Mechanism of glycogen synthase inactivation and interaction with glycogenin

Laura Marr [a], Dipsikha Biswas [b], Leonard A. Daly [c], Christopher Browning [d], John Pollard [d], Catherine Hudson [d], Jay Bertrand [d], Neil A. Ranson [a], Heena Khatter [d], Claire E. Eyers [c], Kei Sakamoto [b], Elton Zeqiraj [a]*

[a] Astbury Centre for Structural Molecular Biology, School of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds, UK

[b] Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Copenhagen, Denmark

[c] Department of Biochemistry and Systems Biology, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, UK

[d] Vertex Pharmaceuticals Inc., Milton Park, Abingdon, UK

*Correspondence to: e.zeqiraj@leeds.ac.uk

Abstract

The macromolecule glycogen is the major glucose reserve in eukaryotes and defects of glycogen metabolism and structure lead to glycogen storage diseases and neurodegeneration. Glycogenesis begins with self-glucosylation of glycogenin (GN), which recruits glycogen synthase (GS). GS is activated by glucose-6-phosphate (G6P) and inactivated by phosphorylation, but how these opposing processes are coupled is unclear. We provide the first structure of phosphorylated human GS-GN complex revealing an autoinhibited GS tetramer flanked by two GN dimers. Phosphorylated N- and C-terminal tails from two GS protomers converge to form dynamic “spike” regions, which are buttressed against GS regulatory helices. This keeps GS in a constrained “tense” conformation that is inactive and more resistant to G6P activation. Mutagenesis that weaken the interaction between the regulatory helix and phosphorylated tails leads to a moderate increase in basal/unstimulated GS activity, supporting the idea that phosphorylation contributes to GS inactivation by constraining GS inter-subunit movement. We propose that multivalent phosphorylation supports GS autoinhibition through interactions from a dynamic “spike” region, thus allowing a “tuneable rheostat” for regulating GS activity. Our structures of human GS-GN provide new insights into the regulation of glycogen synthesis, facilitating future studies of glycogen storage diseases.

Introduction

Glycogen is a branched polymer of glucose that functions as a primary energy store in eukaryotes. In its mature form, the glycogen particle can comprise up to ~50,000 glucose units that can be rapidly utilized when glucose levels are low. Glycogen is stored predominantly in the muscle and liver cells, and to a lesser extent in other organs and tissues including kidney, brain, fat and heart[1].

Glycogen is synthesized through the cooperative action of three enzymes: glycogenin (GN), glycogen synthase (GS) and glycogen branching enzyme (GBE)[2]. GN initiates the process via auto-glucosylation of a conserved tyrosine residue, producing a primer glucose chain of 8-12 resides connected by α-1,4-linkages[3]. This glycogen initiating particle is further extended by GS after its recruitment by the GN C-terminus (GN[34]) allowing the addition of glucose residues using α-1,4-linkages[4,5]. GBE introduces α-1,6-linkages every 6-8 residues to the growing glycogen molecule, thus creating the final globular structure containing GN at the centre[2,6]. Glycogen exists as a population of molecules with varying sizes (10-290 nm) in different tissues and species, although the importance of this variability is not well understood[1,7].

Glycogen synthesis and breakdown are tightly regulated processes, and thus dysregulation of enzymes involved in glycogen metabolism contributes to glycogen storage diseases (GSDs), diabetes, neuroinflammation, neurodegeneration and muscle tissue damage[1,8]. Excessive and/or abnormal glycogen is a common characteristic in most GSDs. Pompe disease (GSDII) is caused by deficiency of acid-α-glucosidase, resulting in accumulation of lysosomal glycogen and consequent lysosomal destruction and dysfunction[9]. Lafora disease is a fatal neurodegenerative condition, of which the hallmark is Lafora bodies, which are hyperphosphorylated and poorly branched, insoluble glycogen deposits[10]. Furthermore, loss of GS-GN interaction results in muscle weakness and cardiomyopathy[11].

Studies using mouse models have found inhibition of glycogen synthesis, and in particular by reducing GS activity, to be beneficial for multiple GSDs[12,13]. To date there is no structure of the GS-GN complex and no structure of human GS. Since inhibition of GS activity is potentially beneficial for GSD patients, obtaining a human GS-GN structure and understanding how GS is regulated is instrumental in developing new therapeutics.

GN is found in two isoforms, GN1 and GN2, encoded by the *GYG1* and *GYG2* genes respectively. While *GYG1* is widely expressed, *GYG2* is restricted to the liver, pancreas and

heart[14,15]. GN belongs to the GT8 family of glycosyltransferases, containing a glycosyl transferase A (GT-A) fold with a single Rossmann fold domain at the N-terminus, which is essential for binding of the glucose donor uridine diphosphate glucose (UDP-G)[16-18]. The C-terminus comprises a highly conserved region of ~34 residues (GN^{34}) which is the minimal targeting region for binding GS[5,19]. Other interaction interfaces have been suggested[20], however, further investigation into the full-length complex is required to determine any additional interaction interfaces. The area between the N-terminal catalytic domain and C-terminal GS binding motif is a linker region that is variable in sequence and in length (**Fig. 1a and Supplementary Fig. 1**).

GS is also found as two isoforms, GS1 and GS2, encoded by the *GYS1* and *GYS2* genes respectively. These are differentially expressed, with *GYS1* being expressed predominantly in skeletal muscle and most other cell types where glycogen is present, while *GYS2* is expressed exclusively in the liver[21-23]. Eukaryotic GS belongs to the GT3 family of glycosyltransferases with a GT-B architecture comprising an N-terminal and a C-terminal Rossmann fold domain, with an interdomain cleft that contains the active site[16,24]. GS is the rate limiting enzyme in glycogen biosynthesis and as such its activity is tightly regulated[25]. GS is inactivated by covalent phosphorylation at numerous N- and C-terminal sites (**Fig. 1a**), and is allosterically activated by glucose-6-phosphate (G6P) binding and/or dephosphorylation[2,26,27]. Human GS (HsGS) phosphorylation sites lie at the N-terminus (sites 2 and 2a) and C-terminus (sites 3a, 3b, 3c, 4, 5, 1a, 1b), and phosphorylation occurs in a hierarchical fashion, whereby the phosphorylation of a specific site is the recognition motif for subsequent phosphorylation[28,29](**Fig. 1a and Supplementary Fig. 2**). How the metazoan GS is inhibited is not clear and while coupling of G6P binding and phosphorylation has been described for the yeast GS paralogues[30] no structural information of the phosphorylated version of the enzyme exists.

The complex interplay between allosteric activation and inhibitory phosphorylation is not yet fully understood, at least in part because of the lack of structural data for the full GS-GN complex. Although a binary GS-GN complex was co-purified over 30 years ago[3], we have yet to confirm the stoichiometry of this complex and identify precisely how the two proteins cooperate to make glycogen.

Here, we report the structural and functional analysis of the full-length human GS-GN complex and the cryo-EM structure of phosphorylated HsGS. The structure reveals phosphoregulatory elements form a flexible inter-subunit “spike” region emanating from two GS protomers which help to keep GS in an inactive conformation through interactions of phosphorylated S641 (site

3a) with arginine residues from GS regulatory helices (arginine cradle). Moreover, low resolution maps of GN bound to GS reveal two flexible GN dimers coordinating a GS tetramer, providing new insights into the stoichiometry and the conformational plasticity of this enzyme complex. Collectively, these results shed light on the regulation of glycogen biosynthesis and the inner workings of how GS and GN cooperate to synthesize glycogen.

Results

The GS-GN complex forms an equimolar 4:4 complex

To characterize the synthesis of glycogen by the GS-GN complex, we expressed and purified human full length GS1 and GN1 in insect cells. Consistent with previous reports, co-expression of GS with GN resulted in improved production yields over the expression of GS alone[31,32]. Purification of the wild-type (WT) complex resulted in a highly glucosylated sample, as shown by a smear on an SDS-PAGE gel corresponding to glucosylated GN detected by Coomassie and periodic acid-Schiff (PAS) staining (**Fig. 1c**). Mutation of the GN auto-glucosylating tyrosine 195[33,34] to a phenylalanine (Y195F), resulted in a non-glucosylated GN species, as shown by a single band for GN migrating at the expected size (~37.5 kDa) on an SDS-PAGE gel and absence of glucosylated species after PAS staining (**Fig. 1b and 1c**).

To determine the stoichiometry of the GS-GN complex, we first performed mass photometry analysis of GS-GN and GS-GN(Y195F) mutant complexes, which enables mass measurements of single molecules in solution. Mass photometry measurements of the GS-GN(Y195F) complex showed a predominant molecular weight of 473 +/- 43 kDa, which is suggestive of a 4:4 stoichiometry (calculated mass of 485 kDa) (**Fig. 1d)**. Analysis of the GS-GN(WT) sample identified a species with an average molecular weight of 534 +/- 110 kDa and the measured peak was broader than the non-glucosylated species (**Fig. 1d**). The increase in molecular weight and distribution compared to the Y195F complex can be attributed to the glucosylation associated with the WT complex.

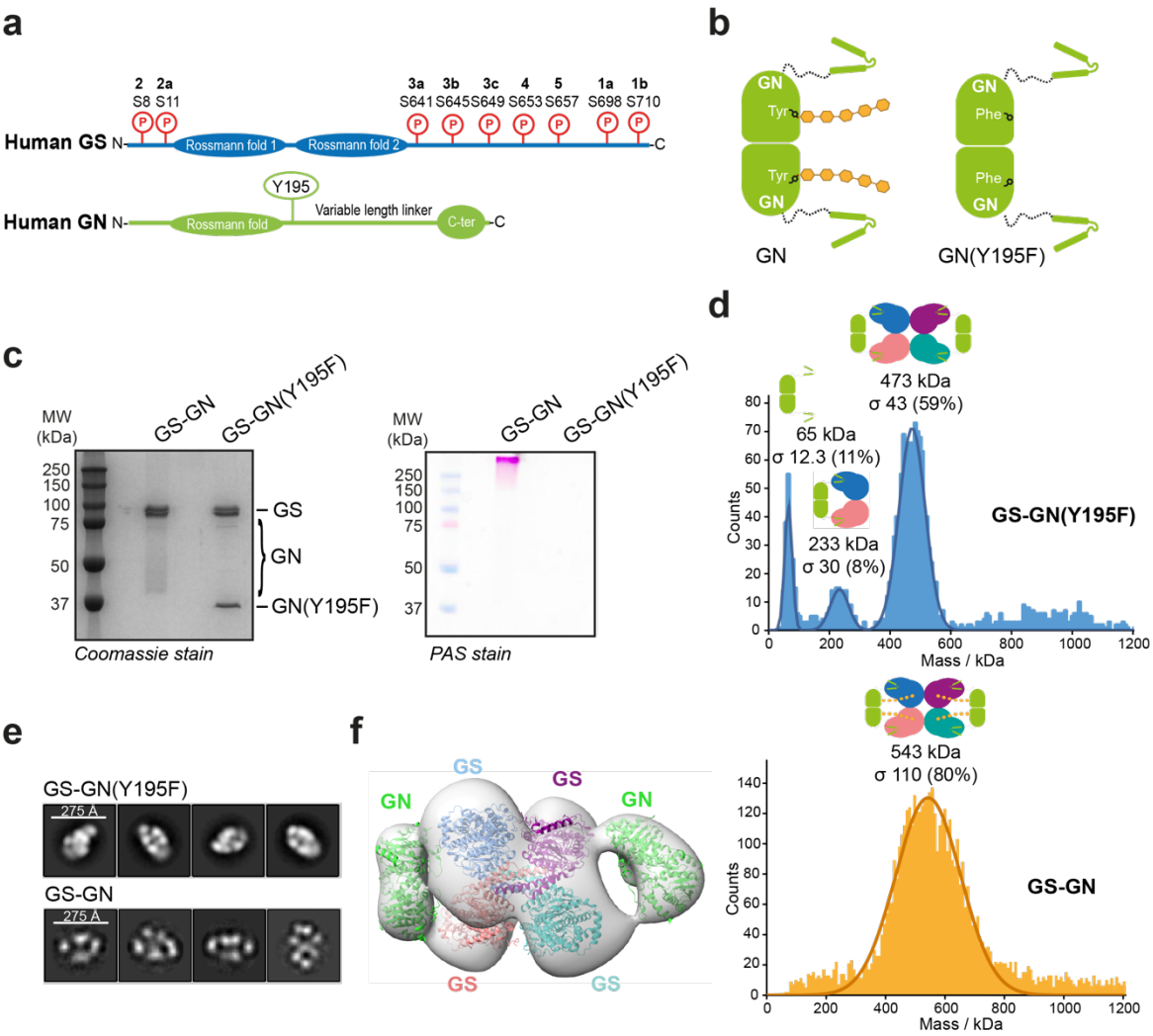


Fig. 1 Structural analysis of the full-length GS-GN complex
a Domain architecture of human GS (top) and human GN (bottom). Known *in vivo* phosphorylation sites of GS are shown in red and are labelled with residue number and classical nomenclature (in bold). GN tyrosine 195 that becomes auto-glucosylated and mutated to a phenylalanine (Y195F) in this study is indicated. Not to scale. **b** Cartoon representation of GN WT and Y195F. **c** SDS-PAGE analysis of GS-GN WT and Y195F complexes (left) and periodic acid-Schiff (PAS) staining of both complexes (right). **d** Mass photometry of GS-GN(Y195F) (top) and WT complex (bottom). Expected stoichiometry for each peak is indicated. **e** Selected 2D class averages after negative-stain electron microscopy (nsEM) analysis of GS-GN(Y195F) (top) and WT (bottom), revealing a GS tetramer and two GN dimers. **f** nsEM final map (C1 symmetry, ~22 Å) is shown in transparent surface, with fitted human GN crystal structure (PDB ID 3T7O) and human GS cryo-EM structure (reported here).

To understand how GS and GN interact within the GS-GN complex, and to reveal the overall shape of the complex we performed negative stain electron microscopy (nsEM) of both the WT and Y195F complexes. 2D class averages show two GN dimers, one on either side of a GS tetramer, for both complexes (**Fig. 1e**). Final 3D maps for both complexes are consistent with the 2D classes, and the density can accommodate a GS tetramer and two GN dimers (**Fig. 1f**). This nsEM confirms a 4:4 stoichiometry and is consistent with previous findings showing that glucose chains are not essential for the GS-GN interaction and that GS can

156 interact with four GN C-terminal peptides simultaneously[5,19]. Surprisingly, GN dimers do not
157 engage GS dimers in an identical fashion, with one GN tilted slightly towards one of the GS
158 protomers (**Fig. 1f**). Collectively, these results confirm that the GS-GN complex has a 4:4
159 stoichiometry where two GN dimers can engage a single GS tetramer. This provides the first
160 glimpse of the GS-GN complex structure and explains how GS can extend glucose chains
161 from two GN dimers simultaneously.
162
163 **The structure of human GS reveals a dynamic "spike" region**
164
165 Previous attempts to crystallise full length GS in complex with full length GN were
166 unsuccessful[19] and led us to pursue structural analysis using cryo-electron microscopy (cryo-
167 EM). NsEM indicated that the position of each GN dimer is different suggesting flexibility of
168 GN in the complex. Cryo-EM analysis of the GS-GN(Y195F) complex confirmed this GN
169 flexibility, evident from the lack of GN signal in 2D class averages (**Supplementary Fig. 4a**)
170 and subsequent 3D maps. Although we could detect GN presence after data processing
171 without the application of symmetry averaging (**Supplementary Fig. 5a and 5b**), it was not
172 possible to trace the connecting residues between the GN globular domain and the C-terminal
173 GN[34] region that binds GS. To gain a higher resolution structure for the human GS, we applied
174 D2 symmetry and achieved a global resolution of 4 Å (**Fig. 2a, Supplementary Fig. 4,**
175 **Supplementary Table 1**). The reconstruction revealed a tetrameric arrangement of HsGS in
176 agreement with the crystal structures of C. elegans (CeGS) and yeast (yGS) (**Supplementary**
177 **Fig. 6a**). Density for the C-terminal GS interacting region of GN is sufficient for a model of the
178 human GN C-terminus to be placed and this is consistent with the CeGS-GN[34] crystal structure
179 (**Fig. 2a and Supplementary Fig. 6b**). Structural analysis of the HsGS-GN(WT) complex
180 revealed a 6 Å map of GS and comparing this to the GS structure from HsGS-GN(Y195F)
181 complex reveals no differences at this resolution (**Supplementary Fig. 5c and 5d**).
182
183 Notably, the EM density maps show density for an inter-subunit apparatus containing the GS
184 phosphoregulatory regions. This occupies the space where the N- and C- terminal regions of
185 two adjacent GS protomers meet and traverse away from the core (**Fig. 2b**). Local resolution
186 analysis shows that this area is resolved at lower resolution than the core of the GS tetramer
187 (**Supplementary Fig. 4b**). Analysis of this ~25 kDa region by focussed 3D classification
188 (without applying symmetry) reveals that the region is highly flexible, as seen by the various
189 different conformations, explaining the lower resolution and uninterpretable density (**Fig. 2c**
190 **and Supplementary Fig. 4f**). Interestingly, these "spike" regions were present in all the
191 refined classes, and suggests that GS exists as a continuum of structures with a core inactive

192 tetramer and “dynamic spikes” buttressed on either side, thus constraining the opening of GS
193 to an active tetramer.
194
195 To explore the flexibility and mobility of GS, we performed 3D variability analysis[35] using
196 cryoSPARC[36]. The dynamic movements of the “spike” region and concurrent movements of
197 the GS tetramer are highlighted in Movie S1. While some flexibility was observed within each
198 GS protomer, an opening of the GS tetramer towards an active, relaxed state was not
199 observed, highlighting the crucial role of the spike regions in maintaining an inactive, tense
200 state GS tetramer.

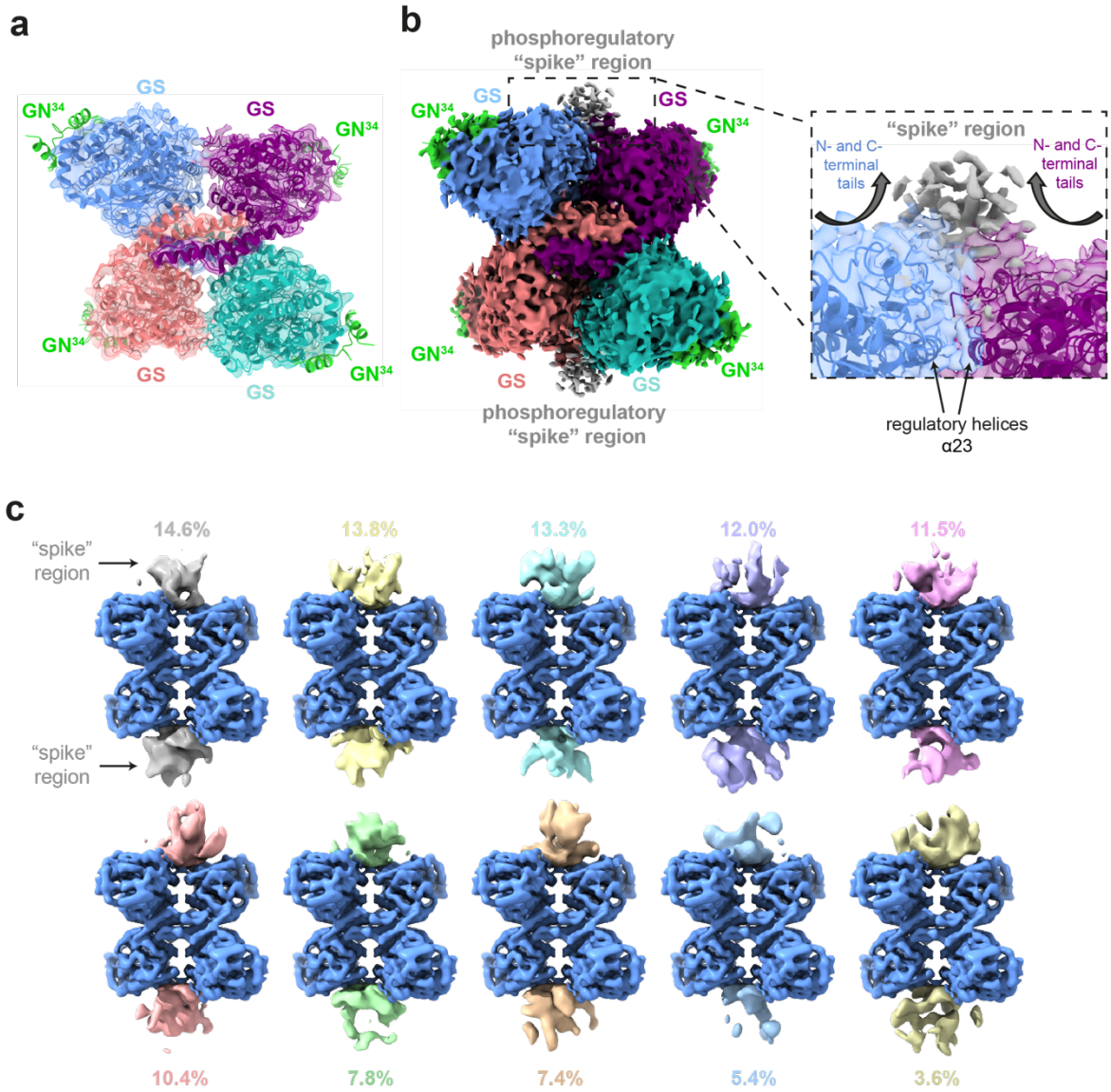


Fig. 2 Cryo-EM structure of human GS-GN complex
a Human GS cartoon model shown in ribbons and a 4 Å reconstructed cryo-EM map shown in transparent surface and coloured by corresponding chain. A model of the last 34 residues of the C-terminus of human GN (GN^{34}) is displayed in green. **b** Cryo-EM density map shown at a lower threshold (left). The N- and C-terminal tails of two protomers converge and form the “spike” region (right). **c** Focussed 3D classification of the phosphoregulatory “spike” region reveals flexibility of this inter subunit domain. 3D classes and percentage population are indicated with different colours and the 3D refined map of the core region is coloured blue.

201 **Phosphorylated human GS is in the inactive state**
202
203 GS is regulated by both allosteric activation by G6P and inhibition via phosphorylation of its
204 N- and C-terminal tails[2]. Mechanistic and structural studies of the yGS have elegantly
205 dissected its allosteric activation by G6P, whereby G6P binding induces large rotational and
206 translational movements, leading to a more accessible active site for the growing glycogen
207 chain to be fed into GS for elongation[27,30]. However, the mechanisms governing GS activation
208 by G6P and inactivation by phosphorylation, and how these opposing regulatory pressures
209 are coupled at the molecular level remains unknown.
210
211 An advantage of expressing HsGS in insect cells is that these preparations are heavily
212 phosphorylated. We mapped phosphorylation sites by tandem mass spectrometry (MS/MS)
213 following proteolysis with either trypsin, chymotrypsin or elastase resulting in a sequence
214 coverage of 97%, which is higher than the 73%[31] and 65%[32] sequence coverage achieved in
215 previous studies (**Table 1 and Supplementary Fig. 7b**). This analysis identified canonical
216 sites 2, 3a, 3b, 4 and 5 (S8, S641, S645, S653, S657), and also non-conventional sites (S412,
217 S652, S727, S731). In addition, we could detect HsGS site 2 (S8) phosphorylation for the first
218 time in a recombinant enzyme preparation. Together, these results show that expression in
219 insect cells is sufficient to achieve a high level of phosphorylation providing suitable enzyme
220 preparations for understanding inactive GS.
221
222 As expected, our highly phosphorylated enzyme has a low basal (-G6P) level of activity (**Fig.**
223 **3a**). Protein phosphatase 1 (PP1) and lambda phosphatase (lambda PP) treatment resulted
224 in faster migration of GS in an SDS-PAGE and also a reduction in signal for specific
225 phosphorylation site antibodies (**Fig. 3b and Supplementary Fig. 7a**). We observed over 10-
226 fold increase in basal activity when GS is dephosphorylated by both PP1 and lambda PP (**Fig.**
227 **3a and 3b**). Both the phosphorylated and dephosphorylated GS forms were similarly active
228 after addition of G6P (**Fig. 3a**), which is consistent with studies using GS from endogenous
229 sources[37,38].
230
231
232
233
234
235

236 **Table 1:** Summary of GS phosphorylation site analysis. LTP and HTP refer to low throughput
237 site determination (methods other than mass spectrometry) and high throughput analysis
238 (mass spectrometry only), respectively[39]

Site	Sequence	Peptide modifications	ptmRS Score	MASCOT score	Enzyme	PhosphoSite
S8	PLNRTLsMS	S7(Phospho)	S7: 100	40	Elastase	15 LTP & 1 HTP
S8	PLNRTLsMS	S7(Phospho); M8(Oxidation)	S7: 100	31	Elastase	
S8	sMSSLPGLEDW	S1(Phospho)	S1: 99.71	13	Chymotrypsin	
S412	ESLLVGsLPDMNKMLDKEDF	S7(Phospho)	S7: 100	31	Chymotrypsin	17 HTP
S412	ESLLVGsLPDMNKML	M14(Oxidation)	S7: 100	23	Chymotrypsin	
S641	QGYRYPRPAsVPPSPS	S10(Phospho)	S10: 99.99	20	Elastase	32 LTP & 30 HTP
S641 & S645	QGYRYPRPAsVPPsPS	S10(Phospho); S14(Phospho)	S10: 100; S14: 100	24	Elastase	(645) 21 LTP & 39 HTP
S652	RHsSPHQSEDEEDPRNGPL	S3(Phospho)	S3: 99.78	34	Elastase	15 HTP
S657	RHssPHQsEDEEDPRNGPL	S3(Phospho); S4(Phospho); S8(Phospho)	S8: 100	19	Elastase	(653) 13 LTP & 27 HTP
S652 & S657	RHsSPHQsEDEEDPRNGPL	S3(Phospho); S8(Phospho)	S3: 99.67; S8: 100	53	Elastase	(657) 14 LTP & 50 HTP
S727	RNSVDTATSSSLSTPSEPLsPTSSLGEER	S20(Phospho)	S20: 100	66	Trypsin	1 LTP & 22 HTP
S727	STPSEPLsPTSSL	S8(Phospho)	S8: 99.63	20	Chymotrypsin	
S731	STPSEPLSPTSsLGEERN	S12(Phospho)	S12: 99.79	73	Chymotrypsin	6 HTP
S731	TPSEPLSPTSsL	S11(Phospho)	S11: 100	22	Elastase	
S727 & S731	NSVDTATSSSLSTPSEPLsPTSsLGEER	S19(Phospho); S23(Phospho)	99.52	60	Trypsin	see above
S727 & S731	STPSEPLsPTSsLGEERN	S8(Phospho); S12(Phospho)	S8: 100; S12: 100	42	Chymotrypsin	see above

239

240 **<u>GS contains a dynamic and tuneable phosphoregulatory apparatus</u>**

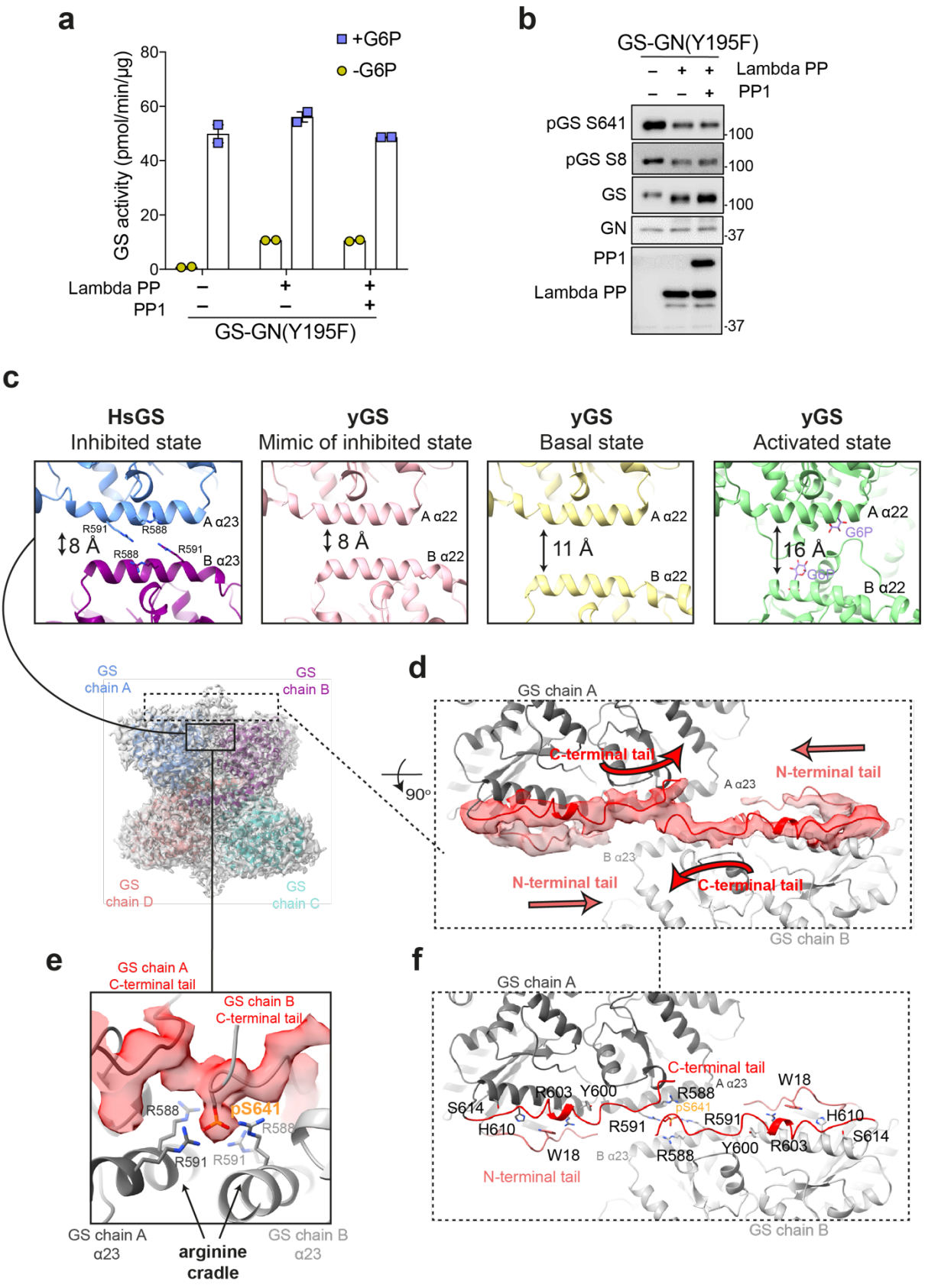
241

242 Comparisons between human, *C. elegans* and yeast GS structures are consistent with the
243 human structure in the inactive state. Each HsGS protomer shows a closed conformation of
244 its active site, and a regulatory loop, that only becomes ordered upon G6P binding, is
245 disordered in the human structure (**Supplementary Fig. 8**). Previous studies have suggested
246 that phosphorylated tails may be able to engage the G6P binding site and directly compete
247 with G6P. However, our EM density maps show no extra density within the G6P binding site
248 (**Supplementary Fig. 8a**). Thus, we see no evidence to support the hypothesis that the
249 phosphorylated tails directly interact with residues lining the G6P pocket to directly compete
250 with G6P binding. Instead, we posit that the phosphoregulatory regions indirectly affect G6P
251 binding by constraining the opening and closing of the GS tetramer.

252

253 When comparing HsGS to previous yeast crystal structures, the distance between regulatory
254 helices (α23 in humans, α22 in yeast) in adjacent monomers changes according to the activity
255 state of GS (**Fig. 3c**). Using the Cα atoms on residue 591 on chain A and residue 580 on
256 chain B (and the corresponding in the yeast proteins) the helices are furthest apart, at 16 Å,
257 when G6P is bound and GS is in its high activity state, as there is better access for accepting
258 the substrate[27,30]. When no G6P is bound and there is no phosphorylation, GS is in the basal
259 state and the helices lie 11 Å apart. In a yGS structure of a mimic of the inhibited state, where
260 residues R589 and R592 were mutated to Ala, the helices are closest together, at 8 Å, leading
261 to the closure of the N-terminal domain and thus acceptor access is hindered[30]. HsGS
262 produced in insect cells is phosphorylated[31,32] (**Fig. 3b, Supplementary Fig. 7**) and the
263 helices lie 8 Å apart, thus confirming that we have reconstructed the first phosphorylated, auto-
264 inhibited structure (**Fig. 3c**).
265 The N- and C-terminal tails of each GS protomer lie parallel to each other, and travel side by
266 side along the GS tetrameric core to reach centre. Here, they meet the N- and C-terminal tails

267 from an adjacent GS protomer, which have travelled from the opposite direction (**Fig. 3d**).
268 From the electron density map generated without the application of symmetry averaging, it
269 appears that one C-terminal tail disengages with the GS core earlier than the other C-terminal
270 tail from the adjacent chain. The C-terminal tail from chain B continues to travel further across
271 the regulatory helices than chain A, prior to traversing away from the core to form the "spike"
272 region (**Fig. 3d**). This allows chain B to engage with the regulatory helices α23, specifically
273 interacting with residues R588 and R591, which come from two GS protomers to form a
274 positively charged pocket we have termed the "arginine cradle" (**Fig. 3e**). Judging by the
275 distance from the last discernible residue of the C-terminal tail and the available cryo-EM
276 density we have modelled a phosphorylated S641 (pS641) in chain B in interacting distance
277 with the arginine cradle (**Fig. 3d**). The position of S641 is consistent with the AlphaFold[40]
278 predicted model of HsGS (**Supplementary Fig. 8d**) suggesting that this interaction may also
279 be possible in the non-phosphorylated state, although the negative charge on the phosphate
280 group would naturally provide stronger interactions with the positively charged arginine cradle.
281 S641 is a major phosphorylation site involved in activity regulation, and interaction of pS641
282 with the arginine cradle from helix α23 provides a crucial activity switch mechanism from a
283 tense (phosphorylated) state to a relaxed (G6P-bound) state. The involvement of helices α23
284 which also interacts with G6P via the nearby arginine residues R582 and R586
285 (**Supplementary Fig. 8a**) is significant, since this could provide a possible explanation
286 between G6P-binding and its ability to override inactivation by phosphorylation.


a
+G6P
-G6P
GS activity (pmol/min/µg)
80
60
40
20
0
Lambda PP
PP1
GS-GN(Y195F)
b
GS-GN(Y195F)
Lambda PP
PP1
pGS S641
pGS S8
GS
GN
PP1
Lambda PP
100
100
100
37
37
c
HsGS
Inhibited state
yGS
Mimic of inhibited state
yGS
Basal state
yGS
Activated state
8 Å
R591
R588
A α23
B α23
8 Å
A α22
B α22
11 Å
A α22
B α22
16 Å
A α22
G6P
B α22
GS chain A
GS chain B
GS chain D
GS chain C
d
GS chain A
C-terminal tail
N-terminal tail
A α23
90°
B α23
N-terminal tail
C-terminal tail
GS chain B
e
GS chain A C-terminal tail
GS chain B C-terminal tail
R588
pS641
R588
R591
R591
GS chain A α23
arginine cradle
GS chain B α23
f
GS chain A
C-terminal tail
A α23
W18
R603
Y600
R588
S614
pS641
H610
R591
R591
H610
S614
W18
B α23
R588
Y600
R603
N-terminal tail
GS chain B

287

Fig. 3 Dynamic and tuneable phosphoregulatory region of human GS
a Activity of GS-GN(Y195F) with and without the addition of lambda phosphatase (lambda PP) and protein phosphatase 1 (PP1). Upon G6P saturation, GS reaches similar activity levels regardless of phosphorylation state. Data are mean +/- S.E.M. from n=2 and representative of two independent experiments. **b** Western blot for the indicated human GS phosphorylation sites and total GS. **c** Comparison of distances between regulatory helices of adjacent monomers of human GS (reported here), low activity inhibited mimic (PDB ID 5SUL), basal state (PDB ID 3NAZ) and G6P activated (PDB ID 5SUK) yeast crystal structure. Distances were measured from the Cα of residues 591 chain A and 580 chain B (and corresponding yeast residues). **d**. The N- and C- terminal tails of one protomer lie parallel and move towards the bottom of the "spike" region, meeting the N- and C-terminal tails of the adjacent GS protomer **e** Electron density for phosphorylated S641 (pS641) interacting with R588 and R591 on the regulatory helices α23. **f** Residues that are interacting with the N- and C-terminal tails that are mutated in this study are shown.

Aligning the structure of the C-terminal domain residues of human, yeast and *C. elegans* GS (residues 294-597 in HsGS, residues 283-598 in yGS and residues 312-679 in CeGS) revealed that helix α3 (human residues 85-96) is shifted towards the C-terminal domain, compared to the yeast basal state. This N-terminal closure observed when comparing the yeast basal and mimic of the phosphorylated state crystal structures was suggested to contribute to the inhibited state of GS[30]. Consistent with these analyses, we observe an even bigger shift of helix α3 and a tight N-terminal closure in the phosphorylated human GS (**Supplementary Fig. 8c**).

The position of the extreme N-terminus is noticeably different in human and *C. elegans* GS structures compared to yeast (**Supplementary Fig. 6c**). The majority of the first β sheet in all structures is in the same place, however human residues before 26 (residue 7 in yeast) move in the opposing direction to yeast (**Supplementary Fig. 6c**). This positioning of the human GS N-terminus is directed towards the regulatory helices (α23). Previous structural investigation of CeGS-GN[34] suggested a hypothesis where phosphorylation could enable the N-terminus to engage with regulatory helices[19]. Our human phosphorylated structure supports this hypothesis, although the current density does not allow model building before residue 14. However, weak density for the N-terminus is present showing that this region also interacts with the globular GS domain and regulatory helices, thus explaining N-terminal phospho tail contribution to GS inactivation (**Fig. 3d and Supplementary Fig. 6d**). Collectively, our structural analyses support a model by which phosphorylated N- and C-terminal tails inhibit the GS tetramer by constraining a tense conformation through inter-subunit interactions.

Dislodging the GS phosphoregulatory regions

Due to the flexibility evident in the N- and C-terminal tails, we were unable to build phosphorylated residues in the cryo-EM map other than pS641. However, we can see the

316 beginning of the flexible phosphoregulatory "spike" region and residues from the GS "core
317 tetramer" which interact with this regulatory region (**Fig. 3f**). To investigate the relationship
318 between allosteric regulation and inhibitory phosphorylation and elucidate the mechanism of
319 inactivation, we mutated residues in GS that contact the beginning of the phosphoregulatory
320 region. We selected residues which are not involved in G6P binding and mutated these in
321 order to "dislodge" the regulatory tails (**Fig. 3f and Supplementary Fig. 8e**). If the
322 phosphorylated tails are indeed holding GS in an inactivated state, weakening the interaction
323 between the core tetramer and the "spike" regions should create an enzyme with higher basal
324 activity in comparison to the WT. Consistent with our hypothesis, we observed a marginal
325 increase in basal (-G6P) GS activity in R588A+R591A, Y600A, R603A, H610E and W18A
326 mutants, that was reflective of the phosphorylated state at residues S8, S641 and S645 (**Fig.
327 4a and 4b**). These mutants were unaffected in terms of GN co-purification and with the
328 exception of R588A+R591A mutant, they had similar melting (T_m) profiles and oligomeric state
329 to the WT GS complexes (**Fig. 4e, Supplementary Fig. 3b, 9a and 9b**). All mutants except
330 Y600A could still be activated to similar levels to the WT upon addition of G6P (**Fig. 4a)**.
331
332 Upon addition of PP1 and lambda PP, the GS mutants R588A+R591A, R603A and W18A
333 were more easily dephosphorylated at S641 and S8 than WT (**Fig. 4d**). For the W18A mutant,
334 this dephosphorylation by both lambda-PP and PP1 resulted in over a 20-fold increase in
335 basal activity, and also an approximately 3-fold increase in comparison to WT (**Fig. 4c**). The
336 GS R603A-GN(Y195F) mutant has a basal activity similar to WT upon dephosphorylation.
337 However, the robust dephosphorylation at S641 and S8 in GS R588A+R591A was not
338 associated with an increase in activity (**Fig. 4c**). As described above, R588 and R591 lie on
339 the regulatory helices and are also involved in inter-subunit interactions and form the arginine
340 cradle that interacts with pS641 (**Fig. 3c**). In the GS R588A+R591A double mutant we see
341 dissociation of the complex in mass photometry (**Supplementary Fig. 9c**). Therefore, the role
342 of these residues in stabilising the GS tetramer may be the cause for the lack of rescue of
343 activity upon dephosphorylation (**Fig. 4c**). Moreover, dephosphorylated GS had a markedly
344 lower T_m (48 ºC) than WT or mutant GS supporting the idea that phosphorylation of the "spike"
345 regions strengthens the inter-subunit interactions within the tetramer and holding the enzyme
346 in the "tense" conformation.

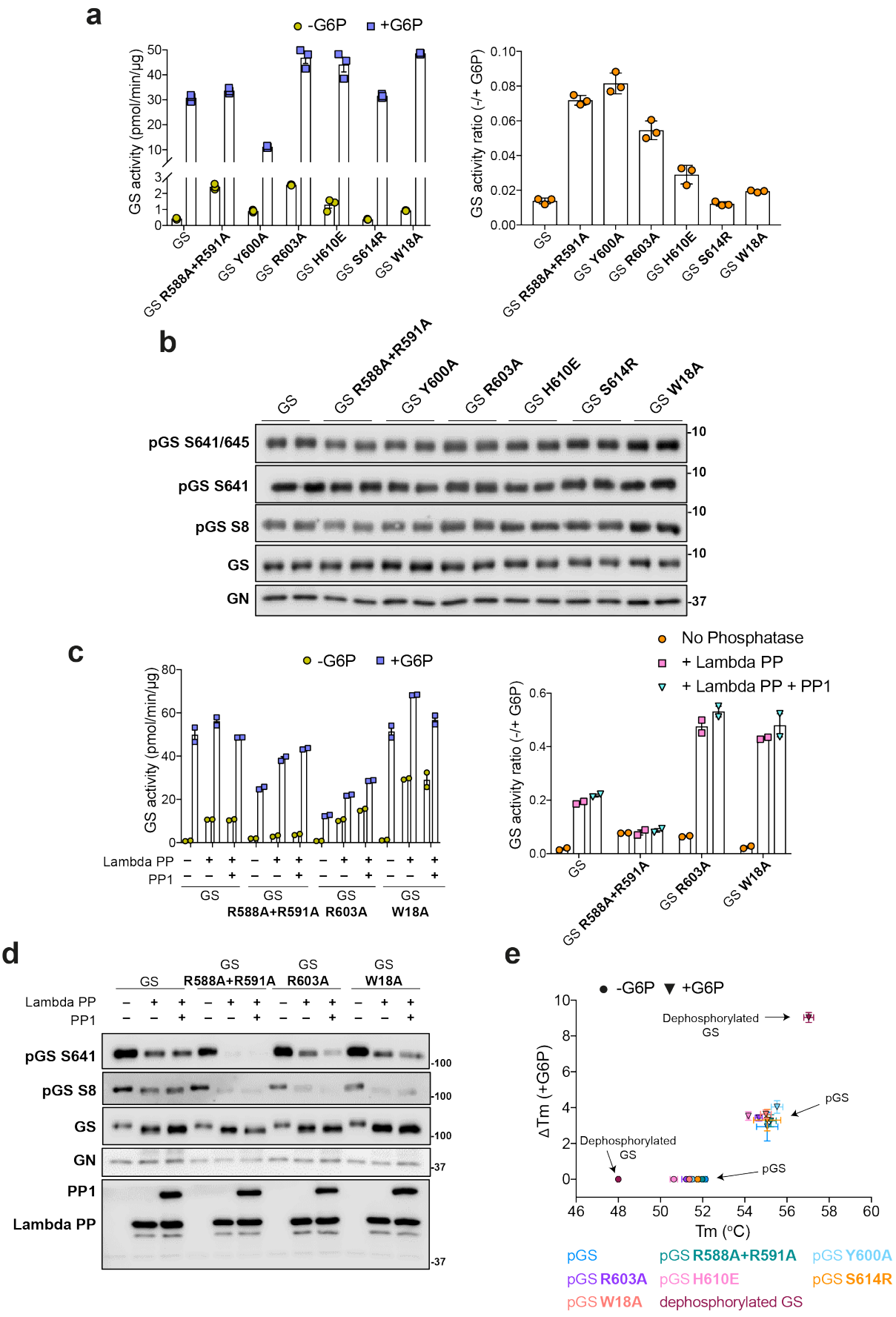

a
-G6P
+G6P
GS activity (pmol/min/µg)
GS activity ratio (-/+ G6P)
GS
GS R588A+R591A
GS Y600A
GS R603A
GS H610E
GS S614R
GS W18A
b
pGS S641/645
pGS S641
pGS S8
GS
GN
c
No Phosphatase
+ Lambda PP
+ Lambda PP + PP1
Lambda PP
PP1
d
pGS S641
pGS S8
GS
GN
PP1
Lambda PP
e
ΔTm (+G6P)
Tm (°C)
Dephosphorylated GS
pGS
pGS R588A+R591A
pGS Y600A
pGS R603A
pGS H610E
pGS S614R
pGS W18A
dephosphorylated GS

Fig. 4 Dislodging the GS phosphoregulatory region increases basal activity and increases accessibility for phosphatases
a Activity of GS WT and indicated mutants in the GS-GN(Y195F) complex in the presence and absence of G6P (left) and -/+ G6P activity ratio (right). Data are mean +/- S.E.M from n=3 and representative of two independent experiments. **b** Western blot for specific human GS phosphorylation sites S641/645, S641, S8, and total GS and GN. **c** Activity of phosphorylated and dephosphorylated GS WT and indicated mutants (left) and -/+ G6P activity ratio (right). Data are mean +/- S.E.M from n=2 and representative of two independent experiments. **d** Western blots of GS WT and mutants in the GS-GN(Y195F) complex after dephosphorylation with PP1 and/or lambda phosphatase (lambda PP). **e** Melting temperature (Tm) of GS WT and mutants in the GS-GN(Y195F) complex. Changes in melting temperature upon addition of 12.5 mM G6P ($\Delta T_m = T_m^{+G6P} - T_m^{-G6P}$). Data in **e** are mean +/- S.E.M from n = 3 experiments carried out in technical duplicates (dephosphorylated GS) and triplicates (WT and mutant GS).

Discussion

Here, we provide structural and biochemical analysis of phosphorylated human GS in the full-length GS-GN complex. Cryo-EM analysis reveals an inter-subunit flexible region which holds GS in an inactive conformation, shedding light on how phosphorylation of the N- and C-terminal tails inhibit the enzyme (**Fig. 5**). In addition, negative stain EM maps reveal two GN dimers binding to a GS tetramer, explaining the conformational plasticity of this octameric enzyme complex and the inner workings of how GS-GN cooperate to synthesize glycogen (**Fig. 1 and Fig. 5**). The two GN dimers neighbouring a GS tetramer do not interact in an identical manner, with one GS dimer tilted closer towards GS in comparison to the other (**Fig. 1f**). This observed flexibility of GN shows the inherited dynamic nature within the complex which can be aided by the variable length linker connecting the catalytic domain and the C-terminal GN[34] region that anchors GS. The precise functional relevance of this movement is yet to be explored, however the linker length was shown to govern glycogen particle size and molecular weight distribution *in vitro*[19]. Thus, the ability for GN to orient and bind flexibly to GS may be a desirable feature facilitating the wide range of size and distribution of glycogen particles seen in multiple species and tissues[4]. Unfortunately, this inherent mobility also prevented a high-resolution reconstruction of GN structure and identification of any additional interfaces between GS and the globular catalytic domain on GN.

GS phosphorylation sites lie outside of the catalytic core on the N- and C- terminal tails (**Fig. 1a and Supplementary Fig. 2**). A mechanism by which phosphorylation could detach this N-terminal region from its position, and move to interact with the arginine residues in the regulatory helices α23 was previously proposed[19]. Our structure of the phosphorylated state of GS supports this hypothesis, as the human N-terminal tail lies in the same direction as *C. elegans* N-terminal region situated against the regulatory helices, although the exact position differs (**Supplementary Fig. 6c**). More importantly, in human GS the C-terminus is

377 responsible for the majority of the observable interactions with helix α23. R588 and R591 are
378 not conserved in *C. elegans* (**Supplementary Fig. 2**), perhaps explaining the positional
379 differences between human and *C. elegans* GS tails. This exchange of interactions between
380 N- and C-terminus suggests a functional redundancy between the multiple phosphorylation
381 sites and a likely explanation for the dynamic nature of the "spike" regions.
382
383 The C-terminus of human GS runs parallel to the N-terminal tail, travelling alongside the GS
384 core, prior to moving away from the core to form the "spike" region (**Fig. 3d**). Unexpectedly,
385 the C-terminal tail of adjacent protomers are not identical. One tail appears to disengage from
386 the GS core sooner than the other. This leaves room for one tail to interact with the regulatory
387 helix, specifically residue pS641 interacting with R588 and R591 on both protomers (**Fig. 3d**
388 **and 3e**). S641 was identified as a site of phosphorylation by both phosphorylation mapping
389 and western blotting (**Table 1, Supplementary Fig. 7b and Fig. 3b).** This negative moiety
390 was coordinated by multiple arginine residues from the arginine cradle, and suggests a crucial
391 role for phosphorylation in interacting with the regulatory helices. A negative charge from
392 pS641 could be strengthening the interaction between two GS protomers, in comparison to
393 basal and activated GS, thus constraining the GS tetramer and leading to an inactive enzyme
394 (**Fig. 3c**).
395
396 The non-identical/different positions of the N- and C-terminal tails as well as the flexibility of
397 the "spike" region indicates a high level of coordination and communication both between the
398 N- and C- termini of a single GS protomer, as well as between protomers (**Fig. 2c and 3d**).
399 Having one tail buttressed against the regulatory helix and the other steering away from the
400 core may allow interchanging of the tails based on the level of phosphorylation, revealing a
401 significance for the existence of multiple phosphorylation sites. Having one tail more
402 accessible to phosphatases may also aid/allow rapid dephosphorylation of GS, leading to an
403 increase in GS activity and thus glycogen synthesis (**Fig. 5**).
404
405 Elucidating the role of the inter-subunit domain that house the phosphorylation sites through
406 mutations that weaken the interactions between the core tetramer and the "spike" regions,
407 resulted in basal activity equal to or higher than the WT, yet retaining activation by G6P (**Fig.**
408 **4a**). The GS Y600A mutant was not activated by G6P to the same extent as WT (**Fig. 4b**),
409 and although Y600 does not directly bind to G6P, UDP or sugars[27,41,42], it is possible that this
410 hydrophobic residue is important for interdomain movements which are required for full GS
411 activation.
412

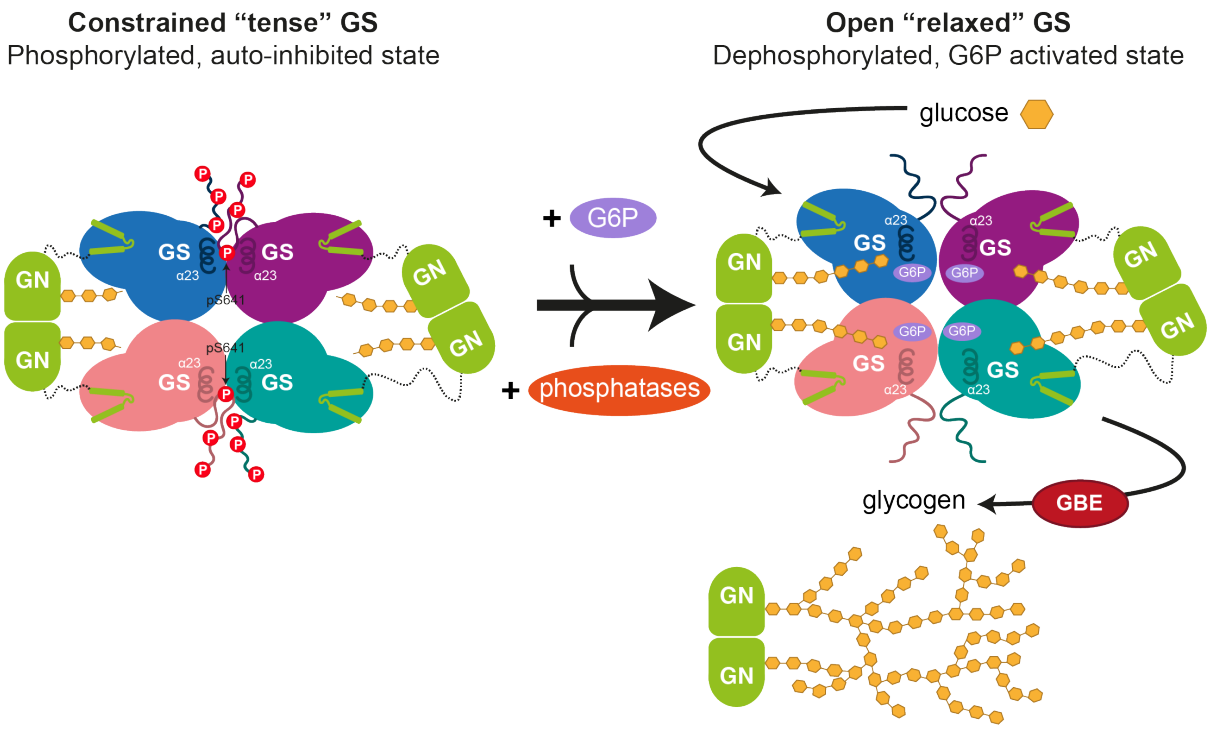


Fig. 5 GS and GN cooperate to synthesize glycogen
Glucose is converted into glycogen through the action of glycogenin (GN), glycogen synthase (GS) and glycogen branching enzyme (GBE). GN interacts with GS to feed the initial glucose chain into the GS active site for elongation. GS is regulated by allosteric activation and inhibitory phosphorylation. pS641 from one C- terminal tail interacts with the regulatory helices α23.

Dephosphorylation of the GS-GN(Y195F) complex resulted in an increase in basal activity, yet there is little difference in the high activity (G6P-bound) state between phosphorylated and dephosphorylated complexes (**Fig. 3a and Fig. 4d**). This is in accordance with previous studies that demonstrate that G6P can overcome inhibition by phosphorylation and restore full activity[2]. C-terminal mutants R588A+R591A and R603A and the N-terminal W18A mutant were more easily dephosphorylated than WT (**Fig. 4c**) suggesting that dislodging of the phosphoregulatory region leads to phosphorylation sites being more accessible to phosphatases.

The robust dephosphorylation of GS R588A+R591A did not result in an increase in basal activity (**Fig. 4c and 4d**), which mirrors previous results in yeast[30]. The analogous mutations were used in yGS resulting in low basal activity, yet it could still be fully activated by G6P[30]. It was proposed that these residues in the regulatory helix are essential for keeping GS in a "spring loaded" intermediate state, and thus charge neutralization by mutation of arginine to alanine leads to the "tense" inactive state[30]. Our activity data agree with this as we don't see an increase in basal activity despite dephosphorylation at S641 and S8, although we do see a marginal difference between this mutant and WT in the phosphorylated state (**Fig. 4a and**

431 **4c**). However, it is important to note that we also see some complex dissociation with the
432 R588A+R591A mutant as evidenced by a larger dissociated complex peak in mass
433 photometry (**Supplementary Fig. 9c**). It is possible that the dephosphorylated R588A+R591A
434 mutant is also unstable and even more dissociated than the phosphorylated mutant, resulting
435 in a less active preparation.
436
437 Mutations of human GS1 and GS2 are prevalent in glycogen storage diseases. Several
438 mutations cause glycogen storage disease 0, which is due to a deficiency in GS and thus
439 reduced levels of glycogen[43]. These mutations are enriched in key binding pockets of GS,
440 affecting UDP-G, G6P and sugar binding. Some mutations affect the interaction between GS
441 and GN[34], consistent with the requirement of this interaction for glycogen synthesis[19]. Thus,
442 the structure provides a valuable resource to understand disease mutations and to assess the
443 impact residue changes have on GS-GN activity and/or interaction (**Supplementary Table 2**
444 **and Supplementary Fig. 8f**).
445
446 Our phosphorylated GS-GN structure reveals that the N- and C-terminal phosphorylation tails
447 constrain inter-subunit movement by coordinating the N-terminal lobes of GS and interaction
448 of the regulatory helices α23 primarily via an arginine-rich pocket (arginine cradle). Thus, GS
449 has evolved a mechanism by which the phosphorylated N- and C-terminal "spike" regions hold
450 GS in an inactive conformation that is relieved by dephosphorylation and/or G6P binding. We
451 propose that the dynamic nature of these "spike" regions serves the purpose of exposing
452 phosphorylated residues to phosphatases, thus allowing a "tuneable rheostat" instead of an
453 on/off switch for regulating GS activity. Collectively, our analyses of the human GS-GN
454 enzyme complexes reveal important mechanistic and structural details that could improve our
455 understanding of glycogen storage diseases.
456
457
458 **Materials and Methods**
459
460 **Materials**
461 Total GN antibody (S197C, third bleed) was obtained from MRC-PPU Reagents and Services.
462 Total GS (#3898) and p-GS S641 (#47043) antibodies were from Cell Signaling Technologies.
463 p-GS S641/S645 (07-817) is from MerckMillipore. p-GS S8 antibody (YZ5716, first bleed) was
464 generated by (YenZym Antibodies Brisbane, CA, USA) by immunisation with a phosphorylated
465 peptide of the mouse GYS1 sequence (PLSRSL-*S-VSSLPG-amide, the prefix * denotes the
466 phosphorylated residue). Secondary antibodies (711-035-152 and 713-035-147) were

467 obtained from BioRad. Glucose-6-phosphate (G6P) (10127647001) is from Roche. All other
468 chemicals if not noted otherwise are from Sigma Aldrich.
469
470 **Cloning, protein expression and purification of GS-GN complex**
471 Genes encoding human GS1 (HsGS:NM 002103) and human GN1 (HsGN:NM 001184720)
472 mutant were cloned into pFL a vector[44]. A single 6 x His purification tag followed by a cleavable
473 site was engineered at the N-terminus of GN WT or GN Y195F mutant. For co-expression of
474 WT GS and mutants the genes encoding human GS1 and human GN1 (Y195F) were cloned
475 in pFastBac vectors, both with a 6 x His purification tag followed by a TEV site at the N-
476 terminus. Recombinant bacmids were generated in DH10Bac™ cells. Virus amplification and
477 cell expression in *Spodoptera frugiperda* (sf9) cells and *Trichoplusia ni* (Tni) cells respectively
478 were carried out using standard procedures[45]. For co-infection of pFastBac clones, a 10:1
479 ratio of the GS:GN P2 virus ratio was used. A PCR based site directed mutagenesis was used
480 to create the following mutants from the pFastBac GS1 construct: W18A, R588A+R591A,
481 Y600A, R603A, H610E, S614R. All of the alterations were confirmed by DNA sequencing.
482
483 Cell pellets containing HsGS-GN, HsGS-GN(Y195F) and mutants were resuspended in lysis
484 buffer (50 mM Tris-HCl pH 7.6, 300 mM NaCl, 20 mM imidazole, 10% glycerol, 0.075% β-
485 mercaptoethanol, 1mM benzamidine, 0.8 mM phenylmethyl sulfonyl fluoride (PMSF), 0.3
486 mg/mL lysozyme). Cells were lysed by sonication (1 second on, 3 seconds off for a total of 5
487 minutes) on ice and the lysate was cleared by centrifugation at 35,000*g* for 30 minutes at
488 4 °C. The clarified lysate was sonicated again (1 second on, 3 seconds off for a total of 1
489 minute), followed by filtering with a 0.45 µM filter (MerckMillipore). Filtered lysate was loaded
490 onto a pre-equilibrated 1 mL or 5 mL HisTrap HP column (GE Healthcare) charged with Ni^{2+}.
491 The loaded column was washed with four column volumes (CV) of low salt buffer (50 mM Tris-
492 HCl pH 7.6, 300 mM NaCl, 20 mM imidazole, 10% glycerol, 0.075% β-mercaptoethanol, 1mM
493 benzamidine), followed by four CV washes of high salt buffer (50 mM Tris-HCl pH 7.6, 500
494 mM NaCl, 20 mM imidazole, 10% glycerol, 0.075% β-mercaptoethanol, 1mM benzamidine)
495 and finally 4 CV washes in low salt buffer. The column was then attached to the AKTA system
496 (GE Healthcare) and washed with low salt buffer. The protein was then eluted by applying an
497 imidazole gradient with elution buffer (50 mM Tris-HCl pH 7.6, 300 mM NaCl, 300 mM
498 imidazole, 10% glycerol, 0.075% β-mercaptoethanol, 1 mM benzamidine). The fractions
499 containing protein were analysed by SDS-PAGE and then pooled and dialysed overnight
500 (10,000 MWCO SnakeSkin dialysis tubing (Thermo Scientific)) at 4 °C in dialysis buffer with
501 TEV protease added (50 mM Tris-HCl pH 7.6, 150 mM NaCl, 20 mM imidazole, 10% glycerol,
502 0.075% β-mercaptoethanol, 1 mM benzamidine). The dialysed protein was re-loaded onto the

503 HisTrap column equilibrated with low salt buffer, for a Ni subtraction step. TEV cleaved protein
504 was eluted in the flow through and low salt washes. The flow through and first low salt wash
505 were pooled and concentrated using a VIVASPIN20 30,000 MWCO (Sartorius, Generon),
506 followed by centrifugation at 17,000*g* for 15 minutes at 4 °C. Protein was then injected onto a
507 16/600 or 10/300 Superdex200 column (GE Healthcare) equilibrated with gel filtration buffer
508 (25 mM HEPES pH 7.5, 150 mM NaCl, 1 mM TCEP, 10% glycerol). Fractions containing
509 protein were analysed by SDS-PAGE, and fractions containing GN were pooled, concentrated
510 and stored at -80 °C. Some fractions containing GS-GN complex were stored separately at -
511 80 °C and the remaining protein was pooled and concentrated before being stored at -80 °C.
512 Proteins were visualized by Coomassie blue staining, and glucosylated species were detected
513 using the periodic acid-Schiff (PAS) method (Glycoprotein staining kit, Thermo Scientific).
514

515 ***In vitro* dephosphorylation of GS-GN**

516 Protein phosphatase 1 (PP1) and lambda phosphatase (lambda PP) were bought from MRC
517 PPU Reagents and Services. Both have an N-terminal GST tag and lambda PP also has a C-
518 terminal 6xHis tag. GS-GN complex was dephosphorylated in reactions containing equal
519 amounts of PP1 and lambda PP in 25 mM HEPES pH 7.5, 150 mM NaCl, 1 mM TCEP, 1 mM
520 $MnCl_2$ and 10% glycerol for 30 minutes at 30 °C. For subsequent differential scanning
521 fluorimetry experiments, the phosphatases were removed by incubating the reactions with
522 GST beads for 1 hour at 4 °C. Reactions were then passed through an equilibrated 0.45 µm
523 Spin-X column (Costar, 0.45 µm cellulose acetate) and eluted by centrifugation at 16,000*g* for
524 2 minutes.
525

526 **Negative stain electron microscopy – grid preparation and data collection**

527 HsGS-GN WT and Y195F were diluted in buffer (25 mM HEPES pH 7.5, 150 mM NaCl, 1 mM
528 TCEP, 10% glycerol) to concentrations between 0.01 and 0.02 mg/mL immediately before grid
529 preparation. Carbon-coated copper grids (Formvar/Carbon 300 mesh Cu, Agar Scientific)
530 were glow-discharged for 30 seconds, 10 mA and 0.39 mBar pressure (PELCO easiGlow, Ted
531 Pella). Grids were incubated for 1 minute with 6 µL sample, washed with H_2O three times and
532 stained twice with 2% w/v uranyl acetate for 20 seconds. Excess liquid was removed by
533 blotting with filter paper. Data was collected on a FEI Technai F20 electron microscope
534 operated at 120 keV, equipped with a FEI Ceta (CMOS CCD) camera.
535

536 **Negative stain electron microscopy – data processing**

537 RELION 3.0 was used for processing of negative stain-EM data[46]. Real-time contrast transfer
538 function (CTF) parameters were determined using gCTF[47]. Approximately 2,000 particles

539 were manually picked, extracted with a box size of 104 $Å^2$, then subjected to reference-free
540 2D classification to produce initial references to be used for auto-picking. The parameters for
541 auto-picking were optimized and 92,580 particles were extracted. The extracted particles were
542 used for iterative rounds of reference-free 2D classification. Based on visual inspection, best
543 quality 2D average classes were selected to generate a *de novo* 3D initial model, which was
544 used as a reference in unsupervised 3D classification. These classes were then subjected to
545 3D refinement to generate a final EM density map.
546
547 **Cryo-electron microscopy – grid preparation and data collection**
548 Quantifoil R2/1 Cu300 or Quantifoil R1.2/1.3 Cu300 (Quantifoil Micro Tools) grids were glow-
549 discharged using the GloQube plasma cleaner (Quorum) at 40 mA for 30 seconds, for GS-
550 GN(Y195F) and GS-GN respectively. A FEI Vitrobot IV was equilibrated at 4 °C at 100%
551 relative humidity. GS-GN(Y195F) complex was diluted in buffer (25 mM HEPES pH7.5, 150
552 mM NaCl, 1 mM TCEP) to 0.23 mg/mL (1.85 µM) containing 3% glycerol immediately before
553 3 µL was added to the grid. GS-GN was diluted to 0.36 mg/mL (2.97 µM) containing 8%
554 glycerol. This was followed by immediate blotting and plunge-freezing into liquid ethane cooled
555 by liquid nitrogen.
556
557 All data was collected on a FEI Titan KRIOS transmission electron microscope at 300 keV.
558 The total electron dose was 85 e/$Å^2$, a magnification of 75,000x was used and a final calibrated
559 object sampling of 1.065 Å/pixel. For GS-GN(Y195F), 8,699 movies were recorded using the
560 EPU automated acquisition software on a FEI Falcon III direct electron detector in integrating
561 mode[48]. Each movie had a total exposure time of 1.8 seconds, collected over 47 fractions with
562 an electron dose of 1.8 e/$Å^2$ per fraction. Three exposures per hole was taken and the defocus
563 values ranged from -1.7 µm to -3.1 µm.
564
565 For GS-GN, 3,009 movies were recorded using the EPU automated acquisition software on a
566 FEI Falcon III direct electron detector in integrating mode[48]. Each movie had a total exposure
567 time of 1.6 seconds, collected over 47 fractions with an electron dose of 1.8 e/$Å^2$ per fraction.
568 One exposure per hole was taken and the defocus values ranged from -1.7 µm to -3.1 µm.
569
570 **Cryo-electron microscopy – data processing**
571 For both GS-GN(Y195F) and GS-GN, drift-corrected averages of each movie were created
572 using MotionCor2[49] and real-time contrast transfer function (CTF) parameters were
573 determined using gCTF[47]. Both motion correction and CTF estimation were carried out on-the
574 fly[48]. Particles were picked using the PhosaurusNet general model in crYOLO v1.3.5[50]. For

575 GS-GN(Y195F), 780,290 particles were picked and for GS-GN, 250,250 particles were picked.
576 Particles were then imported into RELION 3.0[46], extracted with a box size of 220 pixels and
577 subjected to reference-free 2D classification.
578
579 For GS-GN(Y195F), 546,661 particles selected after 2D classification were subjected to 3D
580 classification, applying D2 symmetry. 3D classification gave a multitude of conformations,
581 indicating flexibility and heterogeneity in the dataset. To treat this heterogeneity, iterative
582 rounds of 3D classification were performed, generating 12 classes each round. Carrying all
583 "good"/unambiguous classes forward, 160,687 particles were subjected to 3D refinement and
584 postprocessing, generating a map at 4.11 Å. Followed by iterative rounds of per particle
585 contrast transfer function refinement and Bayesian particle polishing to generate a map at 4
586 Å. Final resolutions were determined using the gold-standard Fourier shell correlation criterion
587 (FSC=0.143). Local resolution was estimated using the local resolution feature in RELION.
588
589 To prevent interpretation of any artefacts created by applying D2 symmetry, the data was also
590 processed in C1 symmetry. The same particles after 2D classification were subjected to 3D
591 classification without applying symmetry. All "good"/unambiguous classes were carried
592 forward for a second round of 3D classification. High quality classes were then subjected to
593 3D refinement and postprocessing, to generate a 4.8 Å map.
594
595 To elucidate the movement of phosphoregulatory regions, an alignment free 3D classification
596 with a mask containing the "spike" density was performed, using a regularisation parameter T
597 of 40[51] (**Fig. 2c and Supplementary Fig. 4f**).
598
599 For GS-GN, 84,557 particles selected after 2D classification were subjected to 3D
600 classification. 36,972 particles (from 2 classes) were subjected to 3D refinement and
601 postprocessing, generating a map at 6.0 Å.
602
603 To explore the heterogeneity in the dataset, the 3D variability analysis[35] tool in cryoSPARC
604 v3.2.0[36] was used. The 160,687 particles after 3D classification were imported into
605 cryoSPARC and homogenous refinement with a mask was performed. The subsequent
606 particles were used in the 3D variability analysis to solve 3 modes, using C1 symmetry and
607 with a filter resolution of 7 Å. The subsequent frames were visualized in Chimera[52]. 1 mode is
608 shown in **Movie S1**.
609

610 **Model building and refinement**

611 A preliminary model of human GS was generated by AlphaFold[40]. This preliminary model was
612 rigid body fitted into the cryoEM density in UCSF Chimera[52]. The model was then built using
613 iterative rounds of manual building in COOT[53] and real space refinement in PHENIX 1.14[54]. A
614 model of the last 34 residues of human GN was created by Phyre2[55] and placed in the
615 approximate location using the *C. elegans* GS-GN[34] crystal structure (PDB ID 4QLB) in
616 Chimera[52] and then fitted into the map by rigid body fitting.

617

618 **Visualisation, structure analysis and sequence alignments**

619 Visualisation and structure analysis were performed using ChimeraX[56] or Chimera[52]. Multiple
620 sequence alignments were performed using MUSCLE[57] and displayed and edited using
621 ALINE v1.0.025[58].

622

623 **Mass photometry**

624 Mass photometry experiments were performed using a Refyn One^{MP} mass photometer.
625 Immediately prior to mass photometry measurements, proteins were diluted in 25 mM HEPES
626 pH 7.5, 150 mM NaCl, 1 mM TCEP for a final concentration of 50 nM. For each measurement,
627 (16 µL) buffer was added to a well and the focus point was found and adjusted when
628 necessary. Protein (4 µL) was then added to the buffer droplet, the sample was mixed and
629 movies of 60 seconds were recorded using AcquireMP. Data were analysed using
630 DiscoverMP.

631

632 **Differential Scanning Fluorimetry**

633 Thermal shift assays were performed using an Applied Biosystems QuantStudio 3 Real-Time
634 PCR system. SYPRO™ Orange (Invitrogen) was used as a fluorescence probe. Proteins were
635 diluted in 25 mM HEPES pH 7.5, 150 mM NaCl, 1 mM TCEP to a final concentration of 1 µM.
636 Varied concentrations of G6P were added and the reaction was incubated at room
637 temperature for 30 minutes. SYPRO Orange was diluted in 25 mM HEPES pH 7.5, 150 mM
638 NaCl, 1 mM TCEP to a final concentration of 2.5 X, in a total reaction volume of 20 µL. The
639 temperature was raised in 0.018 °C intervals from 20 °C to 95 °C. Data were analysed using
640 Protein Thermal Shift v1.4.

641

642 **Tandem mass spectrometry**

643 Concentrated purified protein complexes (6.75 µg) were diluted 30-fold in 25 mM ammonium
644 bicarbonate pH 8.0 before being subject to reduction with dithiothreitol and alkylation with
645 iodoacetamide, as previously described[59]. The eluent was equally divided into three for

646 digestion with either: 33:1 (w/w) trypsin gold (Promega), 25:1 (w/w) chymotrypsin (Promega),
647 or 10:1 (w/w) elastase (Promega), using the manufacturer's recommended temperatures for
648 18 hours with 600 rpm shaking. Digests were then subject to in-house packed, strong cation
649 exchange stage tip clean-up, as previously described by[60]. Dried peptides were solubilized in
650 20 µl of 3% (v/v) acetonitrile and 0.1% (v/v) TFA in water, sonicated for 10 minutes, and
651 centrifuged at 13,000*g* for 15 minutes at 4 °C being separated using an Ultimate 3000 nano
652 system (Dionex) by reversed-phase HPLC, over a 60-minute gradient, as described in[59]. All
653 data acquisition was performed using a Thermo Orbitrap Fusion Lumos Tribrid mass
654 spectrometer (Thermo Scientific), with higher-energy C-trap dissociation (HCD) fragmentation
655 set at 32% normalized collision energy for 2+ to 5+ charge states. MS1 spectra were acquired
656 in the Orbitrap (60K resolution at 200 *m/z*) over a range of 350 to 1400 *m/z*, AGC target =
657 standard, maximum injection time = auto, with an intensity threshold for fragmentation of $2e^4$.
658 MS2 spectra were acquired in the Orbitrap (30K resolution at 200 *m/z*), AGC target = standard,
659 maximum injection time = dynamic. A dynamic exclusion window of 20 seconds was applied
660 at a 10 ppm mass tolerance. Data was analysed by Proteome Discoverer 1.4 using the UniProt
661 Human reviewed database (updated April 2020) with fixed modification =
662 carbamidomethylation (C), variable modifications = oxidation (M) and phospho (S/T/Y),
663 instrument type = electrospray ionization–Fourier-transform ion cyclotron resonance (ESI-
664 FTICR), MS1 mass tolerance = 10 ppm, MS2 mass tolerance = 0.01 Da, and the *ptm*RS node
665 on; set to a score > 99.0.
666
667 **Glycogen synthase activity assay**
668 1 µg of purified protein was diluted in ice cold lysis buffer (270 mM sucrose, 50 mM Tris HCl
669 (pH 7.5), 1 mM EDTA, 1 mM EGTA, 1% (v/v) Triton X-100, 20 mM glycerol-2-phosphate, 50
670 mM NaF, 5 mM $Na_4P_2O_7$, 1 mM DTT, 0.1 mM PMSF, 1 mM benzamidine, 1 mg/mL
671 microcystin-LR, 2 mg/mL leupeptin, and 2 mg/mL pepstatin A) to a total volume of 100 µL. 20
672 µL of the protein was added to 80 µL of the assay buffer (25 mM Tris-HCl (pH 7.8), 50 mM
673 NaF, 5 mM EDTA, 10 mg/ml glycogen, 1.5 mM UDP-glucose, 0.125% (v/v) 2-mercaptoethanol
674 and 0.15 mCi/mmol D-[^{14}C]-UDP-glucose (American Radiolabelled Chemicals, Inc., ARC
675 0154) with 0 and 12.5 mM G6P. Reactions were incubated for 20 minutes at 30 °C with mild
676 agitation at 300 rpm. The reactions were stopped by spotting 90 µL of the reaction mix onto
677 2.5 cm x 2.5 cm squares of filter paper (Whatman 3MM) which were immediately immersed in
678 ice cold 66% ethanol and left to incubate with mild agitation for 20 minutes. The filter papers
679 were washed thrice more in 66% ethanol for 20 minutes each and rinsed in acetone. The dried
680 filters were subjected to scintillation counting.
681

682 **Immunoblotting**

683 Purified proteins were denatured in Laemmli buffer at 100 °C for 5 minutes. 100 ng of the
684 protein was separated by SDS-PAGE on 4-10% gel and transferred onto nitrocellulose
685 membranes (#926-31090, LiCOR). Membranes were blocked for 45 minutes at room
686 temperature in 5% skim milk (Sigma, 1153630500) followed by incubation in TBST (10 mM
687 Tris (pH 7.6), 137 mM NaCl, and 0.1% (v/v) Tween-20) containing 5% (w/v) BSA and the
688 primary antibody overnight at 4 ºC. The membranes were incubated for 45 minutes in HRP
689 conjugated secondary antibodies diluted 1:10,000 in 3% skim milk, after extensive washing in
690 TBST. The membranes were imaged using enhanced chemiluminescence (ECL) reagent (GE
691 Healthcare).

692

693 **Author contributions:**

694 L.M. performed the majority of experiments (molecular biology, protein production, electron
695 microscopy, differential scanning fluorimetry and mass photometry analysis), D.B. performed
696 glycogen synthase activity and western blot assays, L.D. performed phosphorylation mapping
697 and C.B. provided support with structural biology and protein production. L.M. and E.Z. drafted
698 the manuscript with input from D.B., K.S., L.D. and C.E.E. and all authors revised it. J.P.,
699 C.E.E., C.H., J.B., N.A.R., H.K., K.S. and E.Z. provided supervision and project management.
700 E.Z., K.S., C.E.E., C.H. and J.B. designed and interpreted data in consultation with all authors.

701 **Acknowledgments:**

702 L.M. is supported by an MRC Discovery Medicine North (DiMeN) iCASE studentship co-
703 funded by UKRI and Vertex Pharmaceuticals. D.B. is supported by an International
704 Postdoctoral Fellowship funded by the Novo Nordisk Foundation Center for Basic Metabolic
705 Research. C.E. and L.D. were supported by grants from BBSRC (BB/S018514/1,
706 BB/M012557/1, BB/R000182/1). E.Z. is supported by a Sir Henry Dale Fellowship from the
707 Wellcome Trust & Royal Society (200523/Z/16/Z). The Astbury cryo-EM Facility is funded by
708 a University of Leeds ABSL award and a Wellcome Trust (108466/Z/15/Z) grant. The Novo
709 Nordisk Foundation Center for Basic Metabolic Research is an independent Research Center
710 based at the University of Copenhagen, Denmark, and partially funded by an unconditional
711 donation from the Novo Nordisk Foundation (Grant number NNF18CC0034900).

712

713 **Conflict of Interest**

714 The authors report no conflicts of interest.

715

716 **References**

717

718 1. Roach, P. J., Depaoli-Roach, A. A., Hurley, T. D. & Tagliabracci, V. S. Glycogen and
719 its metabolism: some new developments and old themes. *Biochem. J.* **441,** 763–787
720 (2012).
721 2. Roach, P. J. Glycogen and its metabolism. *Curr. Mol. Med.* **2,** 101–120 (2002).
722 3. Pitcher, J., Smythe, C., Campbell, D. G. & Cohen, P. Identification of the 38-kDa
723 subunit of rabbit skeletal muscle glycogen synthase as glycogenin. *European Journal*
724 *of Biochemistry* **169,** 497–502 (1987).
725 4. Zeqiraj, E. & Sicheri, F. Getting a handle on glycogen synthase – Its interaction with
726 glycogenin. *Molecular Aspects of Medicine* **46,** 63–69 (2015).
727 5. Skurat, A. V., Dietrich, A. D. & Roach, P. J. Interaction between glycogenin and
728 glycogen synthase. *Archives of Biochemistry and Biophysics* **456,** 93–97 (2006).
729 6. Pitcher, J., of, C. S. E. J.1988. Glycogenin is the priming glucosyltransferase required
730 for the initiation of glycogen biogenesis in rabbit skeletal muscle. *European Journal of*
731 *Biochemistry* **176,** 391–395 (1988).
732 7. Ryu, J.-H. *et al.* Comparative structural analyses of purified glycogen particles from
733 rat liver, human skeletal muscle and commercial preparations. *International Journal of*
734 *Biological Macromolecules* **45,** 478–482 (2009).
735 8. Ashcroft, F. M., Rohm, M., Clark, A. & Brereton, M. F. Is Type 2 Diabetes a Glycogen
736 Storage Disease of Pancreatic β Cells? *Cell Metabolism* **26,** 17–23 (2017).
737 9. Kohler, L., Puertollano, R. & Raben, N. Pompe Disease: From Basic Science to
738 Therapy. *Neurotherapeutics* **15,** 928–942 (2018).
739 10. García-Gimeno, M., Knecht, E. & Sanz, P. Lafora Disease: A Ubiquitination-Related
740 Pathology. *Cells* **7,** 87–17 (2018).
741 11. Malfatti, E. *et al.* A new muscle glycogen storage disease associated with glycogenin-
742 1 deficiency. *Ann Neurol.* **76,** 891–898 (2014).
743 12. Pursell, N. *et al.* Inhibition of Glycogen Synthase II with RNAi Prevents Liver Injury in
744 Mouse Models of Glycogen Storage Diseases. *Mol Ther* **26,** 1771–1782 (2018).
745 13. Ashe, K. M. *et al.* Inhibition of glycogen biosynthesis via mTORC1 suppression as an
746 adjunct therapy for Pompe disease. *Molecular Genetics and Metabolism* **100,** 309–
747 315 (2010).
748 14. Mu, J., Skurat, A. V. & Roach, P. J. Glycogenin-2, a Novel Self-glucosylating Protein
749 Involved in Liver Glycogen Biosynthesis. *J. Biol. Chem.* **272,** 27589–27597 (1997).

750 15. Curtino, J. A. & Aon, M. A. From the seminal discovery of proteoglycogen and
751 glycogenin to emerging knowledge and research on glycogen biology. *Biochem. J.*
752 **476,** 3109–3124 (2019).
753 16. Lairson, L. L., Henrissat, B., Davies, G. J. & Withers, S. G. Glycosyltransferases:
754 Structures, Functions, and Mechanisms. *Annu. Rev. Biochem.* **77,** 521–555 (2008).
755 17. Gibbons, B. J., Roach, P. J. & Hurley, T. D. Crystal Structure of the Autocatalytic
756 Initiator of Glycogen Biosynthesis, Glycogenin. *Journal of Molecular Biology* **319,**
757 463–477 (2002).
758 18. Chaikuad, A. *et al.* Conformational plasticity of glycogenin and its maltosaccharide
759 substrate during glycogen biogenesis. *Proceedings of the National Academy of*
760 *Sciences* **108,** 21028–21033 (2011).
761 19. Zeqiraj, E. *et al.* Structural basis for the recruitment of glycogen synthase by
762 glycogenin. *Proceedings of the National Academy of Sciences* **111,** E2831–E2840
763 (2014).
764 20. Skurat, A. V., Wang, Y. & Roach, P. J. Rabbit skeletal muscle glycogen synthase
765 expressed in COS cells. Identification of regulatory phosphorylation sites. *J. Biol.*
766 *Chem.* **269,** 25534–25542 (1994).
767 21. Wilamowitz-Moellendorff, von, A. *et al.* Glucose-6-phosphate-mediated activation of
768 liver glycogen synthase plays a key role in hepatic glycogen synthesis. *Diabetes* **62,**
769 4070–4082 (2013).
770 22. Browner, M. F., Nakano, K., Bang, A. G. & Fletterick, R. J. Human muscle glycogen
771 synthase cDNA sequence: a negatively charged protein with an asymmetric charge
772 distribution. *Proceedings of the National Academy of Sciences* **86,** 1443–1447 (1989).
773 23. Nuttall, F. Q., Gannon, M. C., Bai, G. & Lee, E. Y. Primary structure of human liver
774 glycogen synthase deduced by cDNA cloning. *Archives of Biochemistry and*
775 *Biophysics* **311,** 443–449 (1994).
776 24. Buschiazzo, A. *et al.* Crystal structure of glycogen synthase: homologous enzymes
777 catalyze glycogen synthesis and degradation. *The EMBO Journal* **23,** 3196–3205
778 (2004).
779 25. Skurat, A. V., Peng, H. L., Chang, H. Y., Cannon, J. F. & Roach, P. J. Rate-
780 determining steps in the biosynthesis of glycogen in COS cells. *Archives of*
781 *Biochemistry and Biophysics* **328,** 283–288 (1996).
782 26. Hunter, R. W., Treebak, J. T., Wojtaszewski, J. F. P. & Sakamoto, K. Molecular
783 Mechanism by Which AMP-Activated Protein Kinase Activation Promotes Glycogen
784 Accumulation in Muscle. *Diabetes* **60,** 766–774 (2011).

785 27. Baskaran, S., Roach, P. J., Depaoli-Roach, A. A. & Hurley, T. D. Structural basis for
786 glucose-6-phosphate activation of glycogen synthase. *Proc. Natl. Acad. Sci. U.S.A.*
787 **107,** 17563–17568 (2010).
788 28. Roach, P. J. Multisite and hierarchal protein phosphorylation. *J. Biol. Chem.* **266,**
789 14139–14142 (1991).
790 29. Roach, P. J. Control of glycogen synthase by hierarchal protein phosphorylation.
791 *FASEB J* **4,** 2961–2968 (1990).
792 30. Mahalingan, K. K., Baskaran, S., Depaoli-Roach, A. A., Roach, P. J. & Hurley, T. D.
793 Redox Switch for the Inhibited State of Yeast Glycogen Synthase Mimics Regulation
794 by Phosphorylation. *Biochemistry* **56,** 179–188 (2017).
795 31. Hunter, R. W., Zeqiraj, E., Morrice, N., Sicheri, F. & Sakamoto, K. Expression and
796 purification of functional human glycogen synthase-1:glycogenin-1 complex in insect
797 cells. *PROTEIN EXPRESSION AND PURIFICATION* **108,** 23–29 (2015).
798 32. Khanna, M. *et al.* Expression and purification of functional human glycogen synthase-
799 1 (hGYS1) in insect cells. *PROTEIN EXPRESSION AND PURIFICATION* **90,** 78–83
800 (2013).
801 33. Smythe, C., Caudwell, F. B., Ferguson, M. & Cohen, P. Isolation and structural
802 analysis of a peptide containing the novel tyrosyl-glucose linkage in glycogenin. *The*
803 *EMBO Journal* **7,** 2681–2686 (1988).
804 34. Rodriguez, I. R. & Whelan, W. J. A novel glycosyl-amino acid linkage: rabbit-muscle
805 glycogen is covalently linked to a protein via tyrosine. *Biochemical and Biophysical*
806 *Research Communications* **132,** 829–836 (1985).
807 35. Punjani, A. & Fleet, D. J. 3D variability analysis: Resolving continuous flexibility and
808 discrete heterogeneity from single particle cryo-EM. *Journal of Structural Biology* **213,**
809 107702 (2021).
810 36. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms
811 for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14,** 290–296
812 (2017).
813 37. Bouskila, M. *et al.* Allosteric Regulation of Glycogen Synthase Controls Glycogen
814 Synthesis in Muscle. *Cell Metabolism* **12,** 456–466 (2010).
815 38. McManus, E. J. *et al.* Role that phosphorylation of GSK3 plays in insulin and Wnt
816 signalling defined by knockin analysis. *The EMBO Journal* **24,** 1571–1583 (2005).
817 39. Hornbeck, P. V. *et al.* PhosphoSitePlus, 2014: mutations, PTMs and recalibrations.
818 *Nucleic Acids Res* **43,** D512–20 (2015).
819 40. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature*
820 **596,** 583–589 (2021).

821 41. Baskaran, S. *et al.* Multiple glycogen-binding sites in eukaryotic glycogen synthase
822 are required for high catalytic efficiency toward glycogen. *J. Biol. Chem.* **286,** 33999–
823 34006 (2011).
824 42. Chikwana, V. M. *et al.* Structural basis for 2'-phosphate incorporation into glycogen by
825 glycogen synthase. *Proc. Natl. Acad. Sci. U.S.A.* **110,** 20976–20981 (2013).
826 43. Ellingwood, S. S. & Cheng, A. Biochemical and clinical aspects of glycogen storage
827 diseases. *Journal of Endocrinology* **238,** R131–R141 (2018).
828 44. Fitzgerald, D. J. *et al.* Protein complex expression by using multigene baculoviral
829 vectors. *Nat Methods* **3,** 1021–1032 (2006).
830 45. Zeqiraj, E. *et al.* Higher-Order Assembly of BRCC36-KIAA0157 Is Required for DUB
831 Activity and Biological Function. *Mol Cell* **59,** 970–983 (2015).
832 46. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure
833 determination in RELION-3. *eLIFE* **7,** (2018).
834 47. Zhang, K. Gctf: Real-time CTF determination and correction. *Journal of Structural*
835 *Biology* **193,** 1–12 (2016).
836 48. Thompson, R. F., Iadanza, M. G., Hesketh, E. L., Rawson, S. & Ranson, N. A.
837 Collection, pre-processing and on-the-fly analysis of data for high-resolution, single-
838 particle cryo-electron microscopy. *Nat Protoc* **14,** 100–118 (2018).
839 49. Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for
840 improved cryo-electron microscopy. *Nat Methods* **14,** 331–332 (2017).
841 50. Wagner, T. *et al.* SPHIRE-crYOLO is a fast and accurate fully automated particle
842 picker for cryo-EM. *Commun Biol* **2,** 218–13 (2019).
843 51. Scheres, S. H. W. in *Cryo-EM, Part B: 3-D Reconstruction* **482,** 295–320 (Elsevier,
844 2010).
845 52. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research
846 and analysis. *J Comput Chem* **25,** 1605–1612 (2004).
847 53. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of
848 Coot. *Acta Crystallogr D Biol Crystallogr* **66,** 486–501 (2010).
849 54. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for
850 macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66,** 213–221
851 (2010).
852 55. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2
853 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10,** 845–858
854 (2015).
855 56. Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers,
856 educators, and developers. *Protein Sci* **30,** 70–82 (2021).

857 57. Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019.
858 *Nucleic Acids Res* **47,** W636–W641 (2019).
859 58. Bond, C. S. & Schüttelkopf, A. W. ALINE: a WYSIWYG protein-sequence alignment
860 editor for publication-quality alignments. *Acta Crystallogr D Biol Crystallogr* **65,** 510–
861 512 (2009).
862 59. Ferries, S. *et al.* Evaluation of Parameters for Confident Phosphorylation Site
863 Localization Using an Orbitrap Fusion Tribrid Mass Spectrometer. *J Proteome Res*
864 **16,** 3448–3459 (2017).
865 60. Daly, L. A. *et al.* Oxygen-dependent changes in binding partners and post-
866 translational modifications regulate the abundance and activity of HIF-1α/2α. *Sci*
867 *Signal* **14,** (2021).
868