and a lower occupancy fifth site in one of the datasets. Together these derivatives provided phase information with a mean FOM of 0.47 and a Z-score of 43.8. Density modification in RESOLVE (2) improved these phases using the 4-fold averaging matrix determined from the common four tantalum sites and provided an electron density map to 5.5 Å with a FOM of 0.74. This electron density map showed strong density for many helical tubes and some of the beta-sheets. Phase extension from 5.5 Å to 3.0 Å was accomplished using the program DM as implemented in the CCP4 suite. This phase extension protocol utilized 300 steps from 6.0 Å to 3.0 Å and the same 4-fold averaging matrices utilized by RESOLVE. The resulting 3.0 Å electron density map was marginally interpretable and lacked continuity. However, the general relationship of the protein fold to the bacterial starch synthase enzymes could be identified. To improve the map quality we used the phase combination approach implemented in the program PHENIX (3, 4) (maps_only subroutine) where a partial model of the yeast Gsy2p monomer was generated by docking elements of secondary structure from a poly alanine model of Agrobacterium GS (PDB code 1RZU) into a single subunit within the 3.0 Å electron density map. The tetramer was generated by applying the NCS relations to this partial monomer structure. The phase information from this partial model and the experimental phase information from SOLVE were combined to improve the phase information to 3.0 Å. Two iterations of phase combination with increasingly larger partial models within the program PHENIX produced phases that yielded a completely interpretable 3.0 Å electron density map, which was completely retraced using the program COOT (5). A structure-based sequence alignment of the glycogen synthases assisted in assigning sequence register to the initial poly-alanine model.

1. Terwilliger TC & Berendzen J (1999) Automated MAD and MIR structure solution. *Acta Crystallogr D* 55(Pt 4):849–861.
2. Terwilliger TC (2000) Maximum-likelihood density modification. *Acta Crystallogr D* 56(Pt 8):965–972.
3. Adams PD et al. (2002) PHENIX: Building new software for automated crystallographic structure determination. *Acta Crystallogr D* 58(Pt 11):1948–1954.
4. Terwilliger TC et al. (2008) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D* 64(Pt 1):61–69.
5. Emsley P & Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D* 60(Pt 12 Pt 1):2126–2132.
6. DeLano WL (2002) *The PyMOL Molecular Graphics System* (DeLano Scientific, Palo Alto, CA).
7. Bond CS (2003) TopDraw: A sketchpad for protein structure topology cartoons. *Bioinformatics* 19(2):311–312.
8. Thompson JD, Higgins DG, & Gibson TJ (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22(22):4673–4680.
9. Nicholas KB, Nicholas HB Jr., Deerfield DW II (1997) GeneDoc: Analysis and visualization of genetic variation. *EMBNEW.NEWS* 4:14.
10. Cid E, Gomis RR, Geremia RA, Guinovart JJ, & Ferrer JC (2000) Identification of two essential glutamic acid residues in glycogen synthase. *J Biol Chem* 275(43): 33614–33621.
11. Anonymous (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D* 50(Pt 5):760–763.

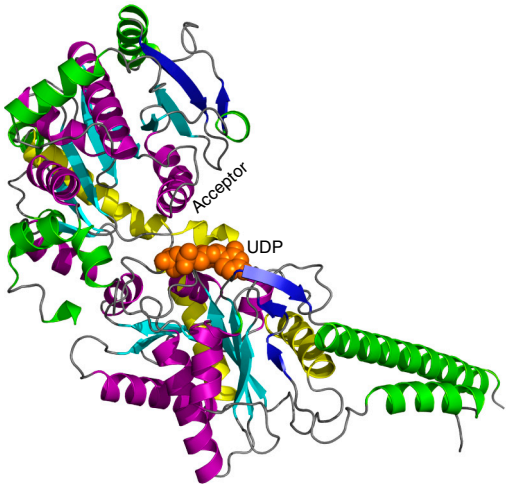

Fig. S1. Ribbon diagram of the Gsy2p monomer. The secondary structural elements of the core Rossmann domains are represented in purple and cyan, the linker sequences connecting the two domains are colored yellow and the unique eukaryotic insertions in green and blue. The UDP molecule is represented in orange space filling models and the regions that contributed to glycogen acceptor binding are labeled. [Produced using Pymol (6) for Windows.]

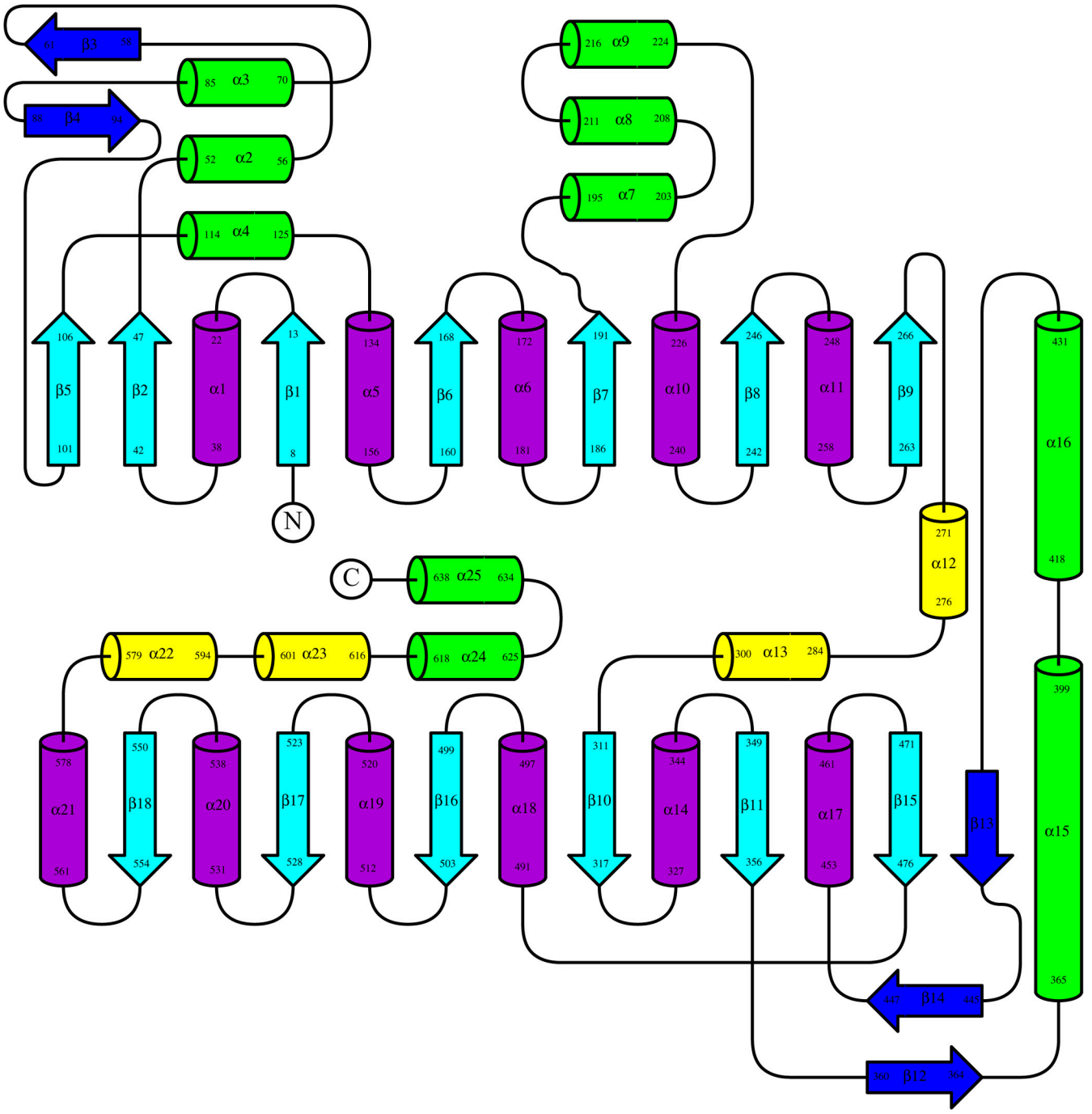

Fig. S2. Topology diagram of the Gsy2p monomer using the same color scheme as in Fig. S1. [Generated using TopDraw (7).]

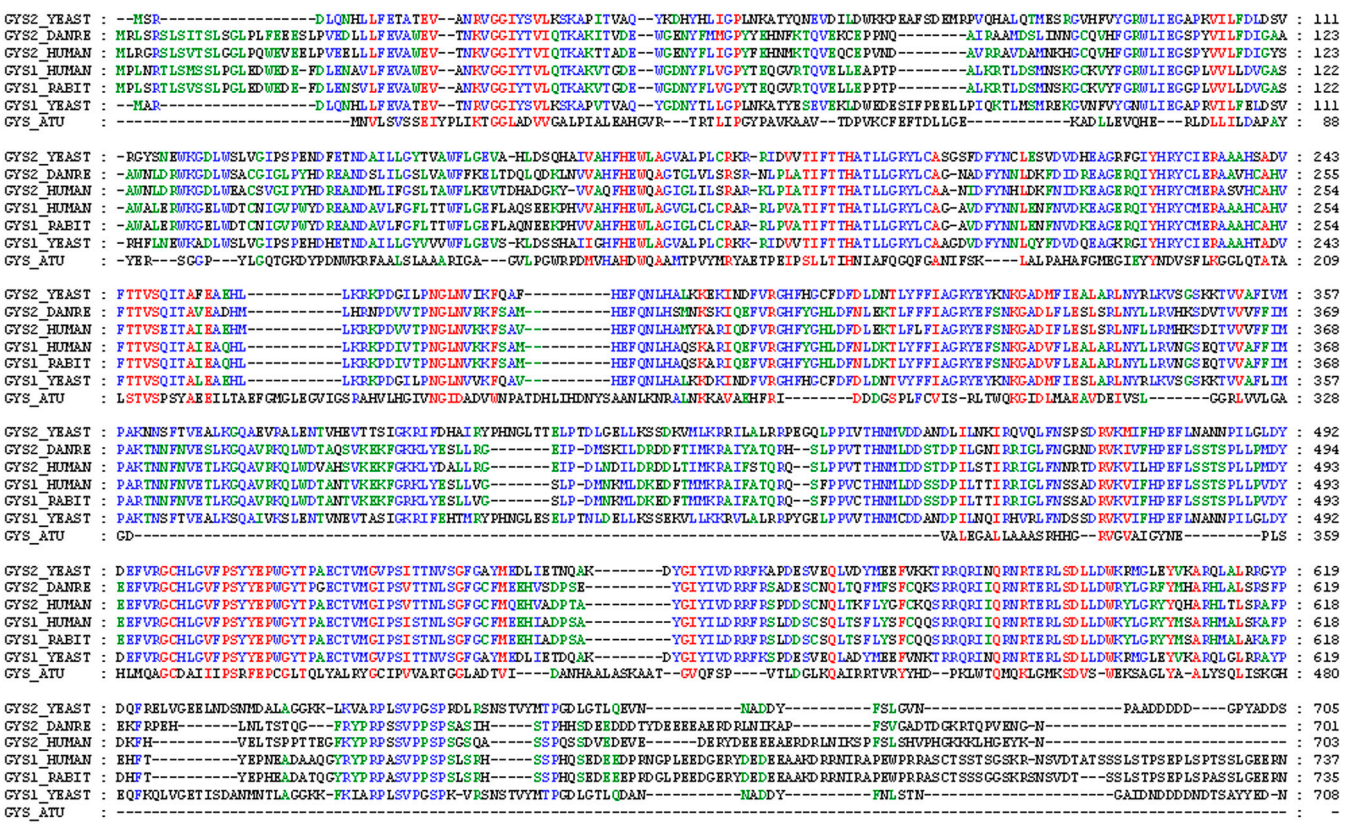

Fig. S3. Glycogen synthase sequence alignment. Structure-based sequence alignment of the Agrobacterium GS enzyme with the representative GS sequences of the eukaryotes. Residues in red, blue, and green represent 100%, 80%, and 60% sequence conservation respectively. The multiple sequence alignment of GS was generated using clustalW (8). The structure-based alignment of Agrobacterium and yeast monomers was generated using LSQ superpose in Coot (5). This alignment was used as the basis to manually edit the multiple sequence alignment produced by clustalW using the program GeneDoc (9).

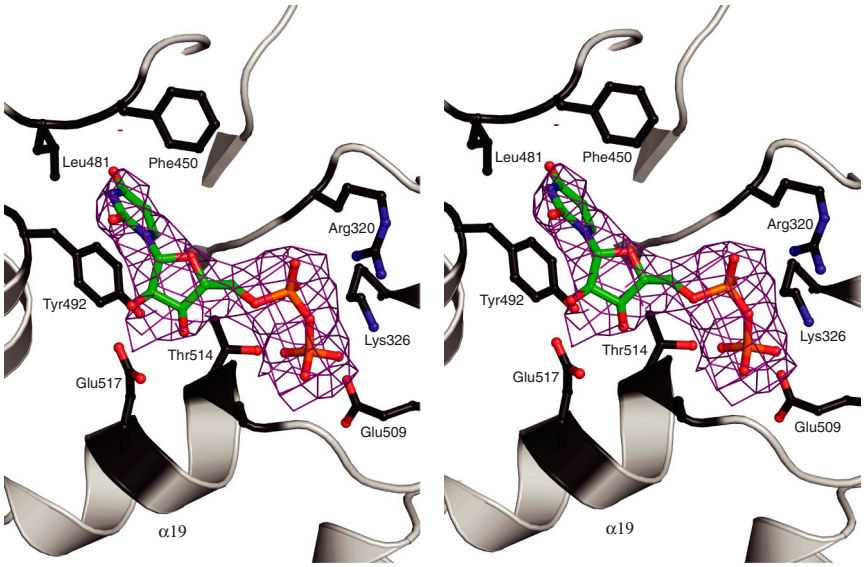

Fig. S4. Stereo diagram of the Gsy2p UDP binding pocket. Ribbon representation of the Gsy2p UDP binding pocket with the bound UDP and interacting residues in ball and stick model. The map shown is the original 2Fo-Fc electron density map for the bound UDP prior to its inclusion in the model (contoured at 1 standard deviation of the map). The residues involved contacting the UDP molecule are labeled. The uridine ring is sandwiched between the aromatic side chains of Phe480 and Tyr492, which is part of the β15-α18 loop whose interactions across the subunit interfaces change upon glucose-6-phosphate binding. The helical dipole of helix α19 and the side chains of Arg320 and Lys326 are positioned to stabilize the UDP leaving group. The proposed nucleophile (Glu509) (10) is positioned near the β-phosphate of UDP to interact with the donor glucose moiety of UDP-glucose, while the proposed general base (Glu517) (10) is instead positioned to form hydrogen bonds with both the 2′ and 3′ hydroxyls of the ribose moiety. [Produced using Pymol (6) for Windows.]

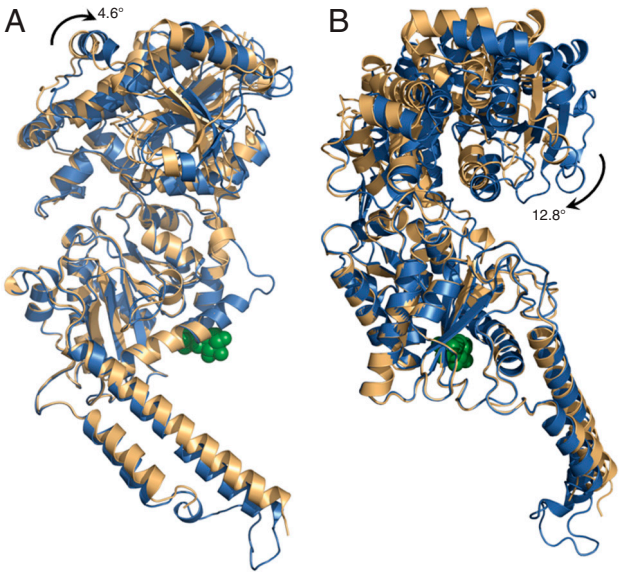

Fig. S5. Comparison of the basal and activated state conformations. (*A*) Ribbon representation of the superposed A subunits from the basal (pale yellow) and activated states (blue). The bound glucose-6-phosphate is displayed using green space filling atoms. (*B*) Ribbon representation of the superposed B subunits from the basal (pale yellow) and activated states (blue). The bound glucose-6-phosphate is displayed using green space filling atoms. The alignments for both panels *A* and *B* were generated by superposing the C-terminal Rossmann domain (residues 312–357 and 453–577) using the superpose program of the CCP4 suite (11). The arrows indicate the direction and extent of additional rotation required to superpose the respective N-terminal domains. [Produced using Pymol (6) for Windows.]