

Molecular insights into biogenesis of glycosylphosphatidylinositol anchor proteins

Yidan Xu

State Key Laboratory of Molecular Biology, CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, University of CAS, Chinese Academy of Sciences

Guowen Jia

The State Key Laboratory of Biotherapy and Cancer Center, Department of Geriatrics and National Clinical Research Center for Geriatrics, West China Hospital, Sichuan University

Tingting Li

University of Chinese Academy of Sciences <https://orcid.org/0000-0003-3146-2761>

Zixuan Zhou

Shanghai Stomatological Hospital, School of Stomatology, Shanghai Key Laboratory of Medical Epigenetics, International Co-laboratory of Medical Epigenetics and Metabolism, Fudan University

Yitian Luo

ShanghaiTech University <https://orcid.org/0000-0001-7808-9322>

Yulin Chao

Shanghai Stomatological Hospital, School of Stomatology, Shanghai Key Laboratory of Medical Epigenetics, International Co-laboratory of Medical Epigenetics and Metabolism, Fudan University

Juan Bao

Shanghai Institute of Biochemistry and Cell Biology, University of Chinese Academy of Sciences

Zhaoming Su

Sichuan University <https://orcid.org/0000-0002-9279-1721>

Qianhui Qu

Stanford University

Dianfan Li (✉ dianfan.li@sibcb.ac.cn)

State Key Laboratory of Molecular Biology, CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, University of CAS, Chinese Academy of Sciences

<https://orcid.org/0000-0003-4729-4678>

Article

Keywords:

Posted Date: March 1st, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1339620/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Nature Communications on May 12th, 2022. See the published version at <https://doi.org/10.1038/s41467-022-30250-6>.

1 **Title: Molecular insights into biogenesis of**
2 **glycosylphosphatidylinositol anchor proteins**

3
4 **Authors:** Yidan Xu^{1, †}, Guowen Jia^{2, †}, Tingting Li^{1, †}, Zixuan Zhou^{3, †}, Yitian Luo¹,
5 Yulin Chao³, Juan Bao¹, Zhaoming Su^{2, *}, Qianhui Qu^{3, *}, Dianfan Li^{1, *}

6
7 **Affiliations:**

8 ¹State Key Laboratory of Molecular Biology, CAS Center for Excellence in Molecular
9 Cell Science, Shanghai Institute of Biochemistry and Cell Biology, University of CAS,
10 Chinese Academy of Sciences (CAS), 320 Yueyang Road, Shanghai 200030, China.

11 ²State Key Laboratory of Biotherapy and Cancer Center, Department of Geriatrics and
12 National Clinical Research Center for Geriatrics, West China Hospital, Sichuan
13 University, Chengdu 610044, China.

14 ³Shanghai Stomatological Hospital, School of Stomatology, Shanghai Key Laboratory
15 of Medical Epigenetics, International Co-laboratory of Medical Epigenetics and
16 Metabolism (Ministry of Science and Technology), Institutes of Biomedical Sciences,
17 Department of Systems Biology for Medicine, Fudan University, Shanghai 200032,
18 China.

19 [†]Equal contribution

20 Correspondence: dianfan.li@sibcb.ac.cn; qqh@fudan.edu.cn; zsu@scu.edu.cn

21 Lead contact: dianfan.li@sibcb.ac.cn

22

23 **Abstract**

24 Eukaryotic cells are coated with an abundance of glycosylphosphatidylinositol anchor
25 proteins (GPI-APs) that play crucial roles in fertilization, neurogenesis, and immunity.
26 Covalent addition of structurally diverse GPI anchorages at the carboxyl termini of
27 target proteins is catalyzed by an endoplasmic reticulum integral membrane GPI
28 transamidase complex (GPI-T) conserved among all eukaryotes. Despite its important
29 role in developmental and cancer biology, a detailed picture of this intricate multi-
30 component machinery remains elusive. Here, we report the cryo-electron microscopy
31 (EM) structure of the human GPI-T complex at a global 2.53 Å resolution, revealing an
32 assembly mechanism whereby the catalytic subunit PIGK and its catalytic dyad are
33 optimally positioned to accommodate its characteristic amphipathic substrates.
34 Structural and functional characterizations suggest a previously unrecognized
35 composite GPI-binding site formed by subunits PIGU/PIGK/PIGT and rationalize
36 genetic diseases associated with GPI-T point mutations. Our work presents an important
37 step towards the mechanistic understanding of GPI-AP biosynthesis.

38 **Introduction**

39 The GPI anchoring represents a ubiquitous, metabolically expensive
40 posttranslational modification of eukaryotic cell surface proteins ¹⁻⁵. Structurally
41 elucidated in the 1980s ⁶, GPI lipids are bioactive ⁷ and chemically diverse with a
42 minimal backbone consisting of a phosphatidylinositol group linked to a polysaccharide
43 core (one glucosamine, four mannoses) modified with phosphoethanolamine (EtNP)
44 (**Fig. 1a**). These complicated anchorages (**Fig. S1a**) place GPI-APs in lipid rafts and
45 confer their unique regulatory properties in developmental and physiological processes
46 ²⁻⁵. Notable GPI-APs include LY6K/TEX101 as key factors for fertilization and
47 biomarkers for infertility ⁸, glypicans/Gas1/RECK that modulate Hedgehog/Wnt/Notch
48 signaling ^{9,10}, CD55/CD59 that inhibit complement cascade in innate immunity ¹¹,
49 alkaline phosphatase as a leading biomarker for hepatic diseases and cancers ¹², and
50 folate receptor 1 that mediates folate uptake and as an important cancer biomarker ¹³.
51 Disruption of the GPI-AP biosynthesis leads to embryonic lethality in animals ¹⁴, while
52 its biogenesis pathway in *Trypanosoma brucei* is a validated drug target for the fatal
53 sleeping sickness ¹⁵.

54 The committed step in GPI-AP biogenesis is catalyzed by the GPI transamidase
55 (GPI-T), which cleaves the C-terminal signal peptide (CSP) of precursor protein and
56 covalently links the EtNP moiety of GPI to the newly exposed carboxyl terminus of the
57 so-called ω -residue (**Fig. 1a**) ^{1,2,16}. GPI-T consists of at least five subunits (**Fig. S1b**),
58 namely PIGK/Gpi8p, PIGT/Gpi16p, PIGU/Gab1p, PIGS/Gpi17p, GPAA1/Gaa1p in
59 human/yeast, respectively (**Fig. S2**) ². For ease of description, we use the human

60 nomenclature hereafter. PIGK and GPAA1 have been proposed to execute the peptide
61 cleavage^{17,18} and GPI addition reactions¹⁹. PIGU and GPAA1 have been suggested to
62 bind GPI^{16,20}. PIGT disulfide-links with PIGK in some species²¹ and may play a
63 structural role. The function of PIGS is less clear although it is essential for GPI-T
64 activity²². All subunits are predicted to contain at least one transmembrane helix (TMH)
65 except for PIGK in some species like *T. brucei*²³. Aberrant activity of GPI-T has been
66 recently implicated in multiple pathologies²⁴⁻³³ such as *neurodevelopmental disorder*
67 with *hypotonia* and *cerebellar atrophy*, with or without *seizures* (NEDHCAS) and
68 cancers. No experimental structural information exists to explain and clarify various
69 and sometimes contradictory models for GPI-T's assembly^{22,34,35}, membrane topology
70^{20,36}, and subunit function¹⁶⁻²⁰.

71 Here, we report the high-resolution cryo-EM structure of the human
72 glycosylphosphatidylinositol transamidase (GPI-T). Combined with functional studies,
73 our work represents an important step towards the mechanistic understanding of the
74 GPI-AP biosynthesis and the pathophysiology associated with GPI-T mutations.

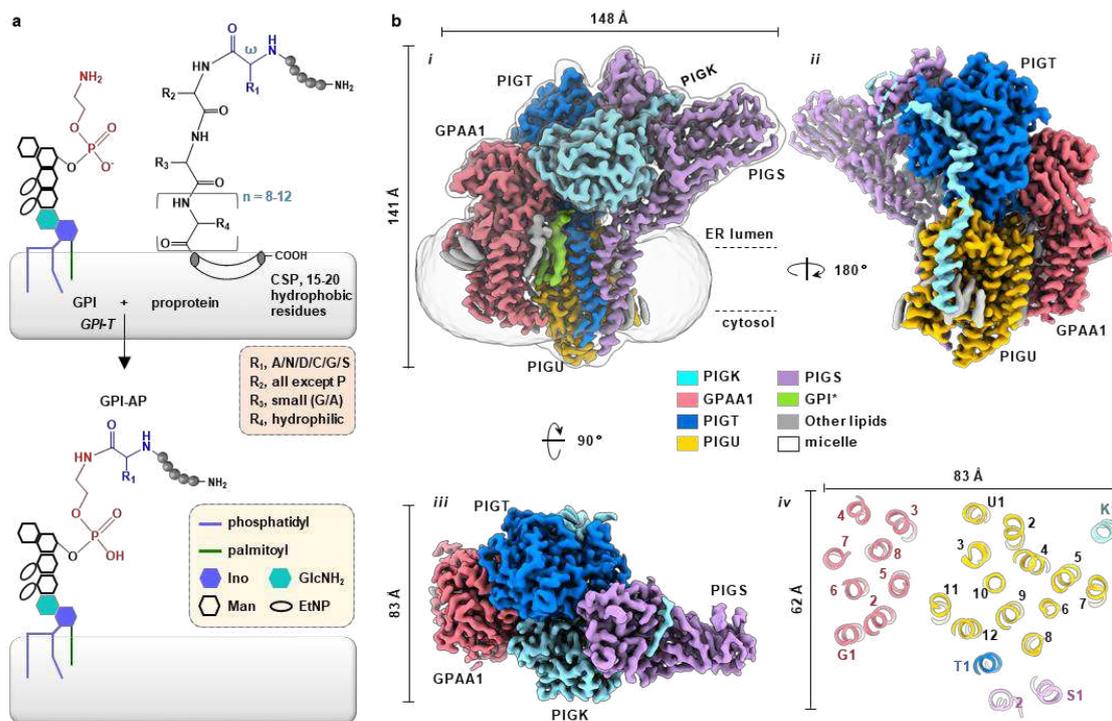
75

76 **Results**

77 **PIGT and PIGU form a platform for complex assembly**

78 To gain insights into the GPI anchoring process, we determined the cryo-EM
79 structure of the transamidase complex at 2.53 Å nominal resolution (**Figs. S3a-c, Table**
80 **S1**). The high-quality map (**Fig. S3d**) sufficed *ab initio* model building, with a total of

81 2,393 residues (94.4% completion), 3 N-glycosylation sites, 4 disulfide bonds, and 22
 82 lipid/detergent molecules. The overall GPI-T architecture assumes the shape of a canon
 83 (the luminal domain, mainly from PIGT/GPAA1/PIGS/PIGK) on a carriage (the
 84 transmembrane domain, TMD, mainly from GPAA1/PIGU) with an approximate
 85 dimension of 148 Å by 141 Å by 83 Å (**Fig. 1b**). The TMD consisting of 24 TMHs is
 86 divided to contain a small entity with eight TMHs from GPAA1 and a large entity with
 87 a TMH core (TMH1-12 of PIGU) surrounded by four satellite TMHs, two from PIGS
 88 and one each from PIGT and PIGK.



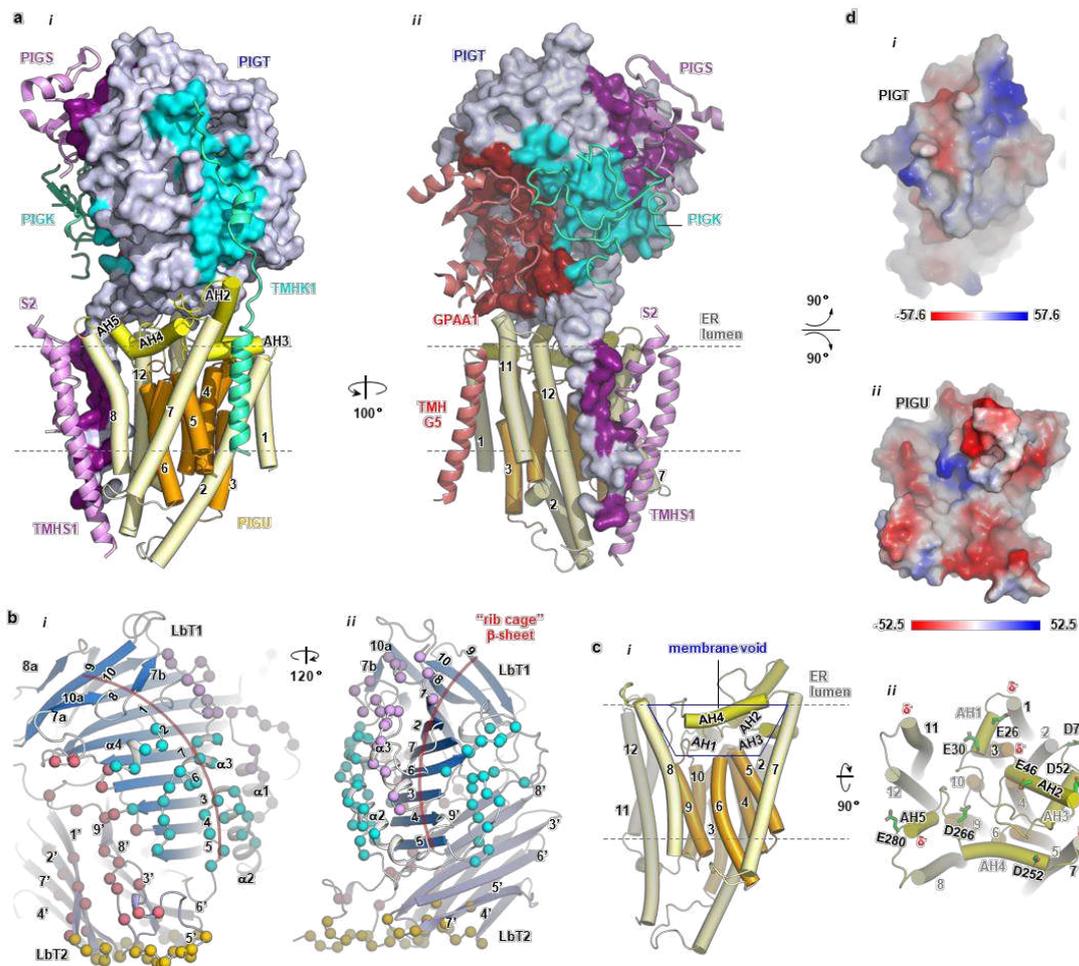
89

90 **Fig. 1. Cryo-EM map of the human GPI-T.** **a** GPI-T replaces the C-terminal signal
 91 peptide (CSP) of proproteins with GPI at the ω residue (blue) by a transamination
 92 reaction. Various parts are denoted in the dashed box. EtNP, ethanolamine; Man,
 93 mannose; Ino, inositol; GlcNH₂, glucosamine. **b** Cryo-EM map (*i-iii*) and cross section
 94 view of the transmembrane domain (*iv*). An asterisk indicates a partial GPI. Numbers

95 in *iv* indicate TMHs and G/T/U/S/K refer to GPAA1 and PIGT/U/S/K, respectively.

96

97 Central to the assembly of the GPI-T complex is a platform formed by PIGT and
98 PIGU (**Fig. 2a**). In the ER lumen, two lobes of PIGT (LbT1 and LbT2) pack together
99 with skeleton features. Specifically, ten twisted antiparallel β -strands of LbT1 stack into
100 a “half-rib” cage that was decorated with loops and short α -helices for interaction with
101 GPAA1/PIGK/PIGS, and LbT2 assumes a stable β -sandwich topology for interaction
102 with GPAA1 and PIGU (**Fig. 2b**). As a result, 25.9% of its total surface area was buried
103 by other subunits (**Fig. 2a**), conforming to its structural role as previously proposed ²².
104 In the ER membrane, PIGU recruits other subunits through an optimal geometry that
105 maximizes interaction surfaces within the membrane and at the luminal surface.
106 Specifically, its 12-TMHs are arranged into two centric rings (**Figs. 2a, c**). The outer
107 ring consisting of TMH1/2/7/8/11/12 interacts with the transmembrane domains of all
108 other four subunits alongside the membrane plane. The inner ring, however, contains
109 six short TMHs that do not fully transverse the membrane, creating a hydrophobic void
110 to strategically attract the five amphipathic helices (AH1-5) (**Fig. 2c**). In turn, these
111 AHs expose several acidic residues and dipole moments (**Fig. 2c**) to create an overall
112 negatively charged surface for electrostatic complementation with the corresponding
113 PIGT surface (**Fig. 2d**). Thus, PIGU is arranged optimally as a docking base for PIGT
114 and together with PIGT to buttress all other subunits.



115

116 **Fig. 2. PIGT and PIGU form a platform for complex assembly.** **a** PIGT (surface

117 representations, blue) and PIGU (cylinder, yellow/orange) form a docking platform for

118 other subunits (GPAA1, red; PIGS, purple; PIGK, cyan, ribbon). Interaction surfaces

119 on PIGT are shaded to match the color of GPAA1/PIGK/PIGS. **b** PIGT displays

120 skeleton features. Its luminal domain contains two lobes (LbT1/2). LbT1 (marine)

121 consists of a central “rib-cage” β -sheet (red curve, β 1-10) sandwiched by connecting

122 loops, α -helices (α 1-4), and β -strands (β 7a, 7b, 8a, and 10a). LbT2 (light blue) consists

123 of two layers of β -sheets with a total of 9 β -strands (β 1’-9’). Spheres indicate residues

124 that interact with GPAA1 (red), PIGK (cyan), PIGS (purple), and PIGU (orange). **c** Side

125 (i) and normal (ii) view of PIGU. PIGU features a membrane core region (orange) with

126 short TMHs enclosed by a ring of TMHs (pale yellow). This arrangement creates a
127 membrane void (blue trapezoid) to hold the five amphipathic helices AH1-5 (yellow).
128 The TMHs and AHs are so arranged such that the C-terminal ends of several α -helices
129 (labeled with black text) expose to the surface (*ii*). The resulting dipole moments (δ^-)
130 and acidic residues make the surface electrostatically negative. **d** “Open-book”
131 representation of the electrostatic potential molecular surface generated using the
132 Adaptive Poisson-Boltzmann Solver module in PyMol.

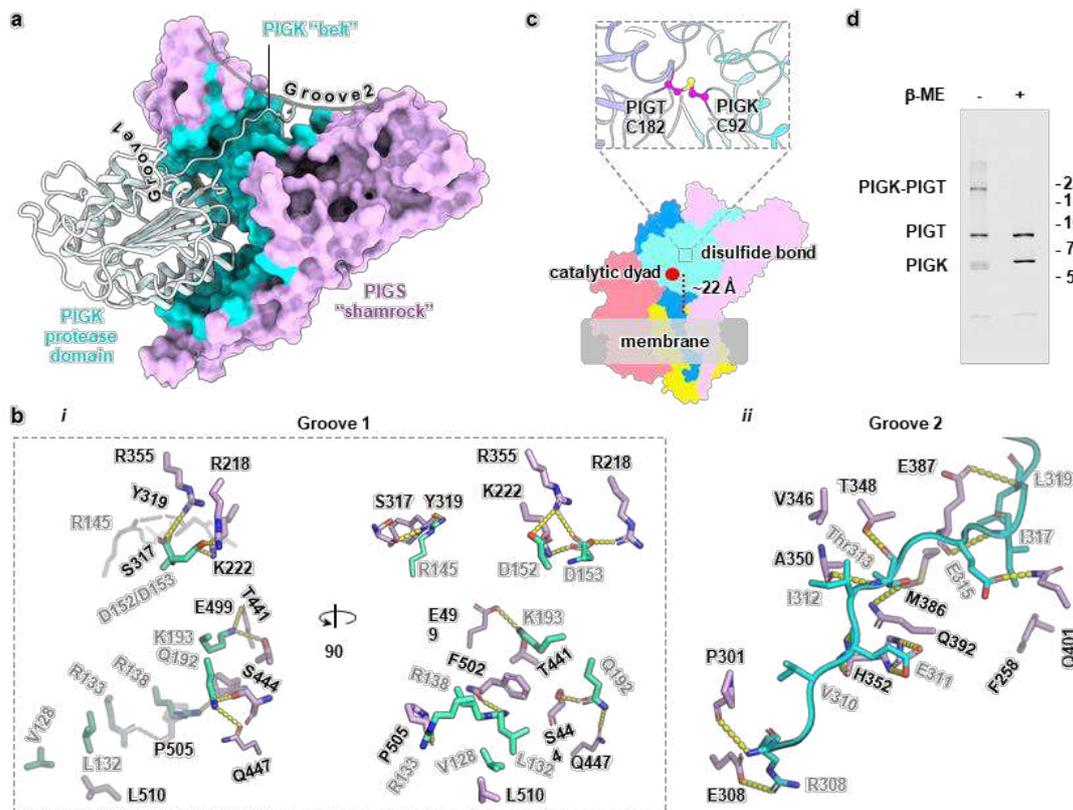
133

134 **A complex assembled for catalytic purposes**

135 The consensually recognized catalytic subunit PIGK^{17,18} is secured onto the
136 platform through multivalent interfaces. Its protease domain sits snugly in a three-sided
137 cavity formed by the PIGT/PIGU platform as well as GPAA1 and PIGS approaching
138 from opposite directions (**Fig. 1b**). In addition, PIGK exploits two grooves on the
139 shamrock-like PIGS for interactions, with the globular luminal domain complementing
140 the concave Groove1 surface and its “belt”-like loop entangling onto the Groove2 (**Fig.**
141 **3a**). This configuration buries a total of 2341.9 Å² surface area and holds PIGK tightly
142 via 18 hydrogen bonds, 6 salt bridges, one cation- π interaction, and several hydrophobic
143 interactions (**Fig. 3b**). Further, the C-terminal portion of PIGK travels down along a
144 shallow surface of PIGT’s backside and buckles its terminal helix into the membrane
145 to interact with TMH5/7 of PIGU (**Figs. 1b, 2a**). Finally, an inter-subunit disulfide bond
146 between PIGK C92 and PIGT C182, which has been reported previously²¹ and

147 visualized in this study, nails PIGK onto PIGT (**Figs. 3c, d**). These interactions place
 148 PIGK at a “mid-air” position (relative to the membrane “ground”) with its catalytic
 149 dyad (C206-H164) ¹⁷ measuring ~22 Å from the membrane (**Fig. 3c**). Because the
 150 distance approximates the hydrophilic “stem” of both GPI and proprotein substrates
 151 that are presumably rooted in the membrane (**Fig. 1a**), this topological arrangement
 152 may confer substrate specificity.

153



154

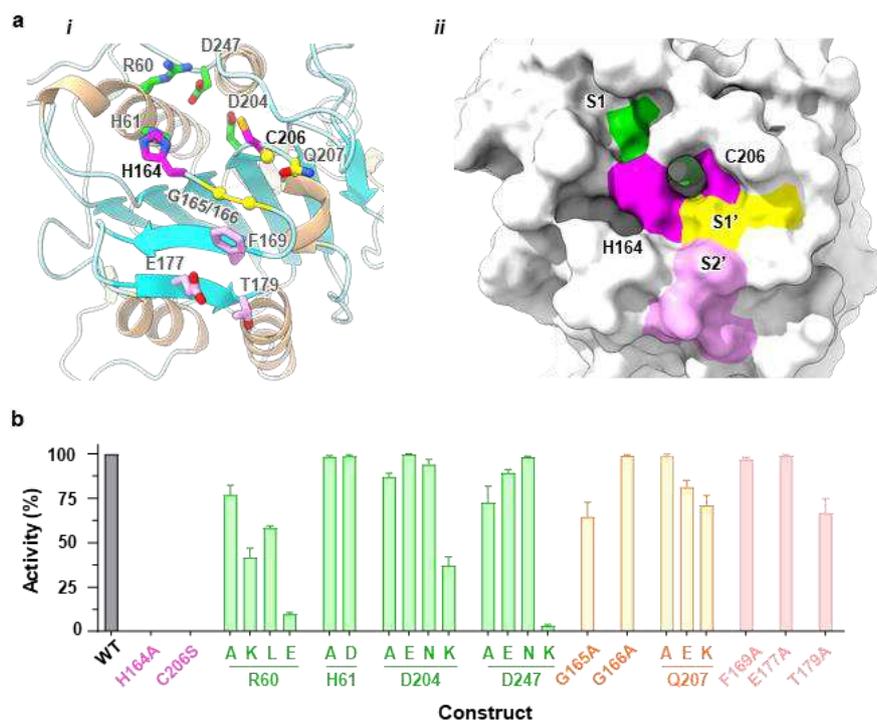
155 **Fig. 3. PIGS and PIGT place PIGK and its catalytic dyad in a position suitable for**
 156 **catalysis. a** PIGS (surface, pink) holds PIGK (ribbon) by embracing the protease
 157 domain using one of the shamrock grooves and hosting the loop region in another
 158 groove. The buried surface is colored cyan. **b** Detailed interactions between PIGK and

159 PIGS at Groove 1 (*i*) and Groove 2 (*ii*). PIGK residues (cyan) are indicated with grey
160 texts and PIGS residues (light purple) are labeled with black texts. Dashed lines (yellow)
161 indicate distances within 3.6 Å. **c** An inter-subunit disulfide bond (PIGT C182 / PIGK
162 C92) (magenta) nails PIGK (cyan) onto the PIGT (blue). This and the other interactions
163 fix PIGK in a position such that the catalytic dyad (red dot) is ~22 Å “above” the
164 membrane interface (grey shading). This geometry would suit interactions with GPI
165 which is expected to insert into the membrane by the phosphatidyl moiety with its
166 hydrophilic glycan chains (measuring ~25 Å in the extended form) stemming from the
167 membrane to meet with the catalytic dyad. **d** Verification of the PIGK-PIGT disulfide
168 bond. SDS-PAGE in-gel fluorescence shows a high molecular-weight band (PIGK-
169 PIGT) in the absence of, but not in the presence of, the reducing agent β-
170 mercaptoethanol (β-ME) at the cost of the individual PIGT/PIGK bands. Other subunits
171 were also co-transfected but were invisible owing to their lack of the TGP-tag ³⁷.
172 Molecular markers are labeled on the right.

173 **A legumain-like active site**

174 PIGK belongs to the C13 cysteine protease family including legumains and
 175 caspases^{17,18}. In the GPI-T structure, PIGK features a central 6-stranded β -sheet with
 176 3 α -helices on either side (**Figs. S4a, b**) and structurally resembles legumains (C α -
 177 RMSD of 2.3 Å) (**Fig. S4d**)³⁸. The catalytic dyad, the trivalent oxyanion hole, and the
 178 residues in the substrate-binding pockets S1/S1'/S2' are mostly conserved and
 179 superimposable (**Figs. 4a, S4d, e**).

180



181 **Fig. 4. Structural and functional resemblance of the PIGK active site to that of**
 182 **legumains.** **a** Cartoon (*i*) and surface (*ii*) representation of the PIGK protease domain
 183 with the catalytic dyad (magenta), S1 (green), S1' (yellow), and S2' (pink) residues
 184 shown as sticks (*i*) or highlighted in colors (*ii*). S1/S1'/S2' sites are superposed from

185 legumain structures^{39,40}. **b** Functional assay of the active site mutants. Apparent activity
186 (% of wild-type, WT) was measured by immune staining of a reporter GPI-AP (CD59)
187 on the surface of PIGK-KO cells transfected with the indicated mutants of the catalytic
188 dyad and S1/S1'/S2' residues. Cells expressing PIGK fused with a thermostable green
189 fluorescence protein (TGP)³⁷ were gated, and the surface staining of CD59 was
190 analyzed. Data represent mean \pm s.e.m. from three independent experiments
191 **(Supplementary Data 1)**.

192

193 To ascertain the functional importance of S1/S1'/S2' residues, we used a cell-based
194 GPI-AP reporter assay²⁰ to probe the apparent activity of the mutants. In the cell-based
195 assay, the surface display of the reporter CD59 was monitored by flow cytometry in
196 GPI-T knockout HEK293 cells upon ectopic expression of mutants. As a control,
197 mutations of the previously identified catalytic dyad¹⁷ (H164A or C206S) abolished
198 GPI-T activity (**Fig. 4b**).

199 Substituting R60 at the S1 site with glutamate almost diminished GPI-T activity
200 (9.8% relative to the wild-type (WT) and same hereafter), and R60A, R60L, or even
201 the conservative R60K mutation reduced the activity by 20%-60% (**Figs. 4b, S5a,**
202 **Supplementary Data 1**), suggesting that both the shape and the positive charge of R60
203 were important for the integrity of S1. The negative charge on D204 or D247 was not
204 essential (D204N, 94.2%; D247N, 98.2%) but a charge reversion was detrimental
205 (D204K, 37.3%; D247K, 3.2%). In contrast, H61 tolerated mutation to alanine or even

206 aspartate (H61A, 98.5%; H61D, 98.7%) (**Fig. 4b**), suggesting a less important role. The
207 mutagenesis results also indicate the functional importance of three residues in the
208 corresponding S1' (G165A, 64.5%; Q207E, 81.2%, Q207K, 71.2%) and S2' site
209 (T179A, 66.7%) (**Fig. 4b**)^{39,40}. Taken together, PIGK may use a similar mechanism as
210 legumains for peptide cleavage and even substrate recognition.

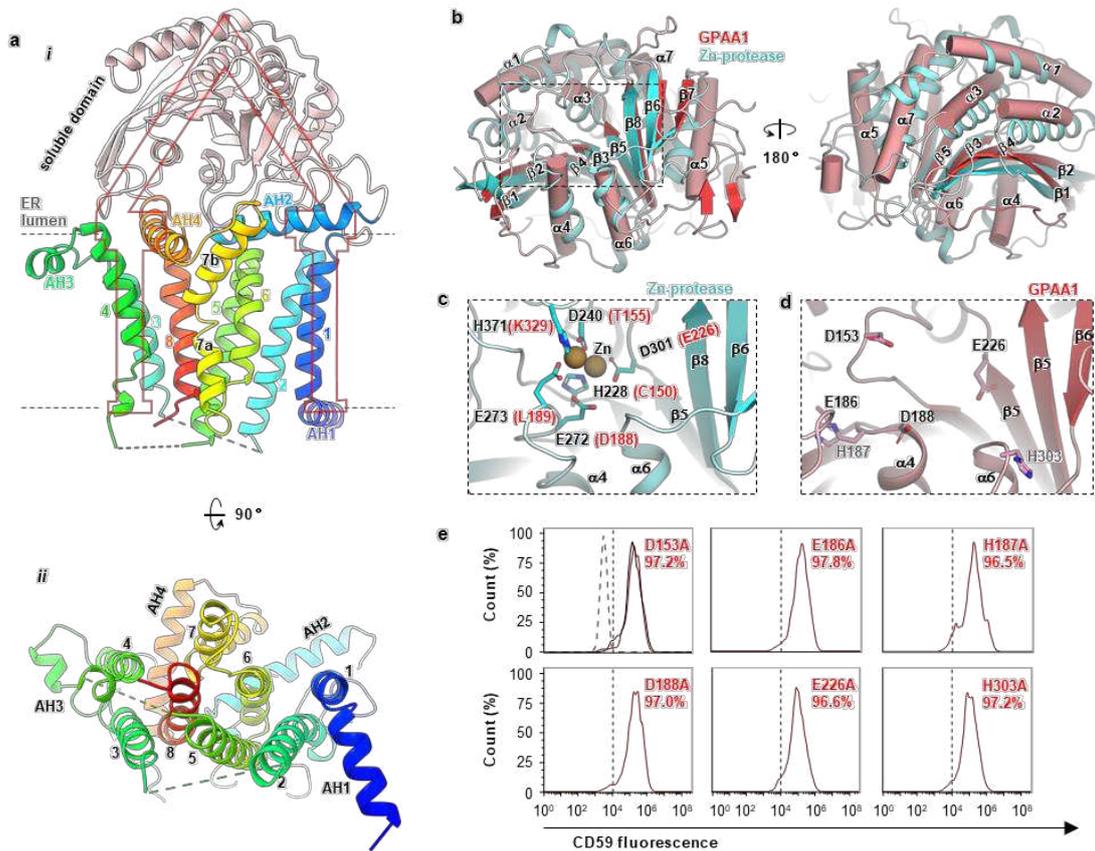
211

212 **GPI-T contains two protease-like domains**

213 Both GPAA1 and PIGS possess a protease-like domain in our GPI-T structure.
214 GPAA1 assumes a portico-shape with eight TMHs, four AHs, and a soluble domain
215 (**Fig. 5a**). Whilst the transmembrane domain shares no recognizable structural
216 homology with known folds, the soluble domain is similarly arranged as the Zn²⁺-
217 protease AM-1⁴¹ (PDB ID 2EK8, C α RMSD of 3.2 Å) (**Fig. 5b**). However, a closer
218 inspection shows the corresponding GPAA1 domain lacks a characteristic catalytic
219 zinc-binding motif that consists of glutamates, aspartates, and histidines (**Fig. 5c**), and
220 mutations of possible Zn-binding motif candidates in the vicinity (D153A, E186A,
221 H187A, D188A, E226A, H303A) (**Fig. 5d**) had no noticeable effect on GPI-T activity
222 in cells (**Fig. 5e**). Likewise, a portion of PIGS (resi. 222-398) folds similarly to a
223 metzincin-type protease AmzA (C α RMSD of 3.6 Å)⁴² (**Fig. S6a**). However, unlike
224 AmzA, PIGS lacks elements for a metzincin active site that contains a conserved
225 methionine near a zinc-binding tri-histidine motif. Further, a cysteine-based zinc-finger
226 in AmzA was compositionally impossible in PIGS (**Figs. S6b, S6c**). Thus, we envisage

227 that the protease-like GPAA1 and PIGS subunits may help recruit protein substrates
 228 instead of performing cleavage, reminiscent of the γ -secretase⁴³.

229



230 **Fig. 5. Structural and functional characterization of GPAA1 reveal a protease-like**

231 **domain.** **a** Side (*i*) and normal (*ii*) view of GPAA1. Numbers indicate TMHs and AH1-

232 3 label the three amphipathic helices. **b** The soluble domain of GPAA1 (red/pink,

233 cylinder) is structurally similar to a Zn-protease AM-1 (cyan, cartoon, PDB ID 2EK8)

234 with a Z-score of 20.6 and C α RMSD of 3.2 Å (from a DALI search)⁴¹. GPAA1 is

235 shown as red/pink cylinder representations and the Zn-protease is shown as cyan

236 cartoon representations. **c** The Zn-binding site of AM-1 (expanded view of the boxed

237 region in **b**) consists of two each of aspartate, glutamate, and histidine residues that are

238 not fully conserved in GPAA1 (in brackets). **d** Arrangement of GPAA1
239 aspartate/glutamate/histidine (D/E/H) residues in the region corresponding to the Zn-
240 binding site in AM-1. Despite having the same composition, these residues are unlikely
241 to form a Zn-binding site because of different spatial arrangements. **e** Mutation of the
242 D/E/H residues in **d** did not reduce CD59-staining in the flow cytometry assay. The
243 function of wild-type GPAA1 and mutants were assessed using FACS analysis which
244 monitors the surface expression of the GPI-AP reporter (CD59) in GPAA1 KO cells
245 transfected with appropriate plasmids. Cells were gated by TGP fluorescence ³⁷ for
246 GPAA1 expression and analyzed for CD59 staining. The dotted line (grey) indicates the
247 CD59-staining background level from cells expressing an unrelated TGP-tagged
248 membrane protein (negative control). Solid lines indicate staining of cells transfected
249 with the wild-type (black) or mutant genes. A vertical dash line marks threshold (CD59-
250 gating) determined from the negative control.

251

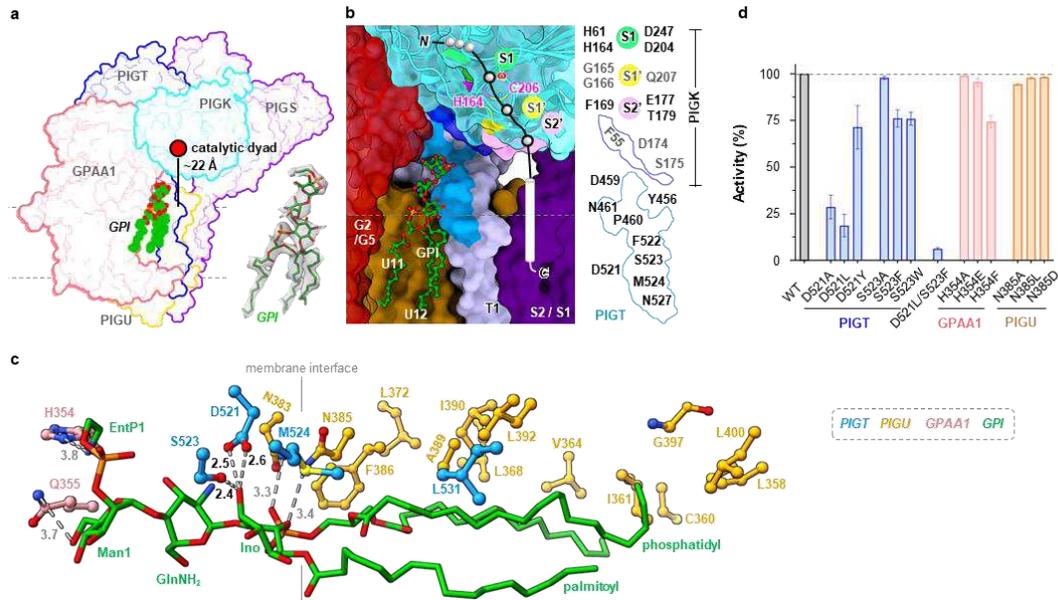
252 **Identification of the GPI-binding site**

253 The lipid substrate GPI is, as aforementioned, expected to be anchored in the
254 membrane via its phosphatidyl group within ~ 25 Å of the catalytic dyad based on the
255 approximate length of the glycans between the reactive EtNP3 and the acyl chains.
256 Intriguingly, densities that fit an almost complete GPI core (palmitoylated
257 phosphatidylinositol, glucosylamine, and EtNP-modified mannose1) (**Fig. 6a**) are
258 observed in a cavity jointly formed by TMH11/12 of PIGU, TMH2 of GPAA1, and the

259 sole TMH of PIGT (**Fig. 6b**) directly “underneath” the catalytic dyad. Although the
260 precise orientation of the EtNP1 moiety is ambiguous at the current resolution, the
261 density for the characteristic tri-acyl chain, the inositol-glucosylamine-mannose glycan
262 core is well-resolved. In the model, the acyl chains of GPI glue to the transmembrane
263 region mostly of PIGU, and the head group forms three strong hydrogen bonds with
264 PIGT D521/S523 and several weak ones with PIGU N383/N385, and GPAA1
265 H354/Q355 (**Fig. 6c**). This composite cavity is additionally filled with a density that
266 fits nicely for a digitonin molecule which also contains a polysaccharide chain like GPI
267 (**Fig. S7a**). Further, several evolutionarily conserved residues (PIGK F55/D174/S175,
268 PIGT Y456/D459/P460/N461/D521/F522/S523/M524/N527) (**Figs. S7c, d**) delineate
269 a surface-exposed patch stretching from the membrane interface to the catalytic dyad
270 (**Fig. 6b**). Notably, this motif is rich in acidic/aromatic residues that are capable of
271 forming H-bonds/“greasy slides” frequently found in protein-sugar interactions ⁴⁴.
272 Making topological sense, the cavity is oriented in a way it would fit a proprotein
273 substrate with the hydrophobic portion of CSP in the membrane while the hydrophilic
274 portion would span the ~22 Å space to reach the active site (**Fig. 6b**).

275

276



277

278 **Fig. 6. Characterization of a composite GPI-binding site.** **a** Cryo-EM density that
 279 fills a nearly complete GPI was observed in the membrane cavity “underneath” the
 280 catalytic dyad. **b** Expanded view of the composite site encompassed by the indicated
 281 TMHs from GPAA1/PIGU/PIGT/PIGS. Evolutionarily conserved residues (**Figs. S7c,**
 282 **d**) in the vicinity are colored marine (PIGT) and blue (PIGK). Subunits are shown as
 283 surfaces except that PIGK was additionally shown as ribbon representations with the
 284 catalytic dyad and S1/S1’/S2’ residues highlighted in indicated colors. **c** Interaction
 285 between the partial GPI (green) and GPI-T (colored-coded as indicated). Distances (Å)
 286 are either indicated by numbers (black/grey for strong/weak H-bonds) for H-bonding
 287 interactions or omitted for hydrophobic interactions (within 5 Å of GPI). A vertical line
 288 marks the membrane boundary. **d** Apparent activity of GPI-T mutants relative to the
 289 wild-type (WT). GPI-T KO cells were gated by TGP-fluorescence³⁷ for subunit
 290 expression and analyzed for surface staining of the reporter GPI-AP (CD59) by flow
 291 cytometry. Data represent mean ± s.e.m. from three independent experiments
 292 (**Supplementary Data 1**).

293 To further characterize the cavity, we generated 30 mutants (**Supplementary Data**
294 **1**) targeting 10 residues in the conserved patch and the hydrogen-bonding network
295 involved in GPI-binding. Three residues, two from PIGT and one from GPAA1,
296 responded appreciably to mutagenesis. D521 formed two strong H-bonds with GPI (**Fig.**
297 **6c**). Consistently, substituting it with an alanine decreased the apparent activity to
298 28.9%, and introducing intended steric clashes by a leucine reduced the activity to
299 18.8%. The H-bond by PIGT S523 was seemingly non-essential (S523A, 98.4%), but
300 mutating to bulky residues reduced apparent activity (S523F, 76.5%; S523W, 76.8%)
301 (**Fig. 6d**) presumably by steric hindrance. Further, a double mutant (D521L/S523F)
302 almost abolished activity (6.2%, **Fig. 6d**). GPAA1 H354 formed a weak H-bond with
303 GPI. Although the alanine mutation did not change the apparent activity, substitution
304 with phenylalanine which has a similar volume with histidine reduced activity by ~25%
305 (**Fig. 6d**). The lack of substantial change on GPI-T activity for individual mutants of
306 other residues (**Fig. S7b**) might be explained by the abundance of weak multivalent
307 interactions with GPI.

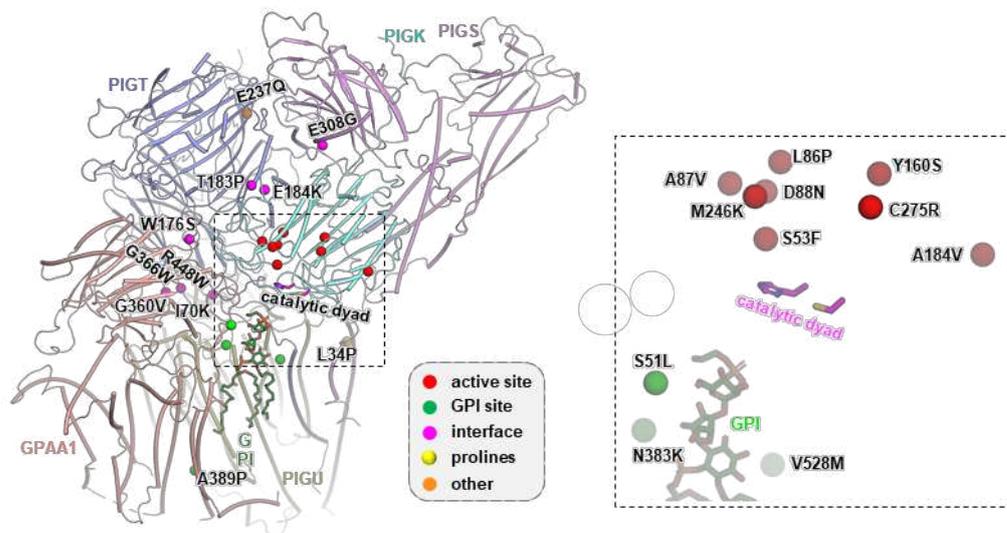
308

309 **Structural interpretation of pathogenic mutations**

310 Our structure provides a framework to propose possible mechanisms for genetic
311 neurodevelopmental disorders such as NEDHCAS associated with GPI-T mutations
312 and reduced surface expression of various GPI-APs²⁴⁻³³. Mapping these mutations onto
313 GPI-T showed a pattern with over half of the mutations (12 out of 22, PIGK

314 S53F/L86P/A87V/D88N/Y106S/A184V/M246K /C275R, PIGT V528M, GPAA1
 315 S51L/A389P, and PIGU N383K) clustering near the catalytic dyad and the GPI-binding
 316 site (**Figs. 7, S8a**). Among the rest of the ten mutations, eight are located at the inter-
 317 subunit interfaces (PIGT T183P/E184K/G360V /G366W/R448W, PIGS E308G,
 318 GPAA1 W176S, and PIGU I70K) (**Figs. 7, S8b**) and hence may disrupt the integrity of
 319 the complex. PIGS L34P occurs in the first transmembrane helix and may cause folding
 320 issues by main chain distortion. PIGK Y160S and GPAA1 W176S may also promote
 321 misfolding by introducing the small hydrophilic serine residue into a hydrophobic
 322 microenvironment (**Figs. S8c, d**).

323



324

325 **Fig. 7. Distribution of genetic mutations on GPI-T.** Mapping the disease mutations
 326 (C α spheres, color-coded by the categories in the grey box) onto the human GPI-T
 327 structure (cylinder representation). The mutations include PIGK S53F/L86P/A87V/
 328 D88N/Y160S/A184V/M246K/C275R that cause a neurodevelopmental syndrome with
 329 hypotonia, cerebellar atrophy, and epilepsy ²⁴, PIGT T183P that causes an intellectual
 330 disability syndrome ²⁷, PIGT E184K/G360V/R448W that are related to the Multiple

331 Congenital Anomalies-Hypotonia Seizures Syndrome 3 ^{28,31,32}, PIGT E237Q/V528M
332 that cause developmental disorders characterizing learning disability, epilepsy,
333 microcephaly, congenital malformations and mild dysmorphic features ²⁹, PIGT
334 G366W that is found in patients with epileptic apnea and multiple congenital
335 anomalies, severe intellectual disability, and seizures ³⁰, PIGS L34P/E308G that are
336 related to a neurological syndrome ranging from fetal akinesia to epileptic
337 encephalopathy ²⁶, PIGU I70K/N383K that are found in patients with severe intellectual
338 disability, epilepsy, and brain anomalies ²⁵, and GPAA1 S51L/W176S/A389P that are
339 also related to developmental disorders featuring global developmental delay, epilepsy,
340 cerebellar atrophy, and osteopenia ³³.

341

342 **Discussion**

343 The yeast GPI-T has been reported to exist in dimer based on native PAGE results
344 ⁴⁵. Although we have not observed dimer particles for the human GPI-T in this study,
345 we note the report of an increasing number of detergent-specific weak oligomers of
346 membrane proteins in the literature ⁴⁶⁻⁴⁹. Future investigations using less disruptive
347 detergents such as steroid glycosides for extraction may be considered. Such detergent
348 screening experiments, and in fact enzyme kinetics studies, would greatly benefit from
349 a robust test-tube biochemical assay with authentic substrates which is yet to be
350 developed.

351 Although the substrate-binding and catalytic mechanisms are yet to be revealed by
352 structural studies of GPI-T with the peptide substrate, the structural comparison to

353 legumains offers mechanistic implications. In legumains, the trivalent oxyanion pocket
354 ⁴⁰ (backbone amide nitrogen of C189, G149, and N δ 1 of the catalytic residue H148) is
355 proposed to polarize the carbonyl oxygen of the P1 residue in the substrate. This
356 increases the electrophilicity of the carbonyl carbon, allowing the catalytic C189-S γ to
357 deprotonate and to launch the nucleophilic attack that proceeds to peptide cleavage. In
358 PIGK, the corresponding PIGK residues are the similarly positioned C206, G165, and
359 H164 (**Fig. S4d, e**), indicating an analogy in the carbonyl activation process. In
360 legumains and caspases, the selectivity of the substrate P1 residue is conferred mainly
361 by the S1 pocket, and the characteristic zwitterionic S1 site (two positively charged
362 residues and two negatively charged residues) of legumains is believed to be suited for
363 its asparagine specificity. In line with this, legumains are unable to process substrates
364 with an aspartate unless low-pH conditions are used to protonate the P1 residue ⁴⁰.
365 Interestingly, despite being capable of processing substrates with various ω -residues
366 (analogous to P1) including aspartate, the corresponding site in PIGK has virtually the
367 same composition as the S1 of legumains with the exception of a conservative change
368 (PIGK D204 versus legumain E187) (**Figs. 4a, S4e**). Whether and how this site confers
369 selectivity to proprotein substrates remains to be investigated structurally and
370 biochemically.

371 In addition to the protease activity, a ligation reaction to conjugate the GPI-peptide
372 amide, presumably also by PIGK, is required for GPI anchoring. In legumains ⁵⁰, the
373 ligase activity involves a recently discovered mechanism where a succinimide residue
374 converted from D147 provides energy. Despite its overall similarity with legumains,

375 PIGK contains a glycine (G163) as a counterpart to the legumain D147 (**Fig. S4a**), thus
376 making the succinimide-mediated ligation less likely in PIGK. Instead, PIGK may use
377 a ping-pong mechanism similar to that of the sortase transamidation ⁵¹: the catalytic
378 dyad acts on the pro-protein, forming a semi-stable thioacyl intermediate that links the
379 ω residue and C206. The incoming amine group on EtNP then attacks this intermediate,
380 producing GPI-AP and freeing PIGK.

381 GPAA1 has been proposed previously to be a catalytic subunit owing to its similar
382 arrangement of secondary structure elements to that of zinc proteases ¹⁹. We verify the
383 structural similarity but clarify that GPAA1 lacks the Zn-binding motif and hence is
384 unlikely a catalytic subunit (**Figs. 5b-e**). Furthermore, the region corresponding to the
385 AM-1 Zn-binding site is remote (~50 Å) from the catalytic dyad in our structure (**Fig.**
386 **S9**).

387 Both GPAA1 and PIGU have been proposed to be involved in GPI-binding. The
388 functional part was narrowed down to the last TMH for GPAA1 ²⁰ which is wrapped by
389 the rest of GPAA1 TMHs (**Fig. 5a**). Because this site is ~20 Å away from the composite
390 GPI-binding site (**Fig. S9**), how GPAA1 TMH8 is involved in GPI-binding is currently
391 unclear. Similarly, residues F274/W275, located in the AH5 of PIGU (**Fig. S2c**), are
392 proposed to be crucial for GPI-binding ¹⁶. Although the functional importance of the
393 two residues is verified in this study (**Supplementary Data 1**), they are also distal (~10
394 Å) to the composite GPI-binding site (**Fig. S9**). It is possible that mutations in these
395 regions compromise GPI-binding by allosteric effects.

396 The density for a partial GPI molecule in the structure provides valuable
397 information for an overview of the active site, but the detailed mechanism, for example,
398 the requirement of the EtNP3 for GPI-T reaction, remains unknown because of the lack
399 of reliable density beyond Man1. In addition to improving resolution, designing
400 mutants to covalently trap GPI or GPI-APs to the active site should help reveal the
401 complete picture of substrate-binding.

402 Finally, it is interesting to note the implications of the glycan-containing digitonin
403 molecule proximal to the composite GPI-binding site. Because lipids mostly diffuse
404 two-dimensionally in the membrane plane, GPI-T may use the cavity as a “hamster”
405 pocket to stuff GPI. By doing so, the relatively low diffusion rate of lipids may be
406 compensated by the locally high concentration of substrates in a reaction.

407

408 **Conclusions**

409 Our study on the human GPI-T complex provides structural evidence of an
410 assembly that suits its catalytic purpose. The subunits PIGU/T/S and GPAA1 jointly
411 hold PIGK in a way the space between the catalytic dyad in PIGK and the membrane
412 interface presumably may act as a molecular ruler to restrict proprotein and GPI
413 substrates. Combined with rational mutagenesis, the structural reveals a composite GPI-
414 binding site where a juxtamembrane portion of PIGT is unexpectedly found to
415 contribute most for the interactions. Consistent with the structural interpretations, most
416 disease mutations occur near the catalytic site, the GPI-binding site, and at the subunit

417 interfaces. Our work also formulates feasible working models for the GPI anchoring
418 process and provides a high-resolution framework to inspire the design of biochemical
419 and biophysical experiments to probe catalytic mechanisms in detail.

420

421 **ACKNOWLEDGEMENTS:**

422 Cryo-EM data were collected at SKLB West China Cryo-EM Center and
423 processed on Duyu High Performance Computing Center in Sichuan University. We
424 thank the staff members of the Center of Cryo-EM of Fudan University and the Cryo-
425 EM Center at National Facility for Protein Science in Shanghai for technical support
426 and assistance. We thank Drs. Feilong Meng, Liu Daisy Liu, and Hai Jiang and Mr. Yin
427 Chen for guidance on CRISPR-Cas9. This work has been supported by the National
428 Natural Science Foundation of China (82151215, 31870726, D.L; 32171194, Q.Q.;
429 82041016, 32070049, Z.S.), the Strategic Priority Research Program of CAS
430 (XDB37020204), CAS Facility-based Open Research Program (2017, D.L.), Science
431 and Technology Commission of Shanghai Municipality (20ZR1466700, D.L.),
432 Ministry of Science and Technology of China (2021YFA1301900, Z.S.), and the start-
433 up funds from Shanghai Stomatological Hospital & School of Stomatology, Fudan
434 University (Q.Q.).

435

436 **AUTHOR CONTRIBUTIONS:**

437 Y.X. established purification protocol. T.L. and Y.X. purified the complex. Y.X.

438 constructed KO cells and performed functional assays. Y.L., Y.C., and Z.Z. prepared
439 and screened cryo-EM grids. G.J. collected cryo-EM data under the supervision of Z.S..
440 Z.Z. and G.J. processed cryo-EM data and produced the final map. J.B. helped with
441 molecular cloning. D.L. initiated the project. D.L. and Q.Q. wrote the manuscript with
442 input from Y.L., T.L., Y.X., Z.Z., and Z.S..

443

444 **COMPETING INTERESTS:**

445 The authors declare no competing interests.

446

447

448 **Methods**

449 **Molecular cloning**

450 The genes encoding the human GPI-T subunits GPAA1 (Genbank ID
451 NP_003792.1), PIGK (NP_005473.1), PIGS (NP_149975.1), and PIGT (NP_057021.2)
452 were PCR-amplified from cDNA clones provided by the authors' institute. The gene
453 encoding the human PIGU (NP_536724.1) was PCR amplified using overlapping
454 oligonucleotides. The PCR products were Gibson assembled⁵² into various versions of
455 the pBTSG³⁷ vector which were modified to carry the following tag sequences at the
456 3'-end of the encoding sequence of the thermostable fluorescence protein (TGP)³⁷ tag:
457 GPAA1, 2×Flag; PIGK, Strep; PIGS, Myc; PIGT, 9×His; PIGU, hemagglutinin (HA).
458 For disulfide cross-linking between PIGT and PIGK, the DNA fragment encoding TGP

459 was removed in the constructs of GPAA1/PIGS/PIGU to avoid background in in-gel
460 fluorescence. The constructs were verified by DNA sequencing.

461 Mutations were made using standard PCR-based site-directed mutagenesis. DNA
462 sequences were verified by Sanger sequencing.

463 The TGP-tags are located in the cytosol side of the GPI-T, opposite to its large
464 luminal domain. Flow cytometry (below) showed that the TGP-tagged GPI-T was
465 functional because the co-transfection of all five TGP-tagged subunits in a cell line with
466 all GPI-T subunits disrupted (see below) restored cell surface expression of the GPI-AP
467 reporter CD59.

468

469 **Generation of GPI-T knock-out (KO) cell lines**

470 To generate cells lines with defective individual subunits, the endogenous genes
471 encoding the five GPI-T subunits were separately disrupted by CRISPR-Cas9 editing
472 using two or three pairs of sgRNA oligos that are designed using the online server
473 (<http://cistrome.org/SSC/>)⁵³ with the following sequences (forward / reverse pairs):

474 5'- CACCGTGTGGGGCTGCTGCTGGCAC-3' / 5'- AAACGTGCCAGCAGCAG
475 CCCCACAC-3' and 5'- CACCGCAGGAGCAAGAAGCCGACAG-3' / 5'- AAAC
476 CTGTCGGCTTCTTGCTCCTGC-3' for GPAA1; 5'- CACCGAATTACCAACATA
477 GAACTCG -3' / 5'- AAACCGAGTTCTATGTTGGTAATTC -3', 5'- CACCGTTC
478 ATATTAGTTTGGCTAGC -3' / 5'- AAACGCTAGCCAACTAATATGAAC -3' and
479 5'- CACCGGCTCTAGCTAGTAGTCAAGT -3' / 5'- AAACACTTGACTACTAGCT

480 AGAGCC -3' for PIGK; 5'- CACCGTGAGCCTCAGGAACAAGCGG -3' / 5'- AAA
481 CCCGCTTGTTTCCTGAGGCTCAC -3' and 5'- CACCGAGTGGAGCGCTGAGAA
482 GAGG-3' / 5'- AAACCCTCTTCTCAGCGCTCCACTC-3' for PIGS; 5'-CACCGC
483 GGTGCAGACCACCTCCCG -3' / 5'- AAACCGGGAGGTGGTCTGCACCGC-3'
484 and 5'- CACCGCACCATCACCTCCAAGGGCA-3' / 5'- AAAGTCCCTTGGG
485 GGTGATGGTGC-3' for PIGT; 5'- CACCGAGTCCTGGATTGCAAATAC-3' / 5'-
486 AAACGTATTTTGCAATCCAGGACTC -3', 5'- CACCGCCTAATTGACTATGCT
487 GAAT-3' / 5'- AACATTCAGCATAGTCAATTAGGC -3' and 5'- CACCGTCTTT
488 GGGTAGTCAAAGTGA-3' / 5'- AAAGTCACTTTGACTACCCAAAGAC-3' for
489 FIGU.

490 The oligos were designed to have sticky ends that are compatible with the Type IIs
491 restriction enzyme *BbsI* (Cat. R3539S, NEB) after annealing. The sgRNA oligo pairs,
492 dissolved in a buffer containing 0.2 M NaCl, 0.1 mM EDTA, and 10 mM Tris HCl pH
493 7.5, were mixed at equimolar concentrations of 10 μ M in a PCR tube. The
494 oligonucleotides were first denatured by heating at 95 $^{\circ}$ C for 3 min, before being
495 subjected to an annealing step with gradual cooling from 94 $^{\circ}$ C to 25 $^{\circ}$ C at 1- $^{\circ}$ C
496 gradients and an 11-s incubation under each temperature. One microliter of the annealed
497 mix was ligated into the vector pX330 (50 ng) pre-digested with *BbsI* using 5 units of
498 T4 ligase (Cat. EL0011, Thermo Fisher Scientific) in a 10- μ L reaction system at room
499 temperature (RT, 20-22 $^{\circ}$ C). The ligation products were transformed into DH5 α and the
500 resulting colonies were sent for DNA sequencing to identify desired constructs.

501 For CRISPR-Cas9 gene editing, 8 μ g of sgRNA-encoding plasmids constructed

502 above, 0.16 μg of pMaxGFP (as a FACS marker later), 16 μL of P3000 (Cat.
503 L3000008, Thermo Fisher Scientific), and 250 μL of Opti-MEM medium (Cat.
504 31985070, Thermo Fisher Scientific) were mixed and added to a separately prepared
505 mix containing 16 μL of Lipofectamine 3000 and 250 μL of Opti-MEM medium.
506 After incubation at RT for 15 min, the mixture was added dropwise to a 6-cm dish
507 containing HEK293 cells with 70%-90% confluency. Cells were cultured in a 5% CO_2
508 incubator at 37 $^\circ\text{C}$ in a Dulbecco's Modified Eagle Medium (DMEM) supplemented
509 with 10% fetal bovine serum (FBS). After 24 h, cells were washed with 2 mL of PBS
510 and digested with 0.5 mL of 0.1% trypsin (Cat. 25200056, Thermo Fisher Scientific)
511 for 3 min at 37 $^\circ\text{C}$ before being re-suspended in 3 mL of DMEM and 10 % FBS to
512 saturate trypsin. Cells were then harvested by centrifugation at RT at 300 g for 5 min,
513 washed with 10 mL of PBS, and resuspended in 0.5 mL PBS for fluorescence assisted
514 cell sorting (FACS). A total of ~800,000 cells were sorted in a BD FACSAria Fusion
515 machine by green fluorescence protein (GFP) and the top 5% (~40,000) were
516 collected. These cells were serially diluted using DMEM supplemented with 10%
517 FBS and seeded into 96-well plates such that an average of 2 or 4 cells were contained
518 in each well with 100 μL of medium. Single cells were allowed to populate for 10-12
519 d in a stationary CO_2 incubator. Wells containing a single colony were selected under
520 a microscope, and the selected cells were washed with 30 μL PBS before being treated
521 with 30 μL of 0.1% trypsin for 2 min at 37 $^\circ\text{C}$. After this, cells were resuspended in
522 200 μL of DMEM and 10% FBS and divided into two parts (160 μL and 70 μL) for
523 further culturing and PCR-identification, respectively. The alternation of genome was

524 screened using the following primer pairs: GPAA1, 5'- AGGACTCCGGGTTTAGGT
525 CT -3' and 5'- GTAGCCCAATCAAGGACCCC -3'; PIGK, 5'- TAAGCGATCTGC
526 CCTACCAC -3' and 5'- CCCACAGGGAAGAATTC GGG -3'; PIGS, 5'- GGCGAA
527 ATGGGTGTCATGTG-3' and 5'- GGCATGCAGATT TCCCTCCT-3'; PIGT, 5'-
528 GACTGTGCTTAAGGAGGGCA -3' and 5'- AGCCTAA CGTTGCCAAACCC -3';
529 PIGU, 5'- GCACAAAATGGTCCGGCAG-3' and 5'- AGGCCCATTAAGGCCAAG
530 TT-3'. Cells pellets were resuspended into a mix containing 10 μ L of 0.1 mM EDTA,
531 1 % Tween-20, 1 \times Taq polymerase buffer (50 mM KCl, 2 mM MgCl₂, 20 mM Tris
532 HCl pH 8.4), and 1 mg mL⁻¹ protease K. The mixture was heated successively at 56
533 °C for 2 h and 95 °C for 30 min in a thermocycler, before PCR amplification for 40
534 cycles. Compared to the wild-type cells, PCR products from knockout (KO) cells
535 lacked a large band (GPAA1, 2 kb; PIGS, 2.7 kb; PIGK, 3.7 kb; PIGT, 3.4 kb; PIGU,
536 2.3 kb) but showed a smaller band (GPAA1, 1 kb; PIGS, 0.7 kb; PIGK/PIGT/PIGU,
537 1.2 kb) owing to the deletion. The genomic deletions were further verified by DNA
538 sequencing of the PCR products using the abovementioned primers.

539 To confirm the lack of GPI-AP surface expression in the individual KO cells, cells
540 identified as positive above were allowed to grow for two more passages before being
541 analyzed by flow cytometry (see the section below). Cells that 1) lacked CD59 on cell
542 surface, and 2) the surface expression of CD59 could be restored by ectopic expression
543 of the corresponding subunits were either used for further flow cytometry assays or
544 cryo-preserved in 10% DMSO, 45% FBS in DMEM for long storage.

545 To generate cells with all five GPI-T subunits disrupted, the procedure above was

546 performed in tandem. Defective GPI-T in this cell line was confirmed by DNA
547 sequencing as above, and by the lack of CD59 staining upon transfection of
548 combination of any four, but not all five, subunits.

549

550 **Flow cytometry**

551 Wild-type or GPI-T KO HEK293 cells were cultured in Dulbecco's Modified
552 Eagle Medium (DMEM) supplemented with 10% FBS at 37 °C in 24-well plates inside
553 a CO₂ stationary incubator. Plasmids (0.5 µg of single plasmids, or a total of 0.5 µg of
554 multiple plasmids) were mixed with 1 µL of P3000, 25 µL of Opti-MEM medium, and
555 added to a separate mix containing 1.5 µL Lipofectamine 3000 and 25 µL Opti-MEM
556 medium. Transfection was carried out as outlined above. Two days after transfection,
557 cells were washed with PBS, treated with trypsin as in the previous section, and washed
558 and resuspended in 0.5 mL PBS. Phycoerythrin (PE)-labeled CD59 antibody (12-0596-
559 42, Thermo Fisher Scientific, 1 : 500 dilution) was incubated with the cells for 15 min
560 in dark. Cells were washed with PBS and resuspended in ~0.3 mL of PBS for flow
561 cytometry (Beckman CytoFlex LX) monitored at two wavelength pairs (488/525 for
562 GFP, 561/585 for PE). Cells (typically 40,000) were gated using the GFP channel (from
563 expression of TGP-tagged GPI-T subunit(s)) and analyzed for positive signal for the
564 PE channel (for surface staining of CD59) using the software FlowJo (BD Life
565 Sciences).

566 For the apparent activity of GPI-T mutants, the percentage of the immune staining

567 of CD59 in singly KO-cells transfected with the mutant was normalized to the negative
568 control (the same cell line transfected with the TGP-tagged Patched, an unrelated
569 membrane protein) and the positive control (the same cell line transfected with plasmids
570 carrying the wild-type subunit gene). The expression and integrity of all mutants (TGP-
571 tagged) was also separately confirmed by SDS-PAGE in-gel fluorescence ³⁷. Data
572 reported in this work were from three independent experiments (**Supplementary Data**
573 **1**).

574

575 **GPI-T expression and purification**

576 GPI-T was expressed in Expi293 cells co-transfected with five plasmids carrying
577 all subunits using polyethylenimine (PEI). Briefly, 0.5 L of cells (>95% viability) at a
578 density of $2 \times 10^6 \text{ mL}^{-1}$ were diluted 2 times and cultured at 37 °C in a 3L flask in a CO₂
579 incubator to re-reach $2 \times 10^6 \text{ mL}^{-1}$ (typically takes one day). Six milligrams of plasmids
580 and 12 mg of PEI were mixed in 100 mL of medium for 20 min at RT before being
581 added into 1 L of cell culture. Sodium valproate (Cat. P4543, Sigma) was added to a
582 final concentration of 2 mM. Cells were harvested after 48 h by centrifugation at 1,500
583 g for 15 min, washed once with PBS buffer, flash-frozen in liquid nitrogen, and stored
584 at -80 °C until use.

585 GPI-T was purified by tandem affinity chromatography followed by gel filtration.
586 All steps were performed at 4 °C. Cells from 4 L of culture were re-suspended with
587 solubilization buffer containing 1 % LMNG/0.1 % CHS, 150 mM NaCl, 1 mM PMSF,
588 1 × protein inhibitor cocktail (Cat. B14001, Bimake), and 50 mM Tris-HCl pH 8.0 and

589 stirred gently for 2 h. Cell debris were removed by centrifuging at 4,4200 g for 1 h. The
590 supernatant was collected, mixed with 4.5 mL pre-equilibrated Strep Tactin beads (Cat.
591 SA053100, Smart-lifesciences) and agitated gently for 2 h. The beads were packed into
592 a gravity column (Cat. 7321010, Bio-Rad) and washed with 5 column volume (CV) of
593 0.01 % LMNG, 0.001 % CHS and 0.1 % Digitonin before being incubated with 0.2 %
594 digitonin in Buffer A (150 mM NaCl, 20 mM Tris HCl pH 8.0) for 1 h. The beads were
595 then washed with 1.5 CV of 0.2 % digitonin and 5 CV of 0.1 % digitonin in Buffer A
596 before being eluted with 5 mM d-desthiobiotin and 0.1 % digitonin in Buffer A. The
597 pooled fractions were adjusted to contain 10 mM imidazole before incubated with 1 mL
598 of Ni-NTA beads (Cat. 1018401, Qiagen) for 1 h with mild agitation. The beads were
599 packed into a gravity column, washed with 5 CV of 10 mM imidazole, before being
600 eluted using 0.3 M imidazole in 0.1 % digitonin in Buffer A. The Pooled fractions were
601 concentrated with a 100-kDa cut-off concentrator (Cat. UFC810096, Merck millipore)
602 and further purified with gel-filtration on a Superose 6 10/300 GL column (Cat. 29-
603 0915-96, Cytiva) in a running buffer containing 0.1 % digitonin in Buffer A. Peak
604 fractions were pooled and concentrated to 14-20 mg mL⁻¹ for cryo-EM grid preparation.
605 Protein concentration was determined by measuring A₂₈₀ using a theoretical molar
606 extinction coefficient of 596,940 M⁻¹ cm⁻¹ assuming an equimolar stoichiometry.

607

608 **Cryo-EM sample preparation and data collection**

609 Purified GPI-T complex (2.5 µL) was applied onto glow-discharged Quantifoil Au
610 R1.2/1.3 (300 mesh) grids, and blotted with filter paper for 3 s with a blotting force of

611 5 at 4 °C, with 100% humidity in a Vitrobot Mark IV (FEI) chamber before being cryo-
612 cooled in liquid ethane.

613 Grids were loaded in a Titan Krios cryo-electron microscope (Thermo Fisher)
614 operated at 300 kV with a 50- μm condenser lens aperture, spot size 5, magnification at
615 165,000 \times (corresponding to a calibrated sampling of 0.85 Å per physical pixel), and a
616 K2 direct electron device equipped with a BioQuantum energy filter operated at 20 eV
617 (Gatan). Movie stacks were collected automatically using EPU2 software (Thermo
618 Fisher) with the K2 detector operating in counting mode at a recording rate of 5 raw
619 frames per second and a total exposure time of 5 s, yielding 25 frames per stack and a
620 total dose of 52.5 e⁻/Å².

621

622 **Cryo-EM data processing**

623 Cryo-EM data were processed using Relion (v3.1)⁵⁴ and CryoSPARC (v3.1)⁵⁵.
624 The frame motion of a total of 4,705 images stacks was corrected by MotionCor2⁵⁶
625 wrapper in Relion. Exposure-weighted averages were then imported to CryoSPARC
626 and the contrast transfer function parameters for each micrograph were estimated by
627 CTFIND4⁵⁷. 2D classification template was generated from a small set of particles
628 via blob-picking. A total of 2,959,791 particles were autopicked using this template and
629 extracted with a box size of 280 pixels, and subjected to several rounds of 2D
630 classification and heterogeneous refinement (3D classification) to remove contaminants
631 or poor-quality particles. A set of 329,617 good GPI-T particles was obtained and

632 converted for Bayesian polishing in Relion, which was subsequently imported back to
633 CryoSPARC for heterogeneous refinement, enabling the production of a 2.66Å map
634 from 179,871 particles by non-uniform refinement (NU-refine). A further round of
635 heterogeneous refinement (3D classification) with phase randomized reference maps
636 (20/30/40Å) narrowed the dataset down to 151,590 particles, which was used to
637 generate the final 2.53Å map via local refinement with a membrane micelle-removed
638 mask. Resolution of these maps was estimated internally in CryoSPARC by gold-
639 standard Fourier shell correlation using the 0.143 criterion. Details for data processing
640 are in Supplementary information (**Fig. S3**) and **Table S1**.

641

642 **Model building**

643 The model was *ab initio* built using Coot⁵⁸ based on the amino-acid sequence and
644 guided by the cryo-EM density using aromatic residues, glycosylation sites, and
645 disulfide bonds as reference markers. The model was refined with Phenix.
646 `real_space_refine`⁵⁹, yielding an averaged model–map correlation coefficient (CCmask)
647 of 0.84. Structures were visualized using UCSF ChimeraX1.1⁶⁰ and PyMol
648 (<https://pymol.org/2/>).

649

650 **Data Availability**

651 The coordinates for GPI-T model have been deposited in the PDB under accession code
652 7WLD. The cryo-EM density map has been deposited in the Electron Microscopy Data
653 Bank with accession code EMD-32582. Data for Figs. 2d/3d/S7d/S8b are included in

654 **Supplementary Data 1.**

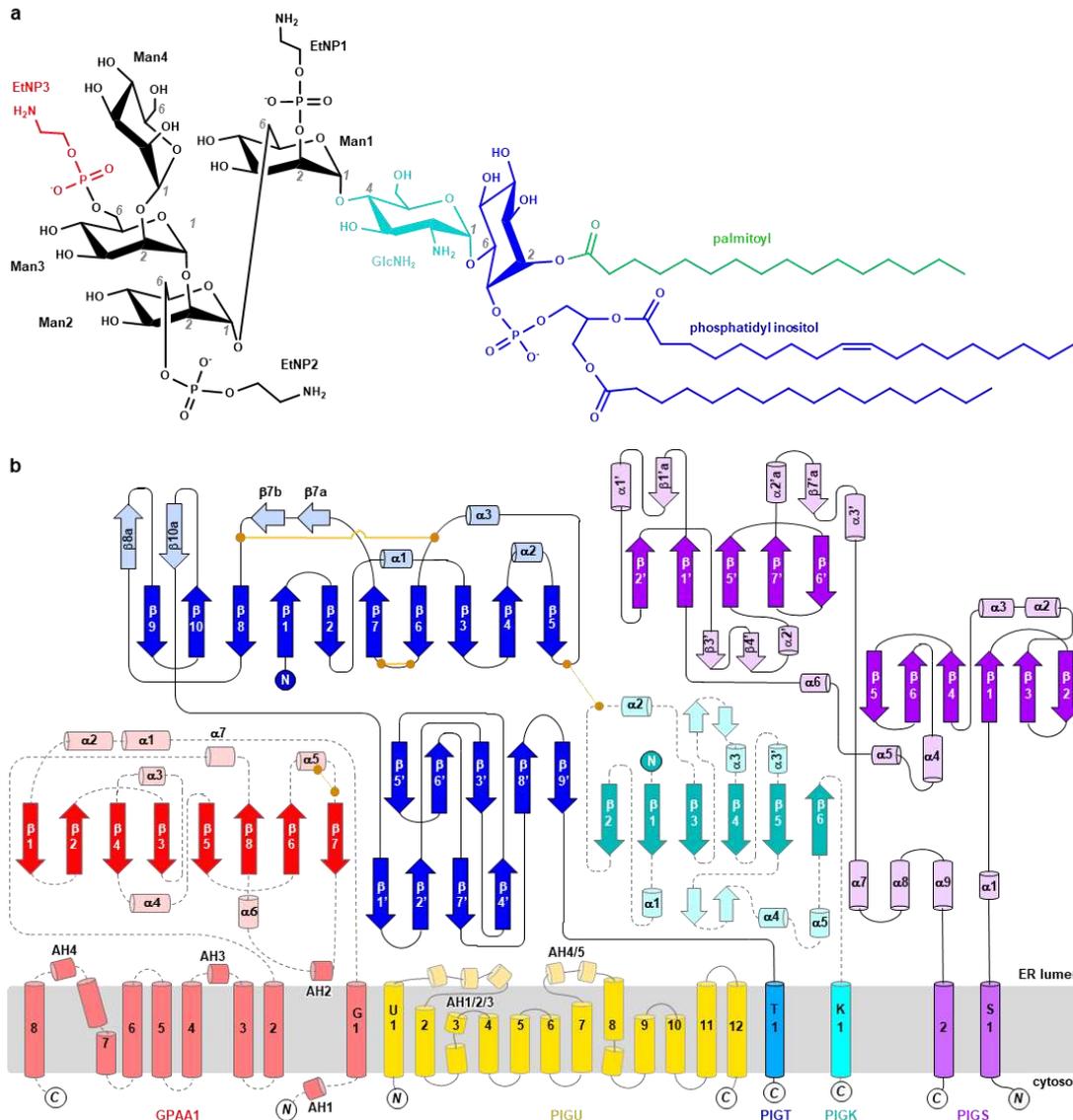
655 **Supplementary Information**

656

657 Fig. S1-S9

658

659 Table S1



660

661

662 **Fig. S1. Chemical structure of GPI and topology of GPI-T. a** The GPI core contains

663 a phosphatidyl inositol (blue), a palmitoyl chain (green), a glucosylamine (GlnNH₂)

664 (cyan), and four mannoses (Man) with possible modification of ethanolamine

665 phosphate (EtnP) on Man1/2. The reactive EtnP3 on Man3 is colored red. The

666 numbering of the carbon atoms on the sugar ring is indicated by a grey number when

667 appropriate. **b** Topology of GPI-T drawn with structural knowledge from this study.

668 Connected orange dots indicate disulfide bonds. Transmembrane helices (TMHs) are

669 indicated by numbers and the ER membrane is shaded grey. Major secondary structures
670 are numbered to match the description in the main text and **Fig. S2**. AH, amphipathic
671 helix. N/C indicates the N- and C-terminal of subunits.

c

			TMH1	AHU1	AHU2	AHU3	TMH2	
<i>H. sapiens</i> (huamn, FIGU)	1	--MAAPLVLV	LVVAIVTRAA	LFRSSLAEFI	SERVEVVSPL	SSWKRVEGL	SLLDLGVSPY	SGAVPHETPL IYLFPHLID YAEIVFMITD
<i>S. cerevisiae</i> (yeast, Gablp)	1	MDSTALKVAL	GCIAIRLAVN	SLFFSLQQQL	DQSVEFSTPV	TSFRSLQEGI	YLLRNNIQVY	NHGVVHHPTI LIFPLSLFN-----
		: *	: *	: *	: *	: *	: *	: *
		: *	: *	: *	: *	: *	: *	: *

		TMH2	TMH3a	TMH3b	TMH4					
<i>H. sapiens</i> (Q9H490)	89	ALTAIALYFA	IQDFNKVVFK	KQKLLLELDQ	YAPDVAEIIR	TPMEMRYIPL	KVALFYLLNP	YTIILSCVAKS	TCAINNTLIA	FFILITIKG-
<i>S. cerevisiae</i> (P41733)	80	SDRLISLIYA	LIDG-----	-----L	IAYQLTEVTR	AFKNLKLKVV	LPGLLYAVNP	LTLLSCLISRS	SIIFTNFAIS	SSLYCILAEG
		: *	: *	: *	: *	: *	: *	: *	: *	: *

		TMH5	TMH6	TMH7	AHU4					
<i>H. sapiens</i>	178	SAPLSAIFLA	LATYQSLYPL	TLFVPGLLYL	LQRQYIPVKM	KSKAFWIFSW	EYAMMYVGS	VVIICLSFFL	LSS--WDFIP	AVYGFILSVP
<i>S. cerevisiae</i> (28.5%/49.3%)	155	NVLLSSVMIS	ISGYLSVYPI	LLLIPLLGML	RS-----	-----WRQRI	LSAIVSILSL	LILLLPSYSI	LGSQSWSPLT	QVYGSIIIFE
		: *	: *	: *	: *	: *	: *	: *	: *	: *

		AHU5	TMH8	TMH9	TMH10					
<i>H. sapiens</i>	266	DLPNIGLWV	YFFAEMPEHF	SLFPVVCVQI	NVPFYTPIPL	IKLKEHPFIP	MPIQIAVIAI	FKSVPTVGVV	ALYMAFFPVW	NHLVPLRNI
<i>S. cerevisiae</i>	233	KVFPNGLMW	YFFIEMFDTF	IPFFKAVENI	FIAVFITPPT	LRVHKQPPYA	FILCIGWIVL	TKPYPSLGDA	GFFPFLPFF	TFPLGYLRYP
		: *	: *	: *	: *	: *	: *	: *	: *	: *

		TMH11	TMH12						
<i>H. sapiens</i>	356	FVLTCTIIVC	SLFPVWHL	WIYAGSANSN	FFYAITLTFN	VGQILLISDY	FYAFRLREYI	LTHGLYLTAK	DGTEAMLVLR
<i>S. cerevisiae</i>	322	IISALLFLHA	IVLAPIFYHL	WVVLGSGNSN	FFYAISLVYA	LAIASILVDL	NWAMLRIEYD	NGINPNKLVK	TQI-----
		: *	: *	: *	: *	: *	: *	: *	: *

d

		TMHS1	α1	β1	β2					
<i>H. sapiens</i> (human, FIGS)	1	MAAAGAAATH	LEVARGKRAA	LFFAAVAIVL	GLPLWKTTE	TYRASLPYSQ	ISGLNALQLR	LMVPEVTVFT	RESVPLDDQE	KLFPVHVER
<i>S. cerevisiae</i> (yeast, Gpi17p)	1	-----M	SNANLRKVVG	FCFVAIYLPF	GVPLWYKLT	VYRASLPINY	IESLQNNKFO	DIHLVPIPVV	KS-----D	TYRFPDVHD-
		: *	: *	: *	: *	: *	: *	: *	: *	: *

		β3	α2	α3	β4	β5	β6	α4		
<i>H. sapiens</i> (Q96S52)	91	EIPLKYMKI	KCRFQKAYRR	ALDHEEERLS	SGSQVEAEM	LDEPQEQAEQ	SLTVYVISEH	SLLPQDMMS	YIGPKRTAVV	RGMHREARN
<i>S. cerevisiae</i> (Q04080)	74	-----	---AIQVQVN	HLLNSQQRV	FWSLQVLPYN	ETIEQMESEG	-----NO	PHVVTIKLDE	FIGYSSADT	KETLWYDDA
		: *	: *	: *	: *	: *	: *	: *	: *	: *

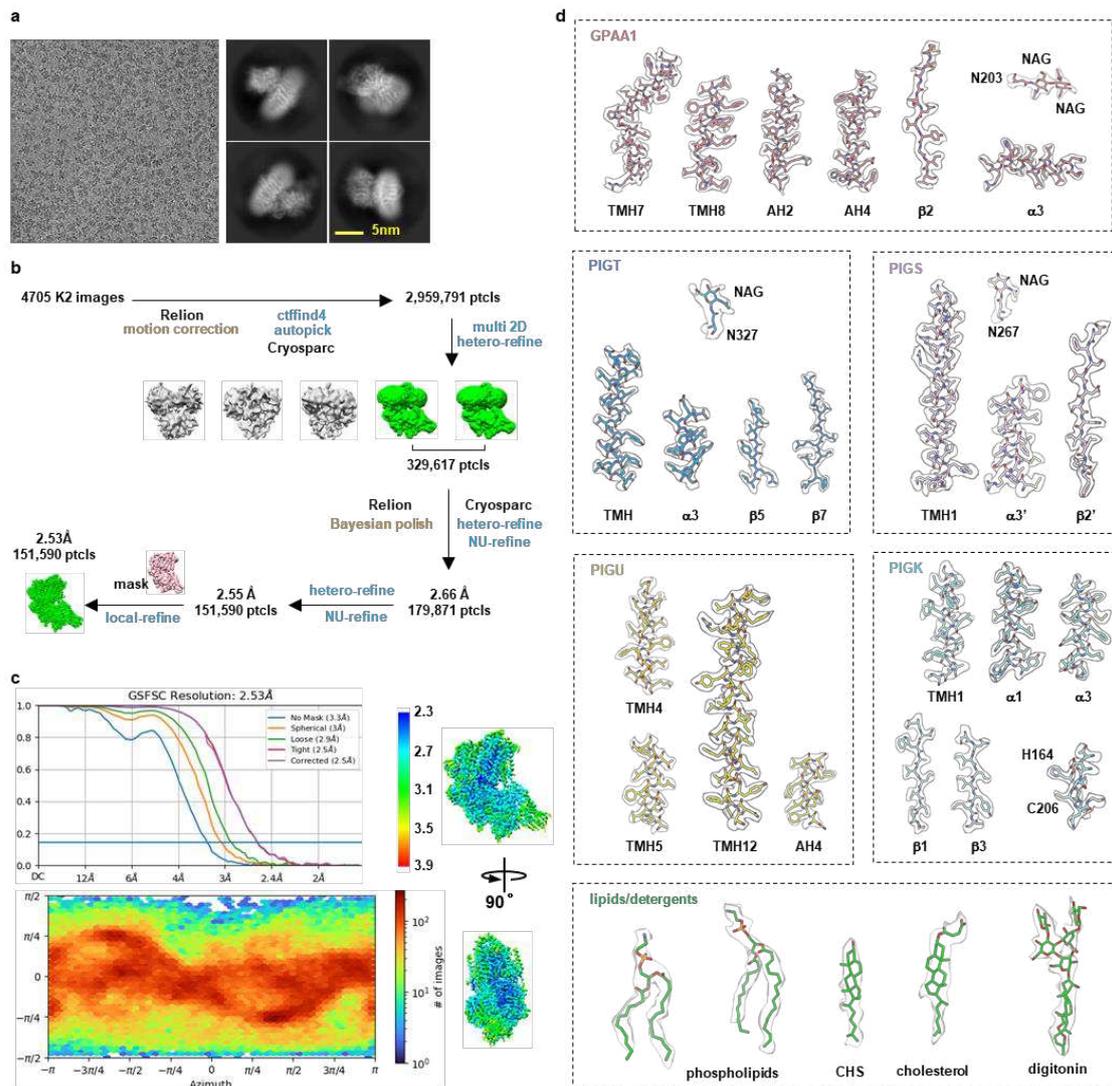
		α4	α5	α6	β1'	β1'a	α1'			
<i>H. sapiens</i>	181	IIGRRIVQVA	QAMSLTEVDL	AAALADHLPE	DKWSAEKRRP	LKSSLGVEIT	FSLNLPDPKS	HDVYWDIEGA	VRRYVQPFNL	ALGAAGNSV
<i>S. cerevisiae</i> (22.1%/38.7%)	143	AVLSNDLPFF	VAQTLVEHTF	QLEWT-HLNK	TCEGVSTNND	VAISYDPNIH	LSVTLLSGDG	NFVAMEIEPT	LTDYFSPFRK	FLSPLVNFIV
		: *	: *	: *	: *	: *	: *	: *	: *	: *

		β2'	β3'	β4'	α2'	β5'	β6'			
<i>H. sapiens</i>	271	DSQILYYAML	GVNFRFDSAS	SSYYLDMHSL	PHVINPVESR	LGSSAASLYP	VLNFLYVPE	LAHSPLYIQD	KDG-----	APVATNAFHS
<i>S. cerevisiae</i>	232	DSSIVYHNDL	NLHSLNGSCT	SVTWFDLS--	-HTIDLSELS	SMAYYPEDS-	ALNLAIVFPF	ASSSPDGLAF	INGTRISDEI	TTLDWNYSLV
		: *	: *	: *	: *	: *	: *	: *	: *	: *

		β7'	α2'a	β7'a	α3'	α7				
<i>H. sapiens</i>	354	PRWGGIMVYN	VDSKTYNASV	LVVRVEVDMV	RVMEVFLAQL	RLLFQIAQFP	LPPKCLLSP	TSEGLMTWEL	DRLWARSVE	NLATATTTLT
<i>S. cerevisiae</i>	318	PQWG-VIIN	KMPLKNSVI	SEDYLEPMXY	RFATIDPQLL	GLTEGSQDLL	SPYITIDS--	-----F	KRLTILQNL	KATEPLMSLV
		: *	: *	: *	: *	: *	: *	: *	: *	: *

		α7	α8	α9	TMHS2					
<i>H. sapiens</i>	444	SLAQLLGKIS	NIVIKDDVAS	EVYKAVAAVQ	KSAEELASG-	--HLASAFVA	SQEAVTSEEL	AFDFPSLLHL	LYFPDDQKFA	IYIPLFLPMA
<i>S. cerevisiae</i>	396	KLTQFPQGMS	IPREVSNNVI	EALDLRLQII	DLNDPQKGG	DIWVNNALHL	SNELVKLCEK	AFNGEMVQQ	NFPQERHMA	VYLPPLGPTIS
		: *	: *	: *	: *	: *	: *	: *	: *	: *

<i>H. sapiens</i>	531	VPILLSLVKI	FLETRKSWRK	-----PE	KTD-----	-----
<i>S. cerevisiae</i>	486	AVMFFGFYNV	MKEKNQSKK	NGTEREVAKE	KLELKEAQL	HAIDGEDEL
		: *	: *	: *	: *	: *



692

693

694 **Fig. S3. Cryo-EM data processing and high-quality density/model fitting**

695 **exemplary views. a** Representative cryo-EM micrograph of GPI-T and selected 2D

696 class averages. **b** The workflow of classification and refinement. **c** The nominal

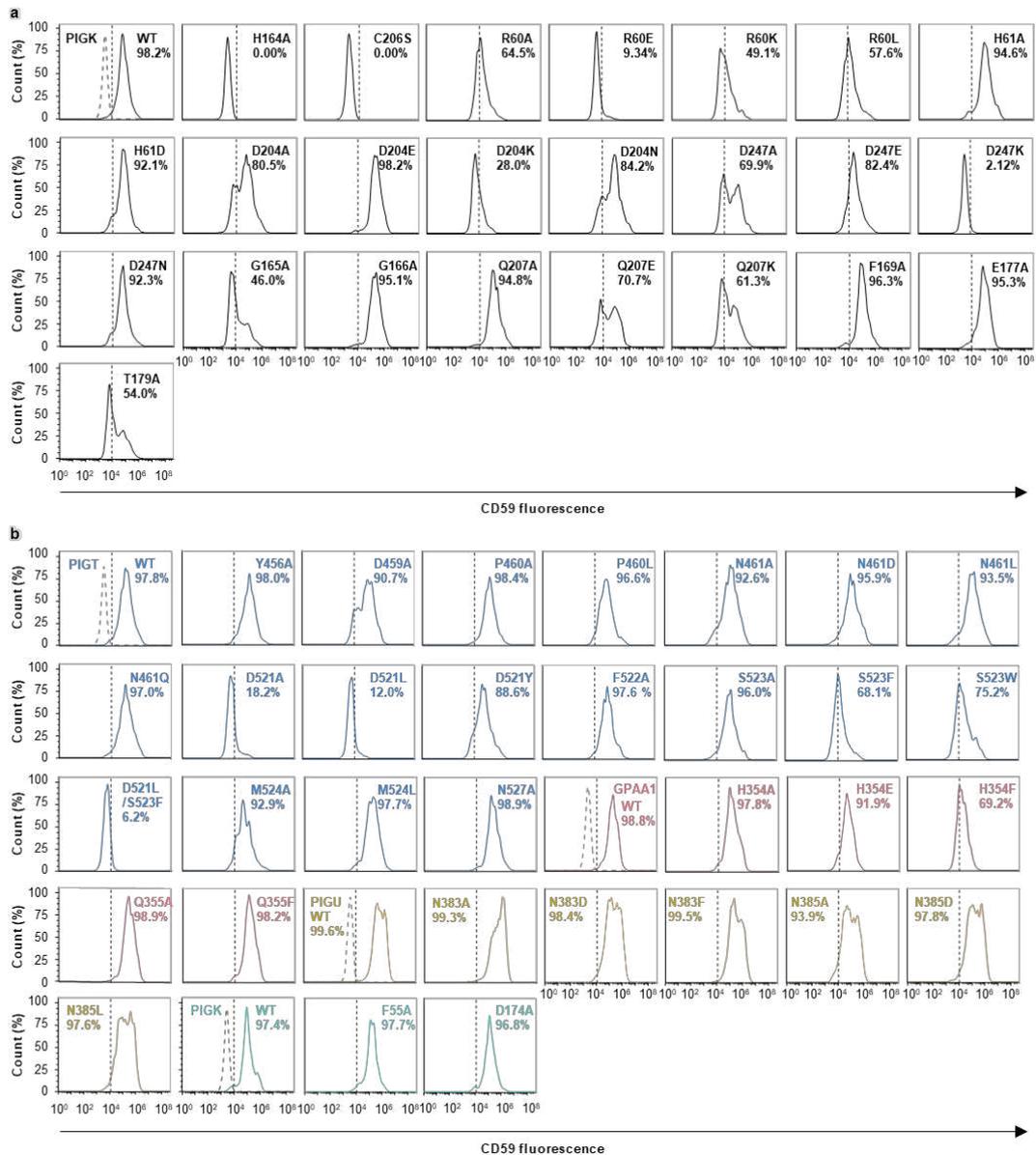
697 resolution of GPI-T was determined by the ‘gold-standard’ FSC curve using the

698 FSC=0.143 criterion, and angular distribution heatmap at the bottom panel calculated

699 in Cryosparc. Local resolution evaluation was shown on right. **d** Cryo-EM map density

700 and model of representative protein parts, glycosylation sites, and lipids and detergents.

712 comparison (**e**). The six β -strands (1-6) in the central sheet and connecting α -helices
713 ($\alpha 1$ - $\alpha 6$) are labeled. The catalytic dyad and S1/S1'/S2' residues are shown as C α
714 spheres (**d**) or sticks (**e**) and colored magenta, green/yellow/pink, respectively.



715

716

717 **Fig. S5. Typical fluorescence-activated cell sorting (FACS) results for the**

718 **constructs described in the main text. a** Representative FACS results of three

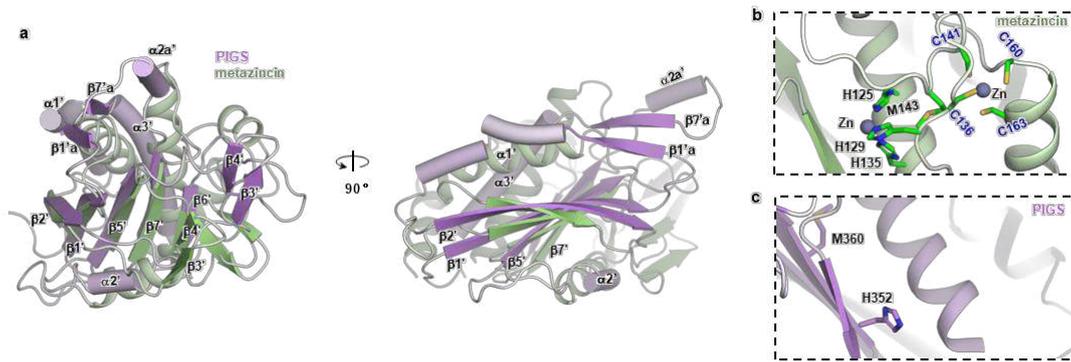
719 independent experiments for the PIGK mutants (the catalytic dyad and S1/S1'/S2' sites).

720 **b** Representative FACS results of three independent experiments for the mutants of the

721 GPI-binding site residues. The percentage shown in this figure represents the

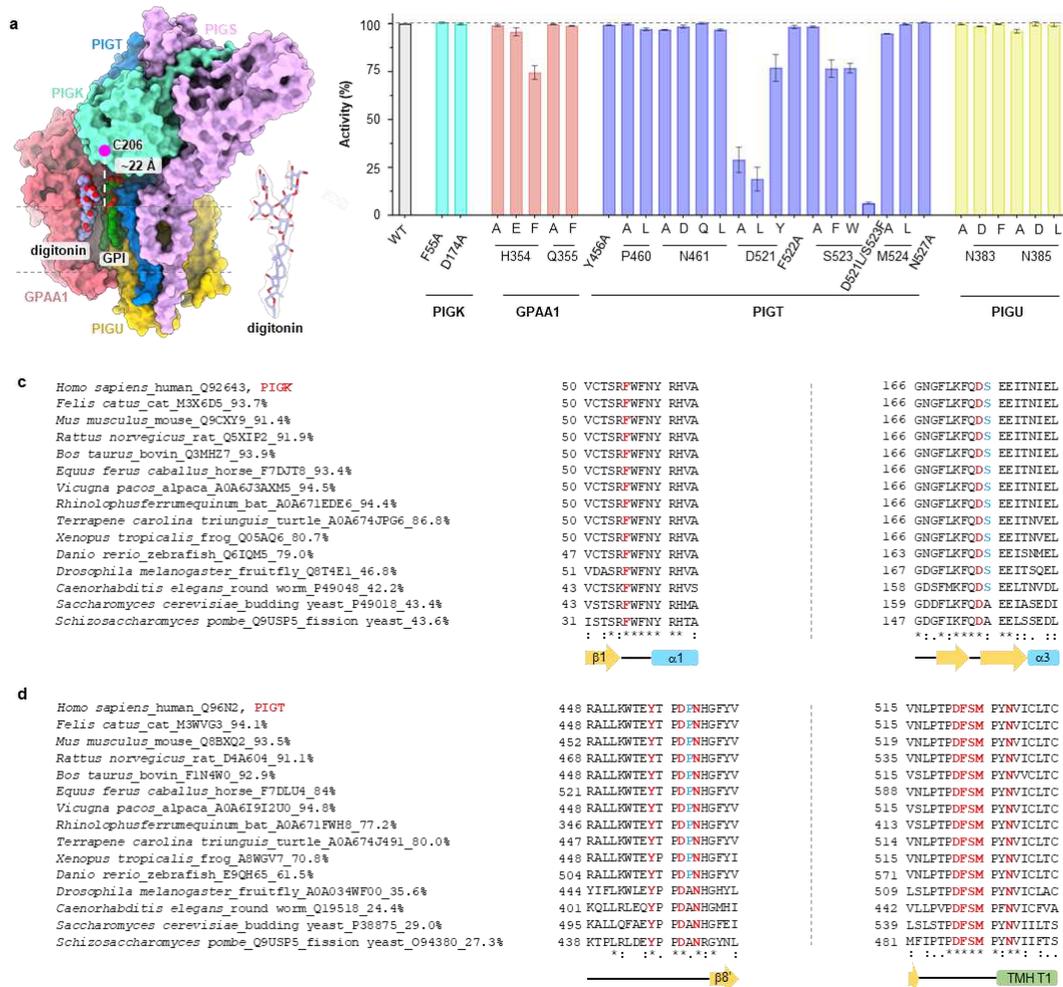
722 subpopulation of CD59-positive cells (as the function of GPI-AP processing) in the

723 GFP-positive population (as a result of subunit expression). To calculate activity, the
724 percentage shown in this figure was normalized using a negative control (grey dashed
725 line and the vertical dash line, cells expressing an unrelated GFP-tagged membrane
726 protein) and a positive control (cells expressing the wild-type). The expression and
727 integrity of the subunits were also separately verified by SDS-PAGE in-gel fluorescence.



728

729 **Fig. S6. PIGS contains a protease-like domain.** **a** The PIGS domain containing
 730 residues 222-398 (purple, cylinder) shares structural similarity with the metazincin
 731 AmzA (green, cartoon, PDB ID 2X7M) with a Z-score of 5.9 and $C\alpha$ RMSD of 3.6 Å
 732 ⁴¹. **b, c** The two Zn-binding motifs in metazincin (**b**) are compositionally impossible for
 733 PIGS (**c**) as only one histidine and one methionine were found in PIGS in the vicinity
 734 (within 10 Å).



735

736 **Fig. S7. The GPI-binding site is delineated with conserved residues and hosts a**

737 **digitonin molecule. a** A digitonin (light blue) which also contains a polyglycan chain

738 is found near GPI (green). The cryo-EM density of the digitonin is shown in the right

739 corner. **b** Apparent activity of GPI-binding site mutants relative to the wildtype (WT).

740 Activity was measured by immune staining of a reporter GPI-AP (CD59) on the surface

741 of KO cells transfected with the indicated mutants. Data represent mean ± s.e.m. from

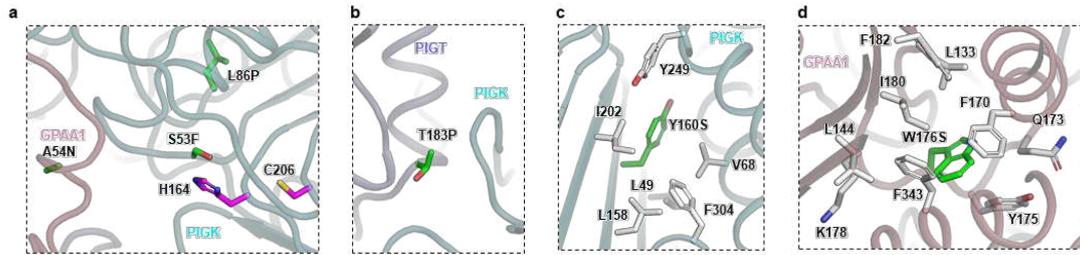
742 three independent experiments (**Supplementary Data 1**). Activity for PIGT

743 D521/S523, GPAA1 H354, and PIGU N385 are re-plotted using data from **Fig. 6d. c,**

744 **d** Sequence alignment of the conserved residues of PIGK (**c**) and PIGT (**d**) around the

745 GPI-binding site (**Fig. 6b**). Protein sequences were selected from evolutionarily

746 representative species. Uniprot IDs of the PIGT/PIGK sequences and their sequence
747 identity to the human ortholog were shown along with the binomial nomenclature and
748 common names. Asterisk, colon, and dot indicate identical, conserved, and semi-
749 conserved substitutions, respectively. Secondary elements were labeled for easier
750 location of the corresponding elements in **Fig. S2**.

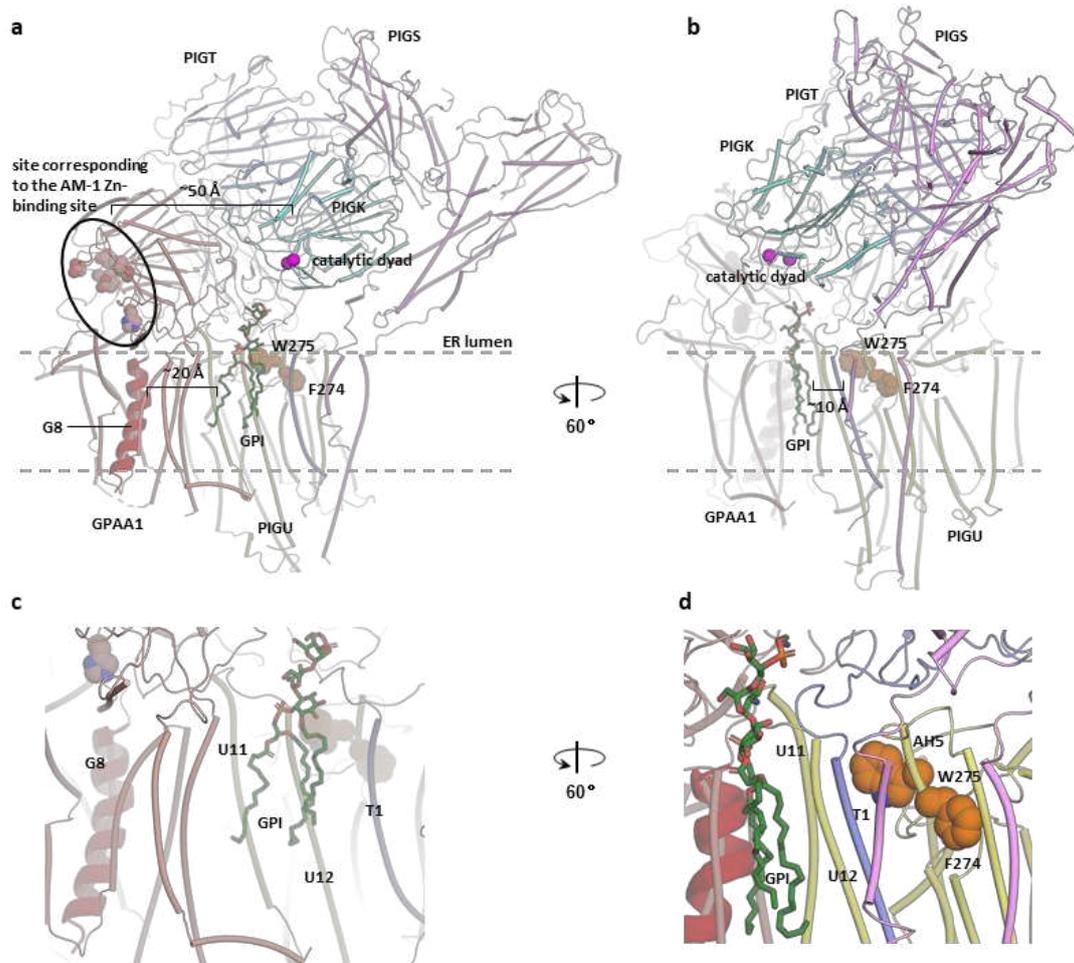


751

752

753 **Fig. S8. Expanded view of the distribution of representative genetic mutations. a,**
 754 **b** Exemplary expanded view of residues close to the catalytic dyad (**a**) and at the inter-
 755 subunit interfaces (**b**). **c, d** The PIGK Y160S (**c**) and PIGT W176S (**d**) may cause
 756 folding issues by introducing a small hydrophilic residue into a large hydrophobic
 757 pocket.

758



759

760

761 **Fig. S9. The site corresponding to the AM-1 Zn-binding site, the TMH8 of GPAA1,**

762 **and F274/W275 in PIGU are not at the proposed active site. Various components**

763 **are labeled appropriately. The catalytic dyad is shown as C α spheres. a, b Overview. c,**

764 **d Expanded view.**

765

766 **Table S1. Cryo-EM data collection and refinement statistics.**

767

Human GPI-T complex	
Data collection and processing	
Magnification	165,000
Voltage (kV)	300
Electron exposure (e ⁻ /Å ²)	52.5
Defocus range (μm)	-2.4 to -0.6
Pixel size (Å)	0.85
Symmetry imposed	C1
Initial particle images (no.)	2,959,791
Final particle images (no.)	151,509
Map resolution (Å)	2.53
FSC threshold	0.143
Map resolution range (Å)	2.3 - 3.9
Refinement	
Model resolution (Å)	2.79
FSC threshold	0.5
Map sharpening <i>B</i> factor (Å ²)	77.5
Model composition	
Non-hydrogen atoms	19,733

Protein residues	2,393
Ligands	27
<i>B</i> factor (Å ²)	
Protein	50.8
Ligand	64.8
R.m.s. deviations	
Bond lengths (Å)	0.019
Bond angles (°)	1.325
Ramachandran plot	
Favored (%)	97.77
Allowed (%)	2.15
Outliers (%)	0.08

768

769

References

- 771 1. Gamage, D.G. & Hendrickson, T.L. GPI transamidase and GPI anchored proteins: oncogenes
772 and biomarkers for cancer. *Crit Rev Biochem Mol Biol* **48**, 446-64 (2013).
- 773 2. Pittet, M. & Conzelmann, A. Biosynthesis and function of GPI proteins in the yeast
774 *Saccharomyces cerevisiae*. *Biochim Biophys Acta* **1771**, 405-20 (2007).
- 775 3. Kinoshita, T. Biosynthesis and biology of mammalian GPI-anchored proteins. *Open Biol* **10**,
776 190290 (2020).
- 777 4. Peter, O. & Menon, A.K. Thematic review series: Lipid Posttranslational Modifications. GPI
778 anchoring of protein in yeast and mammalian cells, or: how we learned to stop worrying and
779 love glycopospholipids. *J Lipid Res* **48**, 993-1011 (2007).
- 780 5. Paulick, M.G. & Bertozzi, C.R. The Glycosylphosphatidylinositol Anchor: A Complex
781 Membrane-Anchoring Structure for Proteins. *Biochemistry* **47**, 6991-7000 (2008).
- 782 6. Ferguson, M.A.J., Homans, S.W., Dwek, R.A. & Rademacher, T.W. Glycosyl-Phosphatidylinositol
783 Moiety That Anchors *Trypanosoma brucei* Variant Surface Glycoprotein to the Membrane.
784 *Science* **239**, 753-759 (1988).
- 785 7. Schofield, L., Hewitt, M.C., Evans, K., Siomos, M.A. & Seeberger, P.H. Synthetic GPI as a
786 candidate anti-toxic vaccine in a model of malaria. *Nature* **418**, 785-9 (2002).
- 787 8. Fujihara, Y. et al. Expression of TEX101, regulated by ACE, is essential for the production of
788 fertile mouse spermatozoa. *Proc Natl Acad Sci USA* **110**, 8111-8116 (2013).
- 789 9. Filmus, J. & Capurro, M. The role of glypicans in Hedgehog signaling. *Matrix Biol* **35**, 248-252
790 (2014).
- 791 10. Cho, C., Wang, Y., Smallwood, P.M., Williams, J. & Nathans, J. Molecular determinants in
792 Frizzled, Reck, and Wnt7a for ligand-specific signaling in neurovascular development. *eLife* **8**,
793 e47300 (2019).
- 794 11. Lubbers, R., van Essen, M.F., van Kooten, C. & Trouw, L.A. Production of complement
795 components by cells of the immune system. *Clin Exp Immunol* **188**, 183-194 (2017).
- 796 12. Nsabimana, A. et al. Alkaline phosphatase-based electrochemical sensors for health
797 applications. *Anal Methods* **11**, 1996-2006 (2019).
- 798 13. Scaranti, M., Cojocar, E., Banerjee, S. & Banerji, U. Exploiting the folate receptor α in
799 oncology. *Nat Rev Clin Oncol* **17**, 349-359 (2020).
- 800 14. McKean, D.M. & Niswander, L. Defects in GPI biosynthesis perturb Cripto signaling during
801 forebrain development in two new mouse models of holoprosencephaly. *Biology Open* **1**,
802 874-883 (2012).
- 803 15. Smith, T.K., Crossman, A., Brimacombe, J.S. & Ferguson, M.A.J. Chemical validation of GPI
804 biosynthesis as a drug target against African sleeping sickness. *EMBO J* **23**, 4701-4708 (2004).
- 805 16. Hong, Y. et al. Human PIG-U and Yeast Cdc91p Are the Fifth Subunit of GPI Transamidase That
806 Attaches GPI-Anchors to Proteins. *Mol Biol Cell* **14**, 1780-1789 (2003).
- 807 17. Meyer, U., Benghezal, M., Imhof, I. & Conzelmann, A. Active Site Determination of Gpi8p, a
808 Caspase-Related Enzyme Required for Glycosylphosphatidylinositol Anchor Addition to
809 Proteins. *Biochemistry* **39**, 3461-3471 (2000).
- 810 18. Ohishi, K. et al. Gaa1p and Gpi8p Are Components of a Glycosylphosphatidylinositol (GPI)
811 Transamidase That Mediates Attachment of GPI to Proteins. *Mol Biol Cell* **11**, 1523-1533
812 (2000).

- 813 19. Eisenhaber, B., Eisenhaber, S., Kwang, T.Y., Grüber, G. & Eisenhaber, F. Transamidase subunit
814 GAA1/GPAA1 is a M28 family metallo-peptide-synthetase that catalyzes the peptide bond
815 formation between the substrate protein's omega-site and the GPI lipid anchor's
816 phosphoethanolamine. *Cell Cycle* **13**, 1912-1917 (2014).
- 817 20. Vainauskas, S., Maeda, Y., Kurniawan, H., Kinoshita, T. & Menon, A.K. Structural Requirements
818 for the Recruitment of Gaa1 into a Functional Glycosylphosphatidylinositol Transamidase
819 Complex. *J Biol Chem* **277**, 30535-30542 (2002).
- 820 21. Ohishi, K., Nagamune, K., Maeda, Y. & Kinoshita, T. Two subunits of
821 glycosylphosphatidylinositol transamidase, GPI8 and PIG-T, form a functionally important
822 intermolecular disulfide bridge. *J Biol Chem* **278**, 13959-67 (2003).
- 823 22. Ohishi, K., Inoue, N. & Kinoshita, T. PIG-S and PIG-T, essential for GPI anchor attachment to
824 proteins, form a complex with GAA1 and GPI8. *EMBO J* **20**, 4088-98 (2001).
- 825 23. Sharma, D.K. et al. Soluble GPI8 restores glycosylphosphatidylinositol anchoring in a
826 trypanosome cell-free system depleted of luminal endoplasmic reticulum proteins. *Biochem*
827 *J* **351 Pt 3**, 717-22 (2000).
- 828 24. Nguyen, T.T.M. et al. Bi-allelic Variants in the GPI Transamidase Subunit PIGK Cause a
829 Neurodevelopmental Syndrome with Hypotonia, Cerebellar Atrophy, and Epilepsy. *Am J Hum*
830 *Genet* **106**, 484-495 (2020).
- 831 25. Knaus, A. et al. Mutations in PIGU Impair the Function of the GPI Transamidase Complex,
832 Causing Severe Intellectual Disability, Epilepsy, and Brain Anomalies. *Am J Hum Genet* **105**,
833 395-402 (2019).
- 834 26. Nguyen, T.T.M. et al. Mutations in PIGS, Encoding a GPI Transamidase, Cause a Neurological
835 Syndrome Ranging from Fetal Akinesia to Epileptic Encephalopathy. *Am J Hum Genet* **103**,
836 602-611 (2018).
- 837 27. Kvarnung, M. et al. A novel intellectual disability syndrome caused by GPI anchor deficiency
838 due to homozygous mutations in PIGT. *J Med Genet* **50**, 521-528 (2013).
- 839 28. Yang, L. et al. Homozygous PIGT Mutation Lead to Multiple Congenital Anomalies-Hypotonia
840 Seizures Syndrome 3. *Frontiers in Genetics* **9**(2018).
- 841 29. Pagnamenta, A.T. et al. Analysis of exome data for 4293 trios suggests GPI-anchor biogenesis
842 defects are a rare cause of developmental disorders. *Eur J Hum Genet* **25**, 669-679 (2017).
- 843 30. Kohashi, K. et al. Epileptic apnea in a patient with inherited glycosylphosphatidylinositol
844 anchor deficiency and PIGT mutations. *Brain Dev* **40**, 53-57 (2018).
- 845 31. Skauli, N. et al. Novel PIGT Variant in Two Brothers: Expansion of the Multiple Congenital
846 Anomalies-Hypotonia Seizures Syndrome 3 Phenotype. *Genes (Basel)* **7**(2016).
- 847 32. Nakashima, M. et al. Novel compound heterozygous PIGT mutations caused multiple
848 congenital anomalies-hypotonia-seizures syndrome 3. *Neurogenetics* **15**, 193-200 (2014).
- 849 33. Nguyen, T.T.M. et al. Mutations in GPAA1, Encoding a GPI Transamidase Complex Protein,
850 Cause Developmental Delay, Epilepsy, Cerebellar Atrophy, and Osteopenia. *Am J Hum Genet*
851 **101**, 856-865 (2017).
- 852 34. Yi, L. et al. Disulfide Bond Formation and N-Glycosylation Modulate Protein-Protein
853 Interactions in GPI-Transamidase (GPIT). *Sci Rep* **8**, 45912 (2017).
- 854 35. Hong, Y. et al. TbGPI16 is an essential component of GPI transamidase in *Trypanosoma*
855 *brucei*. *FEBS Letters* **580**, 603-606 (2006).
- 856 36. Eisenhaber, B., Sinha, S., Wong, W.C. & Eisenhaber, F. Function of a membrane-embedded

- 857 domain evolutionarily multiplied in the GPI lipid anchor pathway proteins PIG-B, PIG-M, PIG-
858 U, PIG-W, PIG-V, and PIG-Z. *Cell Cycle* **17**, 874-880 (2018).
- 859 37. Cai, H. et al. An improved fluorescent tag and its nanobodies for membrane protein
860 expression, stability assay, and purification. *Commun Biol* **3**, 753 (2020).
- 861 38. Fuentes-Prior, P. & Salvesen, G.S. The protein structures that shape caspase activity,
862 specificity, activation and inhibition. *Biochem J* **384**, 201-32 (2004).
- 863 39. Nonis, S.G. et al. Structural and biochemical analyses of concanavalin A circular permutation
864 by jack bean asparaginyl endopeptidase. *Plant Cell* **33**, 2794-2811 (2021).
- 865 40. Dall, E. & Brandstetter, H. Mechanistic and structural studies on legumain explain its
866 zymogenicity, distinct activation pathways, and regulation. *Proc Natl Acad Sci USA* **110**,
867 10940-10945 (2013).
- 868 41. Holm, L. & Rosenström, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38**,
869 W545-W549 (2010).
- 870 42. Waltersperger, S., Widmer, C., Wang, M. & Baumann, U. Crystal structure of archaemetzincin
871 amza from *Methanopyrus kandleri* at 1.5 Å resolution. *Proteins* **78**, 2720-2723 (2010).
- 872 43. Lu, P. et al. Three-dimensional structure of human γ -secretase. *Nature* **512**, 166-170 (2014).
- 873 44. Zimmer, J. Structural features underlying recognition and translocation of extracellular
874 polysaccharides. *Interface Focus* **9**, 20180060 (2019).
- 875 45. Fraering, P. et al. The GPI transamidase complex of *Saccharomyces cerevisiae* contains Gaa1p,
876 Gpi8p, and Gpi16p. *Mol Biol Cell* **12**, 3295-3306 (2001).
- 877 46. Burr, M.L. et al. CMTM6 maintains the expression of PD-L1 and regulates anti-tumour
878 immunity. *Nature* **549**, 101-105 (2017).
- 879 47. Qian, H. et al. Inhibition of tetrameric Patched1 by Sonic Hedgehog through an asymmetric
880 paradigm. *Nature Communications* **10**, 2320 (2019).
- 881 48. Reading, E. et al. The Effect of Detergent, Temperature, and Lipid on the Oligomeric State of
882 MscL Constructs: Insights from Mass Spectrometry. *Chemistry & biology* **22**, 593-603 (2015).
- 883 49. Neumann, B., Chang, C.C.Y. & Chang, T.-Y. A simple method to disrupt and restore subunit
884 interaction of acyl-CoA:cholesterol acyltransferase 1. *MethodsX* **6**, 2242-2247 (2019).
- 885 50. Dall, E., Fegg, J.C., Briza, P. & Brandstetter, H. Structure and mechanism of an aspartimide-
886 dependent peptide ligase in human legumain. *Angew Chem Int Ed* **54**, 2917-2921 (2015).
- 887 51. Jacobitz, A.W., Kattke, M.D., Wereszczynski, J. & Clubb, R.T. Sortase Transpeptidases:
888 Structural Biology and Catalytic Mechanism. *Adv Protein Chem Struct Biol* **109**, 223-264
889 (2017).
- 890 52. Gibson, D.G. et al. Complete Chemical Synthesis, Assembly, and Cloning of a *Mycoplasma*
891 *genitalium* Genome. *Science* **319**, 1215-1220 (2008).
- 892 53. Xu, H. et al. Sequence determinants of improved CRISPR sgRNA design. *Genome Res* **25**,
893 1147-57 (2015).
- 894 54. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in
895 RELION-3. *eLife* **7**, e42166 (2018).
- 896 55. Punjani, A., Rubinstein, J.L., Fleet, D.J. & Brubaker, M.A. cryoSPARC: algorithms for rapid
897 unsupervised cryo-EM structure determination. *Nature Methods* **14**, 290-296 (2017).
- 898 56. Zheng, S.Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved
899 cryo-electron microscopy. *Nature Methods* **14**, 331-332 (2017).
- 900 57. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron

- 901 micrographs. *Journal of Structural Biology* **192**, 216-221 (2015).
- 902 58. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta*
903 *Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).
- 904 59. Afonine, P.V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta*
905 *Crystallogr D Struct Biol* **74**, 531-544 (2018).
- 906 60. Pettersen, E.F. et al. UCSF ChimeraX: Structure visualization for researchers, educators, and
907 developers. *Protein Sci* **30**, 70-82 (2021).
- 908

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FACSdataGPITmutations20220219.xlsx](#)