1   **Free energy landscape for the entire transport cycle of**

2   **triose-phosphate/phosphate translocator**

3   Mizuki Takemoto[1], Yongchan Lee[1], Ryuichiro Ishitani[1,2*], Osamu Nureki[1*]

4   *[1]Department of Biophysics and Biochemistry, Graduate School of Science, The*

5   *University of Tokyo, 2-11-16 Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan.*

6   *[2]Theoretical Molecular Science Laboratory, RIKEN, Wako, Saitama 351-0198, Japan.*

7   *To whom reprint requests should be addressed: Phone: +81-3-5841-4392; E-mails:

8   ishitani@bs.s.u-tokyo.ac.jp and nureki@bs.s.u-tokyo.ac.jp

Abstract

Secondary active transporters translocate their substrates using the electrochemical potentials of other chemicals, undergoing large-scale conformational changes. Despite extensive structural studies, the atomic details of the transport mechanism still remain elusive. Here we performed a series of all-atom molecular dynamics simulations of the triose-phosphate/phosphate translocator (TPT), which exports organic phosphates in the chloroplast stroma in strict counter exchange with inorganic phosphate (P_i). Biased sampling methods, including string method and umbrella sampling, successfully reproduced the conformational changes between the inward- and outward-facing states, along with the substrate binding. The free energy landscape of this entire TPT transition pathway demonstrated the alternating access and substrate translocation mechanisms, which revealed P_i is relayed by positively charged residues along the transition pathway. Furthermore, the conserved Glu207 functions as a "molecular switch", linking the local substrate binding and the global conformational transition. Our results provide atomic-detailed insights into the energy coupling mechanism of antiporter.

Introduction

26   Membrane transporters function as gatekeepers that permeate hydrophilic

27 metabolites through a lipid bilayer. The secondary active transporters catalyze the

28 uptake of compounds essential for cells or organelles or the elimination of unnecessary

29 compounds, using the electrochemical potential of another compound. Generally, these

30 active transporters achieve the transport of substrates by undergoing conformational

31 transitions between the inward-facing (IF) state, the outward-facing (OF) state, and their

32 intermediate, the occluded (Occ) state. This mechanism, called the alternating access

33 mechanism(Drew & Boudker, 2016; Jardetzky, 1966), is universally required for all

34 active transporters, since opening the gates on both sides of the membrane would permit

35 the free diffusion of the substrate across the membrane and hinder the energy coupling.

36 Although the energy coupling mechanism of membrane transporters is essential for all

37 cells and organelles, the molecular basis of the energy coupling for almost all

38 transporters remains elusive.

39   The plastidic phosphate translocator (pPT) family is one of the membrane

40 transporter groups that exchanges phosphorylated C3, C5 and C6 carbon sugars with

41 inorganic phosphate (P_i) on the inner membrane of the plastid in plant and alga

42 cells(Knappe, Flügge, & Fischer, 2003; Weber & Linka, 2011). The pPT family is

43 further divided into four subfamilies, TPT, GPT, XPT and PPT, according to their

44 substrate specificities for triose-phosphate, glucose-6-phosphate, xylulose-5-phosphate

45 and phospho*enol*pyruvate, respectively(Eicks, Maurino, Knappe, Flügge, & Fischer,

46 2002; Fischer et al., 1997; Kammerer et al., 1998; Weber & Linka, 2011). The

47 triose-phosphate/phosphate translocator (TPT) family in green plants exports triose

48 phosphates and 3-phosphoglyceric acids (3-PGA) produced by photosynthesis and

49 imports P_i into stroma(U.-I. Flügge et al., 1991; U. I. I. Flügge et al., 1989; Weber &

50 Linka, 2011). Thus, TPT contributes to efficient carbon fixation and plant

51 growth(Hattenbach, Muller-Rober, Nast, & Heineke, 1997; Häusler, Schlieben, Nicolay,

52 et al., 2000; Häusler, Schlieben, & Flügge, 2000; Heineke et al., 1994; Riesmeier et al.,

53 1993; Schneider et al., 2002). Each pPT subfamily catalyzes the strict 1:1 counter

54 exchange of phosphorylated carbon compounds with P_i to ensure phosphate

55 homeostasis between the stroma and the cytosol(Fliege, Flügge, Werdan, & Heldt, 1978;

56 Weber, Schwacke, & Flügge, 2005). They can also mediate P_i/P_i exchange in a

57 reconstituted system, such as liposomes(Linka, Jamai, & Weber, 2008).

58   Recently, we reported the crystal structure of GsGPT from the thermophilic red

59 alga *Galdieria sulphuraria*(Lee et al., 2017), which is functionally similar to TPT, rather

60 than GPT, despite its name based on the sequence similarity(Linka et al., 2008). GsGPT

61 is composed of 10 α-helical transmembrane (TM) helices (Figure 1a), and adopts the

62 drug metabolite transporter (DMT) superfamily(Jack, Yang, & H. Saier, 2001) fold,

63 similar to that of the recently reported bacterial metabolite transporter, *Starkeya novella*

64 YddG (SnYddG)(Tsuchiya et al., 2016), in which the N-terminal half (TM1-5) is related

65 to the C-terminal half (TM6-10) by two-fold pseudo-symmetry. We obtained the two

66 crystal structures bound with different substrates, P_i and 3-PGA (Figure 1b and Figure

67 1—figure supplement 1). In both crystal structures, the phosphate moieties of the

68 substrates are recognized by the well-conserved basic residues, Lys204, Lys362, and

69 Arg363 (Figure 1b). This central substrate binding site is occluded from the solvent by

70 hydrophobic residues, referred to as the inside and outside gates (Figure 1c). These

71 observations indicated that the crystal structure represents a substrate-bound Occ state.

72 The structural comparison with SnYddG in the OF state revealed the basis of the

73 transport mechanism, in which the opening and closing of these gates regulate the IF

74 and OF conformational changes(Lee et al., 2017). However, the detailed mechanisms of

75 the conformational changes and the substrate binding/releasing by the DMT superfamily

76 have remained elusive.

77   Molecular dynamics (MD) simulations have been used to study the dynamics

78 and conformational transitions of various membrane transporters(Khalili-Araghi et al.,

79 2009; Shaikh et al., 2013). Our previous MD simulation, in which the substrate was

80 removed from the crystal structure of GsGPT, showed the rapid conformational change

81 from the Occ state to the IF and OF states, and provided insights into the transport

82 mechanism of GsGPT(Lee et al., 2017). Generally speaking, a conventional MD

83 simulation cannot sample a large and slow conformational change in the limited

84 computational time, because of the high free energy barriers between the states. In fact,

85 our previous MD simulation of the substrate-bound GsGPT did not reveal any

86 significant conformational change in the 100-ns timescale(Lee et al., 2017). However, in

87 the physiological transport cycle, spontaneous conformational changes between the Occ

88 and IF/OF states, as well as the binding and release of the substrate (Figure 1d), actually

89 occur in a much longer time scale, which is not accessible by the conventional MD

90 simulation methods.

91   This problem can be solved by employing biased sampling methods, such as

92 umbrella sampling(Torrie & Valleau, 1977). In this study, we used bias-exchange

93 umbrella sampling (BEUS)(Moradi, Enkavi, & Tajkhorshid, 2015; Moradi &

94 Tajkhorshid, 2014; Sugita, Kitao, & Okamoto, 2000), an umbrella sampling method

95 combined with Hamiltonian replica exchange sampling(Sugita et al., 2000), which can

96 perform more efficient sampling than normal umbrella sampling. To find the optimum

97 transition pathway of the entire transport cycle of GsGPT, we employed the string

98 method with swarms of trajectories (SMwST)(Pan, Sezer, & Roux, 2008) path-finding

99 algorithm. By combining these sampling techniques, we achieved the *in silico*

100 reconstruction of a feasible transport cycle for GsGPT in the explicit solvent and lipid

101 bilayer, and we observed the substrate binding and releasing events for the first time for

102 the DMT superfamily transporters. Comprehensive structural analyses revealed the

103 strict 1:1 exchange mechanism of GsGPT in atomic detail. We found the unexpected

104 common feature of the conformational transition mechanisms with the well-studied

105 major facilitator superfamily (MFS) transporters(Quistgaard, Löw, Guettou, &

106 Nordlund, 2016), despite their completely different protein folding. Along with the

107 computational simulations, our functional analysis of the mutants provided

108 experimental evidence supporting the MD simulation observations. Taken together, we

109 revealed the atomistic molecular mechanism of GsGPT, which involves the structural

110 coupling between the local conformational change induced by substrate binding and the

111 global conformational transition of the transporter.

112

Results

Reconstruction of the GsGPT transport cycle

To reconstruct the GsGPT transport cycle, we first tried to find the reaction coordinates (collective variables) to induce the conformational transition of GsGPT. Various trials revealed that two distance parameters, D_{in} and D_{out}, were sufficient to induce the transition to the IF and OF states from the Occ state crystal structure (Figure 1—figure supplement 2b-e), where D_{in} and D_{out} represent the center of mass distances between the Cα atoms of the inside and outside gate residues, respectively (see Methods for a detailed definition). Moreover, the one-dimensional collective variable, $\Delta D = D_{out} - D_{in}$, can also be used to induce the conformational change from IF to OF (Figure 1—figure supplement 2f). For simplicity, we used this ΔD to describe the conformational space of GsGPT. Another reaction coordinate for the transition cycle is the position of P_i along the membrane normal axis, referred to as Z_{Pi} (Figure 1—figure supplement 2a). Finally, we performed 12 sets of biased samplings in the two-dimensional collective variable space, ΔD and Z_{Pi}.

The reconstruction of the transport cycle is performed by the following steps (Figure 1e): first of all, we performed the non-equilibrium pulling to induce the transitions from Occ to IF_b^{cent} and OF_b^{cent} (IF- and OF-bound at the central binding site, respectively) (Figure 1d), using the collective variable ΔD (simulation sets 1 and 2 in Table 1). The subsequent non-equilibrium pulling for Z_{Pi} was performed to generate the structures in which P_i is bound near the gate, IF_b^{intra} (IF-bound at intracellular gate) and OF_b^{extra} (OF-bound at extracellular gate) (simulation sets 3 and 4). To search the minimum free energy pathway of the $IF_b^{intra} \leftrightarrow Occ \leftrightarrow OF_b^{extra}$ transition in the ΔD - Z_{Pi}

136 two-dimensional space, we employed the string method with swarms of trajectories

137 (SMwST)(Pan et al., 2008) (simulation set 5), followed by the free energy calculation

138 with a set of bias exchange umbrella sampling (BEUS)(Moradi et al., 2015; Moradi &

139 Tajkhorshid, 2014; Sugita et al., 2000) along the refined transition string (simulation set

140 6) using "path collective variables" (pathCV(s) and pathCV(ζ); equations (1) and

141 (2))(Branduardi, Gervasio, & Parrinello, 2007). Next, to obtain the *apo* states of GsGPT,

142 IF_a and OF_a in Figure 1c, we pulled the P_i from the IF_b^{cent} and OF_b^{cent} states to the bulk

143 solvent region (simulation sets 7 and 9), and performed free energy calculations along

144 the Z_{Pi} variable (simulation sets 8 and 10). The results of these three sets of free energy

145 calculations were combined with the recently developed *post-hoc* string method

146 (PHSM)(Moradi et al., 2015) to estimate the minimum free energy pathway for the full

147 transition cycle, $IF_a \leftrightarrow Occ \leftrightarrow OF_a$. The extracted transition pathway was further

148 refined with SMwST (simulation set 11). Finally, the free energy calculation for the

149 entire transport cycle (simulation set 12) was performed along this optimized transition

150 pathway, using pathCV(s) and pathCV(ζ). We obtained the entire transport cycle of

151 GsGPT with 1,000,000 configurations, and calculated the free energy along the

152 optimum transition pathway using the weighted histogram analysis method

153 (WHAM)(Kumar, Rosenberg, Bouzida, Swendsen, & Kollman, 1992) (Figure 1f and

154 Video 1). (For detailed parameters of the simulations, see Methods.)

155

156 **Overall conformational change of GsGPT**

157 To investigate the overall conformational change of GsGPT, representative

158 structures corresponding to the free energy basins of the IF_a and OF_a states (Figure 1f)

159   were compared with that of the Occ state (Figure 2a, b). The major differences between

160   the IF and OF states were the outward-gate distance between the hairpin of TM3-4 and

161   TM1 (Figure 2a), and the inward-gate distance between the hairpin of TM8-9 and TM6

162   (Figure 2b). The superimposition of the IF_a and OF_a states onto the Occ state revealed

163   that the TM bundles composed of TM1, 8 and 9 (bundle1) and TM6, 3 and 4 (bundle2)

164   are directly involved in the overall conformational change. In contrast, the other helices

165   (TM2, 5, 7 and 10) do not participate in the open/closed conformational changes, and

166   function as a scaffold for the conformational changes of bundle1 and bundle2. Hereafter,

167   we refer to bundle1 and bundle2 (*i.e.*, TM1, 3, 4, 6, 8, and 9) as the core domain.

168   The molecular surfaces for the representative structures illustrate that the

169   substrate binding site is exposed to opposite sides of the membrane in an alternating

170   fashion, known as the alternating access mechanism(Drew & Boudker, 2016; Jardetzky,

171   1966) (Figure 2c). To confirm whether all of the structures obtained from the simulation

172   satisfy the alternating access mechanism, we performed pore radius calculations along

173   the transport pathway by HOLE(Smart, Neduvelil, Wang, Wallace, & Sansom, 1996) for

174   all structures obtained from the simulation (for representative results, see Figure

175   2—figure supplement 1). The outward and inward minimum radii were defined as the

176   minimum radii of the pore in the regions of -20<z<-7 and 7<z<20, respectively, where z

177   represents the coordinate along the pore pathway (z = 0 corresponds to the central

178   binding site, and z > 0 corresponds to the outward direction). The plot of the outward *vs.*

179   inward minimum radii illustrates that the conformational transition between the IF and

180   OF states always visits the Occ-state region, and none of the structures visit the region

181   where the substrate binding site is exposed to both the inward- and outward-side solvent

182   (Figure 2d). This result indicates that the obtained transition pathway satisfies the

183   alternating access, suggesting the validity of the simulated transition pathway.

184   Next, to reveal the mechanism of the overall conformational change, we

185   investigated the internal conformational change within the core domain helices, bundle1

186   and bundle2. Bundle1 and bundle2 in the IF_a conformation both superimpose well onto

187   those in the OF_a conformation, respectively (Figure 2—figure supplement 2a, b),

188   suggesting the semi-rigid body movement of the core domains along the transport cycle.

189   In fact, the distributions of the root-mean-square deviation (RMSD) values of the

190   bundle1 and bundle2 Cα atoms against the crystal structure illustrated that the RMSD

191   values for only bundle1 or bundle2 remained lower than that of bundle1+2 over the

192   entire transport cycle (Figure 2—figure supplement 2c). These results suggest that the

193   conformational transition between IF and OF is governed by the semi-rigid body tilting

194   motions of bundle1 and bundle2, rather than the internal bending and straightening

195   motion of the TM helices. To confirm this mechanism, we projected the free energy

196   onto the two-dimensional space consisting of the first principal component of the Cα

197   atoms, or PC1 (Figure 2—figure supplement 3a-c), and the inter-helical angle between

198   TM1 and TM6 (Figure 2e). The PC1 correlates well with the ΔD, pathCV(s), and replica

199   number (Figure 2—figure supplement 3d-f), and represents the global conformational

200   transition as well as the ΔD. The inter-helical angle between TM1 and TM6 represents

201   the inter-domain angle between bundle1 and bundle2. This free energy landscape shows

202   three distinct free energy basins corresponding to the IF, Occ and OF states, suggesting

203   that the tilting motions of bundle1 and bundle2 play a major role in the overall

204   conformational change.

205   These results are in contrast with the previously suggested mechanism

206   hypothesized from the static crystal structures of SnYddG and GsGPT, in which the

207   bending motion of the TM helices governs the conformational transition(Lee et al.,

208   2017). In the previous report of the bacterial DMT superfamily transporter, SnYddG,

209   the bending and straightening of the core bundle, especially TM3-4 and TM8-9, were

210   suggested to regulate the opening and closing of the inside and outside gates(Tsuchiya

211   et al., 2016). However, the present results suggest that the tilting motion of the

212   semi-rigid core domain, which resembles the rocker-switch mechanism proposed for the

213   MFS transporters(Huang, Lemieux, Song, Auer, & Wang, 2003; Law, Maloney, & Wang,

214   2008), is sufficient to explain the conformational transition of GsGPT. Thus, the

215   alternating access mechanism of the DMT superfamily is unexpectedly similar to that of

216   the MFS transporters, despite their completely different TM topologies.

217

218   Hydrogen bonds at the inside and outside gates are dispensable for the

219   conformational change

220   In the crystal structure of GsGPT, the well conserved Lys271 and Lys128

221   residues form hydrogen bonds with the main chain carbonyl atoms on the opposite

222   bundle, and lock the inside and outside gates, respectively (Figure 3a). In our previous

223   analysis based on the structural comparison with SnYddG, the disruption and formation

224   of these hydrogen bonds were suggested to regulate the gate opening and closing(Lee et

225   al., 2017). However, in the present simulation, the two hydrogen bonds of Lys128

226   shown in Figure 3a were not always maintained, even in the IF state (Figure 3b; cyan),

227   and the two hydrogen bonds of Lys271 were also not always maintained, even in the OF

228   state (Figure 3b; magenta). Furthermore, the liposome-based transport assays of K271A,

229   K128A and their double mutant revealed that all of these mutants still retain the P_i/ P_i

230 homo-exchange activity, as compared with the previously reported mutants in the

231 substrate binding site, K204A, K362A and R363A(Lee et al., 2017) (Figure 3c). These

232 results suggest that the hydrogen bond formation by Lys271 or Lys128 is not crucial for

233 the conformational change. Our investigations are consistent with the previous

234 functional analysis of the GsGPT homologue, the apicoplast phosphate translocator of

235 *Toxoplasma gondii* (TgAPT), which revealed that the K67A mutation of TgAPT,

236 corresponding to the K128A mutation of GsGPT, retained 10-20% activity(Brooks et al.,

237 2010). Our simulation suggests that Lys128 and Lys271 play another role in the

238 conformational transition of GsGPT, rather than the regulation of gate opening and

239 closing by hydrogen bond formation.

240

241 **Substrate binding and translocation mechanism**

242 To gain detailed insight into the mechanism for substrate binding and

243 translocation, we projected the free energy onto the PC1- Z_{Pi} plane, and compared the

244 representative structures of the free energy basins (Figure 4). The translocation of P_i

245 from the outside (intermembrane side) to the inside (stroma side) can be described by

246 the following reaction steps (Video 1). (i) At first, P_i in the outside solution is captured

247 by the positively charged residues, and passed to Lys128 (Figure 4a and Figure

248 4—figure supplement 1a). (ii) In addition to Lys128, P_i is captured by Lys204, one of

249 the central substrate binding residues (Figure 4b and Figure 4—figure supplement 1b).

250 (iii) The salt bridge between P_i and Lys128 is disrupted, and then P_i binds to Lys362 and

251 Arg363 (Figure 4c and Figure 4—figure supplement 1c, d). (iv) The hydrogen bond

252 between P_i and Tyr339 on TM9 is formed (Figure 4d and Figure 4—figure supplement

253   1e), which enables the outward halves of the core domains to approach each other, and

254   the conformational change from the OF to Occ states is facilitated. (Figure 4d). This

255   configuration is the most stable state in our simulation (Figure 1f and Figure 4),

256   consistent with the fact that the two crystal structures of GsGPT were obtained in the

257   substrate-bound Occ state. Note that two distinct configurations were degenerate in this

258   Occ state in terms of the binding modes of Arg266: in one state, Arg266 binds to P_i with

259   a water-mediated indirect interaction, as observed in the crystal structure (Figure

260   4—figure supplement 2a); in the other, Arg266 directly forms a salt bridge with P_i

261   (Figure 4—figure supplement 2b). These two states can be discriminated by the slight

262   difference of Z_{P_i}, and this difference might determine the direction of the

263   conformational transition from the Occ state to the IF or OF state. (v) The inside gate is

264   disrupted, and the conformational change from the Occ to IF state occurs (Figure 4e).

265   (vi) P_i is dissociated from the central binding site and bound to both Arg266 and Lys271

266   (Figure 4f and Figure 4—figure supplement 1f, g). (vii) The salt bridge between P_i and

267   Arg266 is disrupted, and P_i is released to the inside solution mediated by Lys271 and

268   other basic residues located on the inward surface, $e.g.$ Lys400 (Figure 4g). Each free

269   energy basin corresponds to a distinct substrate binding mode, and the transport

270   mechanism can be explained as the relaying of P_i by these pore lining, basic residues

271   (except for Tyr339). The free energy barriers between the P_i binding modes correspond

272   to the energy required for disrupting the salt bridge between P_i and these basic residues.

273   The minimum distance plot between P_i and these pore lining residues shows

274   that the residues in the central binding site, Lys204, Tyr339, Lys362 and Arg363,

275   interact strongly with the substrate in all of the OF, Occ and IF states (Figure 4—figure

276   supplement 1b-e). This observation is consistent with our previous mutational analysis,

— 13 —

277   in which the mutations of these residues, K204A, Y339F, K362A and R363A, abolished

278   the P_i/P_i homo-exchange activity(Lee et al., 2017). It is also interesting to note that

279   Lys271 and Lys128 in the inside and outside gates strongly interact with P_i in the

280   transport cycle. Given that the hydrogen bond formation of Lys128 and Lys271 is

281   dispensable for the conformational transition (Figure 3), the decreases of the transport

282   activity in the K128A, K271A and K128A/K271A mutants may be due to the loss of

283   these interactions. Thus, we can conclude that Lys128 and Lys271 primarily function as

284   the initial binding sites for the substrates, rather than in the inside and outside gate

285   formation. To further support this relaying mechanism of P_i, we performed an additional

286   mutational analysis of GsGPT using the liposome-based assay (Figure 4h). The result

287   revealed that the R266A mutant also completely abolishes the P_i exchanging activity,

288   showing the importance of Arg266 for the transport mechanism. In addition, the

289   conservative mutants, R266K and R363K, exhibited significantly decreased activities

290   (Figure 4h), suggesting that the shapes of these side chains are also important for the

291   recognition of the tetrahedral arrangement of the substrate.

292

293   **Coupling mechanism of substrate binding and protein conformational**

294   **change**

295   To reveal the conformational coupling mechanism of GsGPT, we focused on

296   Glu207 on TM4, which is the only conserved acidic residue in the positively-charged

297   substrate binding pocket (Figure 5a). Calculations of the minimum distances between

298   the neighboring basic residues, Lys204 and Arg363, revealed that Glu207 exchanges its

299   salt bridge partner during the transport cycle (Figure 5a, b). In the IF_a state, Glu207

300 mainly forms an intra-helical salt bridge with Lys204 (Fig 5b; magenta line). This

301 intra-helical salt bridge fixes the side chain of Lys204 to the Glu207 side, thereby

302 causing the electrostatic repulsion between Lys204 and Arg363 to prevent the

303 conformational transition to Occ and stabilize the IF conformation (Figure 5a; left). By

304 contrast, in the OF_a state, Glu207 forms a stable inter-helical salt bridge with Arg363 on

305 TM9 (Fig 5b; cyan line). This inter-helical salt bridge between Glu207 and Arg363

306 clamps the inward halves of TM4 and TM9 together (Figure 5a; right), thereby fixing

307 the arrangement of bundle1 and bundle2 to the OF conformation. P_i binding to these

308 basic residues weakens the salt bridges between the basic residues and Glu207, which

309 may facilitate the exchange of the salt bridge partner of Glu207 (Figure 5a; center).

310   This salt-bridge rearrangement involving Glu207 leads to a plausible mechanism

311 of the conformational coupling, which can explain how the local conformational change

312 upon the substrate binding is amplified to the global motions of bundle1 and bundle2. In

313 this mechanism, Glu207 plays a role as a "switch" that turns on the global tilting

314 motions of TM4 and TM9 upon substrate binding. Since TM4 and TM9 constitute the

315 (semi-)rigid bodies of bundle1 and bundle2, respectively, these tilting motions of TM4

316 and TM9 directly result in the global IF/OF conformational change of GsGPT (Figure

317 5—figure supplement 1). This mechanism is in contrast to our previous working model

318 based on the static crystal structure, in which the electrostatic repulsion by the

319 positively charged residues in the substrate binding pocket is simply neutralized by the

320 binding of the negatively-charged substrate, thereby enabling the global IF/OF

321 conformational change(Lee et al., 2017). To further support the importance of Glu207,

322 we performed liposome-based assays of Glu207 mutants. The results revealed that the

323 E207A mutation abolishes the P_i/P_i homo-exchange activity (Figure 5c), in spite of its

324   stable membrane expression (Figure 3—figure supplement 2). Moreover, the

325   conservative mutation, E207D, significantly reduces the exchange activity (Figure 5c),

326   probably due to the insufficient side chain length to form the salt bridge network (Figure

327   5a). Taken together, in addition to the previous mechanism based on the electrostatic

328   repulsion and neutralization, the switch mechanism by Glu207 plays a pivotal role in

329   the coupling mechanism of the local and global conformational changes, thus allowing

330   the strict 1:1 exchange of the substrates by GsGPT.

331

Discussion

Our simulation, combining string method and umbrella sampling, successfully reconstructed the entire transport cycle of GsGPT from a single crystal structure. The computational analysis for all configurations obtained from the simulation revealed the antiport mechanism of GsGPT in atomic detail.

In the current study, we used a highly symmetric substrate, P_i, to simplify the collective variable used in the simulations. However, we can gain insight into other pPT family protein substrates that have more complicated structures, such as 3-PGA. Although pPT family proteins transport various types of phosphosugars, depending on their subtypes(Lee et al., 2017; Weber & Linka, 2011), all of the substrates have a common structural feature: one oxygen atom of P_i is replaced with a sugar moiety. In our simulation, the substrate binding pocket in all of the binding modes can accommodate a compound in which one oxygen atom of P_i is replaced with a sugar moiety (Figure 4a-g), as observed in our previous 3-PGA bound crystal structure(Lee et al., 2017). In addition, our previous study revealed that the recognition and discrimination of the sugar moiety of the substrates is achieved by the side chains located on the opposite side of the P_i binding residues, such as His185 (Figure 1—figure supplement 1b), which are not involved in the currently proposed transport mechanisms. The conformational transition is regulated by the binding of the phosphate moiety, and the binding manner of the phosphate moiety is the same in both the P_i and 3-PGA bound crystal structures(Lee et al., 2017) (Figure 1 and Figure 1—figure supplement 1). Therefore, it is highly likely that the transport mechanism of other substrates by GsGPT is similar to that of P_i observed in our simulations. Moreover, given that the P_i binding

355 residues observed in our simulation (Lys128, Lys204, Arg266, Lys271, Tyr339, Lys362

356 and Arg363) are well conserved among the pPT family members(Lee et al., 2017), and

357 that the free energy landscape is governed almost entirely by the binding mode of P_i to

358 these conserved residues, all of the transporters belonging to the pPT family may have

359 similar free energy landscapes to that of GsGPT (Figure 4). Thus, our results can be

360 generalized to the transport mechanism for all pPT family transporters.

361   Our previous research on the bacterial DMT superfamily protein, SnYddG,

362 revealed that the best-characterized small drug resistance (SMR) family protein, *E. coli*

363 EmrE(Chen et al., 2007; Henzler-Wildman, 2012), shares a similar TM topology with

364 SnYddG, suggesting an evolutional relationship between SMR and other DMT

365 proteins(Tsuchiya et al., 2016). EmrE forms a dimer of four TM segments, and

366 exchanges protons and cationic drugs(Chen et al., 2007; Henzler-Wildman, 2012). The

367 structural comparison of EmrE with SnYddG and GsGPT revealed that EmrE lacks the

368 TM helices corresponding to TM2 and TM7 in SnYddG and GsGPT(Tsuchiya et al.,

369 2016). Our simulation showed that these TM helices work as a scaffold, and are not

370 directly involved in the conformational transition. Thus, the transport mechanism of

371 EmrE can also be explained as the rocker-switch movement of a pair of three TM

372 bundles (TM1, 2 and 3 of each protomer), corresponding to bundle1 and bundle2 of

373 GsGPT (Figure 2c).

374   It is worth comparing the conformational regulation mechanism of GsGPT with

375 that of GlpT, which belongs to MFS and functions as a glycerol-3-phosphate/P_i

376 exchanger(Huang et al., 2003; Law, Maloney, et al., 2008). GlpT has a different

377 evolutionary origin and protein folding from those of GsGPT, but is functionally similar

378 to GsGPT. The previous biochemical and MD simulation studies(Law, Almqvist, et al.,

379 2008; Moradi et al., 2015) revealed that the inter- and intra-domain salt bridges between

380 two basic (R45 and K46) and two acidic (D274 and E299) residues stabilize the

381 different conformations during the transport cycle, and substrate binding weakens these

382 interactions, allowing the conformational transition. In the case of GsGPT, our

383 simulations revealed that the inter- and intra-helical salt bridges involving Glu207,

384 which connect bundle1 and bundle2, stabilize the IF and OF conformations (Figure 5),

385 and these interactions are weakened by substrate binding. Furthermore, the overall

386 conformational change is described by the rocker-switch motions of bundle1 and

387 bundle2 (Figure 2), which are widely observed in MFS transporters, including

388 GlpT(Huang et al., 2003; Law, Maloney, et al., 2008). Therefore, the conformational

389 change mechanism by the rocker switch motion and the conformational regulation

390 mechanism by the salt bridge formation and disruption in GlpT and GsGPT are quite

391 similar to each other. These similarities between GlpT and GsGPT over protein

392 superfamily, in terms of both functions and mechanisms, may represent an example of

393 the convergent evolution of the transporter mechanisms.

394

Methods

Simulation system setup

All MD simulations were performed with NAMD 2.9-2.11(Phillips et al., 2005), along with its Colvars Module(Fiorin, Klein, & Hénin, 2013). The simulation system was 96x96x100 Å^3, and contained GsGPT (PDB ID: 5Y78), a molecule of P_i, a POPC bilayer, 150 mM NaCl, and TIP3 water molecules. All of the water molecules that were visible in the crystal structure were kept. The missing hydrogen atoms were added by the psfgen plugin of VMD(Humphrey, Dalke, & Schulten, 1996). The net charge of the system was neutralized by adding 150 mM NaCl. The molecular topology and force field parameters from Charmm36 were used(Best et al., 2012; Klauda et al., 2010). The systems were first energy minimized for 1,000 steps with fixed positions of the non-hydrogen atoms, and then for another 1,000 steps with 10 kcal/mol restraints for the non-hydrogen atoms, except for the lipid molecules within 5.0 Å from the protein. Next, equilibrations were performed for 0.1 ns under NVT conditions, with 10 kcal mol^{-1} Å^{-2} restraints for the heavy atoms of the protein. Finally, equilibration was performed for 1.0 ns under NPT conditions with the 1.0 kcal mol^{-1} Å^{-2} restraints for all Cα atoms of the protein, followed by a 100 ns equilibration without the restraints. The equilibrated system was further used for the next simulation, as indicated in Table 1 and below. For all simulations in Table 1, the SHAKE algorithm was used to constrain all hydrogen-containing bonds to enable a 2 fs timestep. Constant temperature was maintained with Langevin dynamics at 310 K, and constant pressure was maintained with a Nosé -Hoover Langevin piston(Feller, Zhang, Pastor, & Brooks, 1995) at 1 bar. Electrostatic interactions were calculated by the particle mesh Ewald method(Darden,

418   York, & Pedersen, 1993).

419

420   **Collective variables**

421   In order to induce the conformational change from the Occ state crystal structure,

422   we used the collective variables of ΔD and Z_{Pi}. ΔD was defined as follows: $\Delta D =$

423   D_{out} - D_{in}, where D_{out} was defined as the distance between the Cα mass centers of the

424   two residue groups, 126-130 and 191-200, and D_{in} was defined as the distance between

425   the Cα mass centers of the two residue groups, 269-273 and 347-355. Z_{Pi} was used to

426   track the position of the substrate P_i, and defined as the projection of the distance vector

427   between the phosphorus atom of P_i and the reference point onto the membrane normal

428   axis. The reference point of Z_{Pi} was defined as the center of mass of the Cα atoms of the

429   P_i binding residues in the crystal structure, *i.e.*, K204, K362 and R363. In all of our

430   simulations, P_i was also restrained within a cylinder with a 10 Å radius, which was

431   aligned to the membrane normal axis and centered to the mass center of the protein,

432   using a half-harmonic potential with a spring constant of 10 kcal mol^{-1} Å^{-2}.

433   To sample the configurations for the entire transport cycle of GsGPT, we used

434   the following strategy. Generally, it is difficult to sample sufficiently in an

435   *n*-dimensional collective variable space, $(z_1, z_2, ..., z_n) = \{z\}$, due to the computational

436   cost. A practical solution is to perform the sampling only around a particular pathway

437   (or string), $s(\{z\})$. In this study, we explored the minimum free energy path of the

438   transport pathway with the string method with the swarms of trajectories (SMwST)(Pan

439   et al., 2008) in a two-dimensional collective variable space, $(\Delta D, Z_{Pi})$. To achieve the

440   conformational sampling along the refined transition path by the bias exchange

441   umbrella sampling (BEUS)(Moradi et al., 2015; Moradi & Tajkhorshid, 2014; Sugita et

442   al., 2000) (simulation sets 6 and 12), the so-called path collective variables

443   (pathCV)(Branduardi et al., 2007) were used:

$$s(\boldsymbol{z}) = \frac{1}{M-1} \frac{\sum_{i=1}^{M}(i-1)e^{-\lambda(z-z_i)^2}}{\sum_{i=1}^{M} e^{-\lambda(z-z_i)^2}} \tag{1}$$

$$\zeta(\boldsymbol{z}) = -\frac{1}{\lambda}\ln\left(\sum_{i=1}^{M} e^{-\lambda(z-z_i)^2}\right) \tag{2}$$

444   where λ is the average RMSD between images, M is the number of images, $\boldsymbol{z}$ is a vector

445   for the current values of the collective variables, and z_i is the collective variables vector

446   for the i-th image. In this paper, $\boldsymbol{z}$ was represented as a two-dimensional vector (ΔD,

447   Z_{Pi}).

448

449   **Sampling protocol**

450   At first, to generate the IF_b^{cent} and OF_b^{cent} states from the Occ state crystal

451   structure, we performed a non-equilibrium pulling simulation (or steered MD

452   simulation(Izrailev, Stepaniants, Balsera, Oono, & Schulten, 1997)) using the collective

453   variable ΔD (simulation sets 1 and 2). These simulations were performed by moving the

454   center harmonic restraint with a force constant of 10 kcal mol^{-1} Å^{-2} and using the ΔD

455   ranges from 0 to -10 Å and 0 to 10 Å, to generate IF_b^{cent} and OF_b^{cent}, respectively. The

456   final structures from simulation sets 1 and 2 were used as the initial structures of the

457   simulation sets 3 and 4, respectively, in which we induced the conformational

458   transitions to IF_b^{intra} and OF_b^{extra} using the collective variable Z_{Pi} within the ranges from

459   0 to -15 Å and 0 to 15 Å, respectively.

460   As the $IF_b^{intra} \leftrightarrow OF_b^{extra}$ transition was expected to undergo complicated

461   conformational changes, including large scale substrate translocation and a significant

462   local conformational change of the protein, we first refined the $IF_b^{intra} \leftrightarrow OF_b^{extra}$

463   transition with SMwST, prior to reconstructing the entire conformational transition, IF_a

464   $\leftrightarrow$ Occ $\leftrightarrow OF_a$. The final frames of these simulation sets were used to generate the

465   initial string of the $IF_b^{intra} \leftrightarrow OF_b^{extra}$ transition. We defined the initial $IF_b^{intra} \leftrightarrow OF_b^{extra}$

466   transition string with 40 images: 10 frames from simulation set 3 ($Z_{Pi} = 0$, -1.5, -3.0, $\cdots$,

467   -15 Å), 10 frames from simulation set 4 ($Z_{Pi} =$ 0, 1.5, 3.0, $\cdots$, 15 Å) and 20 frames

468   from simulation set 2 (ΔD = -10, -9, $\cdots$, 10 Å). This initial string was refined with

469   SMwST in the two-dimensional (ΔD, Z_{Pi}) collective variable space (simulation set 5).

470   The selected images were minimized with 10,000 steps, followed by 1 ns of

471   equilibration in the NPT ensemble with ΔD and Z_{Pi} restraints to the initial values.

472   Twenty swarms of 0.1 ps unrestrained simulation were followed by 20 ps of restrained

473   simulation, and the force constants for both ΔD and Z_{Pi} were 1.0 kcal mol^{-1} Å^{-2}. The

474   string was converged after 234 iterations. The last 20 strings were averaged and used as

475   the $IF_b^{intra} \leftrightarrow OF_b^{extra}$ transition path in simulation set 6. To perform the free energy

476   calculation with BEUS along the refined transition string, we utilized the path collective

477   variables(Branduardi et al., 2007) (pathCV(s) and pathCV(ζ)) defined by equations (1)

478   and (2) with the lambda value of 1.0. The 20 windows for the collective variable s =

479   0.025, 0.075, $\cdots$ 0.975 (increased by 0.05) were used for BEUS with a harmonic force

480   constant of 200 kcal mol^{-1} Å^{-2}. The collective variable ζ was restrained to 0.0 with a

481   harmonic force constant of 200 kcal mol^{-1} Å^{-2}. The replica exchange was attempted

482   between neighboring replicas with an interval of 2.0 ps for each unique pair. The atomic

483   coordinates were sampled every 2.0 ps.

484   Next, to induce the conformational transition to the *apo* structure, the final

— 23 —

485   frames of simulation sets 1 and 2 were further employed in the non-equilibrium pulling

486   simulation, using the collective variable Z_{Pi} (simulation sets 7 and 9) with the ranges

487   from 0 to -45 Å and from 0 to 45 Å, respectively. BEUS with 16 windows along Z_{Pi} for

488   the transition from IF_b^{cent} to IF_a was performed with window centers (and force

489   constants) of -2.0, -4.5, -6.5, -8.0, -9.0, -10.5, -12.0, -14.0, -15.5, -18.0, -20.5, -22.0,

490   -24.0, -28.0, -33.0 and -40.0 Å (0.2, 0.5, 1.0, 1.0, 1.0, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.1,

491   0.1, 0.05 and 0.01 kcal mol^{-1} Å^{-2}) (simulation set 8). BEUS with 16 windows along Z_{Pi}

492   for the transition from OF_b^{cent} to OF_a was also performed with window centers and force

493   constants of 0, 2.5, 4.0, 6.0, 8.5, 10.5, 12.5, 14.0, 16.0, 18.0, 20.0, 22.0, 25.0, 29.0, 33.0

494   and 40 Å (0.5, 1.0, 1.0, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.1, 0.1, 0.05 and 0.05

495   kcal mol^{-1} Å^{-2}) (simulation set 10).

496   The structures and free energy information obtained from simulation sets 6, 8

497   and 10 were utilized for the PHSM(Moradi et al., 2015) to generate the initial string for

498   the transition $IF_a \leftrightarrow OF_a$. The PHSM was performed as described previously(Moradi et

499   al., 2015). First, the initial string was generated by a non-parametric version of the

500   lowest free energy pathway (LFEP)(Ensing, Laio, Parrinello, & Klein, 2005; Moradi et

501   al., 2015). The following 37 metrics were used for the non-parametric LFEP and PHSM

502   algorithm: (i) Z_{Pi}, (ii) first 8 Cα-based principal components, (iii) D_{in} and D_{out}, (iv) four

503   atomic distances of K128-F192O, K128-V195O, K271-L347O and K271-V350O, (v)

504   minimum distances (D_{min}) between P_i and pore-lining polar residues: N124, K128,

505   H201, K204, R266, K271, Y339, K362 and R363, (vi) D_{min} values of internal salt

506   bridges E207-K204, E207-K362 and E207-R363, and (vii) χ_1 angles for all 10 residues

507   listed in (v) and (vi). A distance was defined in this 37-dimensional collective variable

508   space of (i) to (vii), using a diagonal metric matrix with the values 1/Å^2, 0.04/Å^2, 1/Å^2,

509   $1/\text{Å}^2$, $0.01/\text{Å}^2$, $1/\text{Å}^2$ and $0.0025/(1°)^2$ for the elements in (i) to (vii), respectively. These

510   weights were roughly chosen based on the relative variance of the collective variables.

511   The initial string generated by the non-parametric LFEP was used for the PHSM

512   algorithm with 400 image centers. The tube thickness $\varepsilon = 0.5$ was used in the PHSM

513   algorithm, since it gives the smoothest pathway as compared with other results based on

514   $\varepsilon = 0.1, 0.2, \dots 1.0$.

515   The string of the $IF_a \leftrightarrow OF_a$ transition generated by the PHSM algorithm was

516   further refined with SMwST (simulation set 11) in the two-dimensional collective

517   variable space (ΔD, Z_{Pi}). We selected 40 images from the string refined with the PHSM

518   algorithm at even intervals. The selected images from PHSM were minimized and

519   equilibrated in the same manner as the simulation set 5. In this analysis, 20 swarms of 5

520   ps unrestrained simulation were performed, followed by 10 ps of restrained simulation

521   with force constants of 10.0 kcal/mol for each collective variable. The string was

522   converged after 50 iterations. The final 10 strings were averaged, and used for the final

523   BEUS sampling.

524   The final BEUS sampling along the refined $IF_a \leftrightarrow OF_a$ transition string

525   (simulation set 12) was performed using the pathCV(s) and pathCV(ζ) with $\lambda = 0.1$

526   (equations (1), (2)). The 40 images corresponding to the image points of the averaged

527   string were selected and equilibrated for 0.1 ns with the collective variables (ΔD, Z_{Pi})

528   restrained. In total, 40 windows, s = 0.015, 0.035, … 0.99 (increased by 0.025) for the

529   pathCV(s) were used for BEUS, with a force constant of 500 kcal mol^{-1}. The pathCV(ζ)

530   is restrained to the region $\zeta < 1.0$ with half-harmonic restraints with a force constant of

531   100 kcal mol^{-1}. The exchange was attempted between the neighboring replicas with an

532   interval of 10 ps for each unique pair. The sampling was performed for 50 ns, and the

— 25 —

533   coordinates for all atoms were sampled every 2 ps, to obtain 1,000,000 structures (= 40

534   replicas x 50 ns / 2 ps) in total.

535

536   **Data analysis**

537   Data analysis was performed using MDAnalysis(Michaud-Agrawal, Denning,

538   Woolf, & Beckstein, 2011), mdtraj(McGibbon et al., 2015), and HOLE(Smart et al.,

539   1996), along with in-house python and C++ codes. All molecular graphics were

540   illustrated with CueMol2 (http://www.cuemol.org/). Plot graphics were generated with

541   ggplot2(Wickham, 2009), seaborn(Waskom et al., 2016) and matplotlib(Hunter, 2007).

542   To calculate the potential of mean force (PMF) from the BEUS simulations

543   along the path collective variable s in Figure 1f, the C implementation of the WHAM

544   algorithm(Kumar et al., 1992) (Grossfield, Alan, "WHAM: the weighted histogram

545   analysis method", version 2.0.9 http://membrane.urmc.rochester.edu/content/wham) was

546   used. To perform the *post-hoc* string method analysis, the weights for each structure w^t

547   were required(Moradi et al., 2015). They were calculated by the following generalized

548   WHAM equation(Bartels, 2000; Moradi et al., 2015):

$$w^t = \left(\sum_j T_j \, exp\left[-\beta(U_j(\zeta^t) - F_j) \right] \right)^{-1} \qquad (3)$$

$$exp[-\beta F_i] = \sum_t w^t exp[-\beta U_i(\zeta^t)] \qquad (4)$$

549   where $\beta = 1/k_B T$ is the inverse temperature, T_j is the number of samples collected for

550   image j, F_j is the PMF for the j-th umbrella window, ζ^t is the collective variable at

551   time t, and $U_j(x) = 1/2 \, k_j(x - x_j^{center})^2$ is the j-th umbrella potential used in the

552   BEUS simulation. The projection of the PMF from one collective variable space to

553   another one, ξ, was performed with the following equation(Moradi et al., 2015):

$$G(\xi) = -\frac{1}{\beta} \ln \left(\sum_t w^t K(\xi(x^t) - \xi) \right) \qquad (5)$$

554   where K(.) is a kernel function. In this study, the RBF kernel was used.

555   The principal component analysis (PCA) was performed on the Cα atoms of

556   GsGPT. The coordinates of the trajectory were superimposed on the averaged structure

557   of simulation set 12. The covariance matrix of the atomic fluctuations was calculated,

558   and its eigenvalues and eigenvectors were determined. The calculation of the number of

559   hydrogen bonds in Figure 3 was performed by the following equation:

$$\sum_{i \in group} \frac{1 - \left(d_i / d_0 \right)^n}{1 - \left(d_i / d_0 \right)^m} \qquad (6)$$

560   where d_i is the distance between two hydrogen-bonding atoms, d_0 is a cutoff distance,

561   and n and m are integers that determine the behavior of the switching function. The sum

562   is taken over all of the hydrogen bonding atom pairs. In this study, $d_0 = 3.3$ Å, $n = 6$ and

563    $m = 8$ were used.

564

565   **Transport assay**

566   The liposome assays were performed as previously described(Lee et al., 2017).

567   In brief, the liposomes reconstituted with yeast membranes expressing WT or mutant

568   GsGPT were preloaded with 30 mM NaH$_2$PO$_4$, and were mixed with an equal volume of

569   extra-liposomal solution containing 1 mM [^{32}P]-NaH$_2$PO$_4$ (0.1 mCi ml^{-1}). The reaction

570   was terminated by passing the liposomes through AG-1 X8 resin. The total amounts of

571   incorporated P_i were measured at 30 min. The control experiment was performed with

572   membranes from yeast cells harboring the empty vector. The mean count value of the

573   control experiment was subtracted from each count value for each experiment set, and

574   the values shown in the figures were normalized with the mean value of the WT. The

575   expression of each mutant was verified by a western-blot analysis, using an anti-His-tag

576   polyclonal antibody (code PM032; MBL).

577

578

References

Bartels, C. (2000). Analyzing biased Monte Carlo and molecular dynamics simulations. *Chemical Physics Letters*, *331*(5–6), 446–454. http://doi.org/10.1016/S0009-2614(00)01215-X

Best, R. B., Zhu, X., Shim, J., Lopes, P. E. M., Mittal, J., Feig, M., & MacKerell, A. D. (2012). Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ 1 and χ 2 Dihedral Angles. *Journal of Chemical Theory and Computation*, *8*(9), 3257–3273. http://doi.org/10.1021/ct300400x

Branduardi, D., Gervasio, F. L., & Parrinello, M. (2007). From A to B in free energy space. *The Journal of Chemical Physics*, *126*(5), 54103. http://doi.org/10.1063/1.2432340

Brooks, C. F., Johnsen, H., van Dooren, G. G., Muthalagi, M., Lin, S. S., Bohne, W., … Striepen, B. (2010). The Toxoplasma Apicoplast Phosphate Translocator Links Cytosolic and Apicoplast Metabolism and Is Essential for Parasite Survival. *Cell Host and Microbe*, *7*(1), 62–73. http://doi.org/10.1016/j.chom.2009.12.002

Chen, Y.-J., Pornillos, O., Lieu, S., Ma, C., Chen, A. P., & Chang, G. (2007). X-ray structure of EmrE supports dual topology model. *Proceedings of the National Academy of Sciences*, *104*(48), 18999–19004. http://doi.org/10.1073/pnas.0709387104

Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An N log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics*, *98*(12), 10089. http://doi.org/10.1063/1.464397

Drew, D., & Boudker, O. (2016). Shared Molecular Mechanisms of Membrane Transporters. *Annual Review of Biochemistry*, *85*, 543–72. http://doi.org/10.1146/annurev-biochem-060815-014520

Eicks, M., Maurino, V., Knappe, S., Flügge, U.-I., & Fischer, K. (2002). The plastidic pentose phosphate translocator represents a link between the cytosolic and the plastidic pentose phosphate pathways in plants. *Plant Physiology*, *128*(2), 512–22. http://doi.org/10.1104/pp.010576

Ensing, B., Laio, A., Parrinello, M., & Klein, M. L. (2005). A recipe for the computation of the free energy barrier and the lowest free energy path of concerted reactions. *Journal of Physical Chemistry B*, *109*(14), 6676–6687. http://doi.org/10.1021/jp045571i

Feller, S. E., Zhang, Y., Pastor, R. W., & Brooks, B. R. (1995). Constant pressure

614   molecular dynamics simulation: The Langevin piston method. *The Journal of*

615   *Chemical Physics*, *103*(11), 4613. http://doi.org/10.1063/1.470648

616 Fiorin, G., Klein, M. L., & Hénin, J. (2013). Using collective variables to drive molecular

617   dynamics simulations. *Molecular Physics*, *111*(22–23), 3345–3362.

618   http://doi.org/10.1080/00268976.2013.813594

619 Fischer, K., Kammerer, B., Gutensohn, M., Arbinger, B., Weber, A., Häusler, R. E., &

620   Flügge, U. I. (1997). A new class of plastidic phosphate translocators: a putative link

621   between primary and secondary metabolism by the phosphoenolpyruvate/phosphate

622   antiporter. *The Plant Cell*, *9*(3), 453–62. http://doi.org/10.1105/tpc.9.3.453

623 Fliege, R., Flügge, U. I., Werdan, K., & Heldt, H. W. (1978). Specific transport of

624   inorganic phosphate, 3-phosphoglycerate and triosephosphates across the inner

625   membrane of the envelope in spinach chloroplasts. *BBA - Bioenergetics*, *502*(2),

626   232–247. http://doi.org/10.1016/0005-2728(78)90045-2

627 Flügge, U.-I., Weber, A., Fischer, K., Lottspeich, F., Eckerskorn, C., Waegemann, K., &

628   Soll, J. (1991). The major chloroplast envelope polypeptide is the phosphate

629   translocator and not the protein import receptor. *Nature*, *353*(6342), 364–367.

630   http://doi.org/10.1038/353364a0

631 Flügge, U. I. I., Fischer, K., Gross, A., Sebald, W., Lottspeich, F., & Eckerskorn, C.

632   (1989). The triose phosphate-3-phosphoglycerate-phosphate translocator from

633   spinach chloroplasts: nucleotide sequence of a full-length cDNA clone and import of

634   the in vitro synthesized precursor protein into chloroplasts. *The EMBO Journal*, *8*(1),

635   39–46. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2714257

636 Hattenbach, A., Muller-Rober, B., Nast, G., & Heineke, D. (1997). Antisense Repression

637   of Both ADP-Glucose Pyrophosphorylase and Triose Phosphate Translocator

638   Modifies Carbohydrate Partitioning in Potato Leaves. *Plant Physiology*, *115*(2),

639   471–475. http://doi.org/10.1104/pp.115.2.471

640 Häusler, R. E., Schlieben, N. H., & Flügge, U. I. (2000). Control of carbon partitioning

641   and photosynthesis by the triose phosphate/phosphate translocator in transgenic

642   tobacco plants (Nicotiana tabacum). II. Assessment of control coefficients of the

643   triose phosphate/phosphate translocator. *Planta*, *210*(3), 383–90.

644   http://doi.org/10.1007/PL00008146

645 Häusler, R. E., Schlieben, N. H., Nicolay, P., Fischer, K. L., Fischer, K. L., & Flügge, U.

646   I. (2000). Control of carbon partitioning and photosynthesis by the triose

647   phosphate/phosphate translocator in transgenic tobacco plants (Nicotiana tabacum

648   L.). I. Comparative physiological analysis of tobacco plants with antisense

649   repression and overexpression of t. *Planta*, *210*(3), 371–82.

650   http://doi.org/10.1007/PL00008145

651   Heineke, D., Kruse, A., Flügge, U.-I. I., Frommer, W. B., Riesmeier, J. W., Willmitzer, L.,
652   & Heldt, H. W. (1994). Effect of antisense repression of the chloroplast
653   triose-phosphate translocator on photosynthetic metabolism in transgenic potato
654   plants. *Planta*, *193*(2), 174–180. http://doi.org/10.1007/BF00192527

655   Henzler-Wildman, K. (2012). Analyzing conformational changes in the transport cycle of
656   EmrE. *Current Opinion in Structural Biology*, *22*(1), 38–43.
657   http://doi.org/10.1016/j.sbi.2011.10.004

658   Huang, Y., Lemieux, M. J., Song, J., Auer, M., & Wang, D.-N. (2003). Structure and
659   mechanism of the glycerol-3-phosphate transporter from Escherichia coli. *Science*
660   *(New York, N.Y.)*, *301*(5633), 616–20. http://doi.org/10.1126/science.1087619

661   Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics.
662   *Journal of Molecular Graphics*, *14*(1), 33–38.
663   http://doi.org/10.1016/0263-7855(96)00018-5

664   Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science &*
665   *Engineering*, *9*(3), 90–95. http://doi.org/10.1109/MCSE.2007.55

666   Izrailev, S., Stepaniants, S., Balsera, M., Oono, Y., & Schulten, K. (1997). Molecular
667   dynamics study of unbinding of the avidin-biotin complex. *Biophysical Journal*,
668   *72*(4), 1568–1581. http://doi.org/10.1016/S0006-3495(97)78804-0

669   Jack, D. L., Yang, N. M., & H. Saier, M. (2001). The drug/metabolite transporter
670   superfamily. *European Journal of Biochemistry*, *268*(13), 3620–3639.
671   http://doi.org/10.1046/j.1432-1327.2001.02265.x

672   Jardetzky, O. (1966). Simple allosteric model for membrane pumps. *Nature*, *211*(5052),
673   969–970. http://doi.org/10.1038/211969a0

674   Kammerer, B., Fischer, K., Hilpert, B., Schubert, S., Gutensohn, M., Weber, A., &
675   Flügge, U. I. (1998). Molecular characterization of a carbon transporter in plastids
676   from heterotrophic tissues: the glucose 6-phosphate/phosphate antiporter. *The Plant*
677   *Cell*, *10*(1), 105–17. http://doi.org/10.2307/3870632

678   Khalili-Araghi, F., Gumbart, J., Wen, P. C., Sotomayor, M., Tajkhorshid, E., & Schulten,
679   K. (2009). Molecular dynamics simulations of membrane channels and transporters.
680   *Current Opinion in Structural Biology*, *19*(2), 128–137.
681   http://doi.org/10.1016/j.sbi.2009.02.011

682   Klauda, J. B., Venable, R. M., Freites, J. A., O'Connor, J. W., Tobias, D. J.,
683   Mondragon-Ramirez, C., … Pastor, R. W. (2010). Update of the CHARMM
684   All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *The*
685   *Journal of Physical Chemistry B*, *114*(23), 7830–7843.

686   http://doi.org/10.1021/jp101759q

687   Knappe, S., Flügge, U.-I., & Fischer, K. (2003). Analysis of the plastidic phosphate
688   translocator gene family in Arabidopsis and identification of new phosphate
689   translocator-homologous transporters, classified by their putative substrate-binding
690   site. *Plant Physiology*, *131*(3), 1178–1190. http://doi.org/10.1104/pp.016519

691   Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., & Kollman, P. A. (1992).
692   THE weighted histogram analysis method for free-energy calculations on
693   biomolecules. I. The method. *Journal of Computational Chemistry*, *13*(8), 1011–
694   1021. http://doi.org/10.1002/jcc.540130812

695   Law, C. J., Almqvist, J., Bernstein, A., Goetz, R. M., Huang, Y., Soudant, C., … Wang, D.
696   N. (2008). Salt-bridge Dynamics Control Substrate-induced Conformational
697   Change in the Membrane Transporter GlpT. *Journal of Molecular Biology*, *378*(4),
698   826–837. http://doi.org/10.1016/j.jmb.2008.03.029

699   Law, C. J., Maloney, P. C., & Wang, D.-N. (2008). Ins and Outs of Major Facilitator
700   Superfamily Antiporters. *Annual Review of Microbiology*, *62*(1), 289–305.
701   http://doi.org/10.1146/annurev.micro.61.080706.093329

702   Lee, Y., Nishizawa, T., Takemoto, M., Kumazaki, K., Yamashita, K., Hirata, K., …
703   Nureki, O. (2017). Structure of the triose-phosphate/phosphate translocator reveals
704   the basis of substrate specificity. *Nature Plants*.
705   http://doi.org/10.1038/s41477-017-0022-8

706   Linka, M., Jamai, A., & Weber, A. P. M. (2008). Functional characterization of the
707   plastidic phosphate translocator gene family from the thermo-acidophilic red alga
708   Galdieria sulphuraria reveals specific adaptations of primary carbon partitioning in
709   green plants and red algae. *Plant Physiology*, *148*(3), 1487–96.
710   http://doi.org/10.1104/pp.108.129478

711   McGibbon, R. T., Beauchamp, K. A., Harrigan, M. P., Klein, C., Swails, J. M.,
712   Hernández, C. X., … Pande, V. S. (2015). MDTraj: A Modern Open Library for the
713   Analysis of Molecular Dynamics Trajectories. *Biophysical Journal*, *109*(8), 1528–
714   1532. http://doi.org/10.1016/j.bpj.2015.08.015

715   Michaud-Agrawal, N., Denning, E. J., Woolf, T. B., & Beckstein, O. (2011).
716   MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *Journal*
717   *of Computational Chemistry*, *32*(10), 2319–27. http://doi.org/10.1002/jcc.21787

718   Moradi, M., Enkavi, G., & Tajkhorshid, E. (2015). Atomic-level characterization of
719   transport cycle thermodynamics in the glycerol-3-phosphate:phosphate antiporter.
720   *Nature Communications*, *6*, 8393. http://doi.org/10.1038/ncomms9393

721   Moradi, M., & Tajkhorshid, E. (2014). Computational Recipe for Efficient Description of

722   Large-Scale Conformational Changes in Biomolecular Systems. *Journal of*
723   *Chemical Theory and Computation*, *10*(7), 2866–2880.
724   http://doi.org/10.1021/ct5002285

725 Pan, A. C., Sezer, D., & Roux, B. (2008). Finding transition pathways using the string
726   method with swarms of trajectories. *Journal of Physical Chemistry B*, *112*(11),
727   3432–3440. http://doi.org/10.1021/jp0777059

728 Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., … Schulten,
729   K. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational*
730   *Chemistry*, *26*(16), 1781–802. http://doi.org/10.1002/jcc.20289

731 Quistgaard, E. M., Löw, C., Guettou, F., & Nordlund, P. (2016). Understanding transport
732   by the major facilitator superfamily (MFS): structures pave the way. *Nature*
733   *Publishing Group*, 1–10. http://doi.org/10.1038/nrm.2015.25

734 Riesmeier, J. W., Flügge, U. I., Schulz, B., Heineke, D., Heldt, H. W., Willmitzer, L., &
735   Frommer, W. B. (1993). Antisense repression of the chloroplast triose phosphate
736   translocator affects carbon partitioning in transgenic potato plants. *Proceedings of*
737   *the National Academy of Sciences of the United States of America*, *90*(13), 6160–4.
738   http://doi.org/10.1073/pnas.90.13.6160

739 Schneider, A., Häusler, R. E., Kolukisaoglu, U., Kunze, R., van der Graaff, E., Schwacke,
740   R., … Flügge, U.-I. (2002). An Arabidopsis thaliana knock-out mutant of the
741   chloroplast triose phosphate/phosphate translocator is severely compromised only
742   when starch synthesis, but not starch mobilisation is abolished. *The Plant Journal :*
743   *For Cell and Molecular Biology*, *32*(5), 685–99.
744   http://doi.org/10.1046/j.1365-313X.2002.01460.x

745 Shaikh, S. a, Li, J., Enkavi, G., Wen, P.-C., Huang, Z., & Tajkhorshid, E. (2013).
746   Visualizing functional motions of membrane transporters with molecular dynamics
747   simulations. *Biochemistry*, *52*(4), 569–87. http://doi.org/10.1021/bi301086x

748 Smart, O. S., Neduvelil, J. G., Wang, X., Wallace, B. A., & Sansom, M. S. P. (1996).
749   HOLE: A program for the analysis of the pore dimensions of ion channel structural
750   models. *Journal of Molecular Graphics*, *14*(6), 354–360.
751   http://doi.org/10.1016/S0263-7855(97)00009-X

752 Sugita, Y., Kitao, A., & Okamoto, Y. (2000). Multidimensional replica-exchange method
753   for free-energy calculations. *The Journal of Chemical Physics*, *113*(15), 6042.
754   http://doi.org/10.1063/1.1308516

755 Torrie, G. M. M., & Valleau, J. P. P. (1977). Nonphysical sampling distributions in
756   Monte Carlo free-energy estimation: Umbrella sampling. *Journal of Computational*
757   *Physics*, *23*(2), 187–199. http://doi.org/10.1016/0021-9991(77)90121-8

758 Tsuchiya, H., Doki, S., Takemoto, M., Ikuta, T., Higuchi, T., Fukui, K., … Nureki, O.
759   (2016). Structural basis for amino acid export by DMT superfamily transporter
760   YddG. *Nature*, *534*(7607), 417–420. http://doi.org/10.1038/nature17991
761 Waskom, M., Botvinnik, O., drewokane, Hobson, P., David, Halchenko, Y., … Lee, A.
762   (2016). seaborn: v0.7.1 (June 2016). http://doi.org/10.5281/zenodo.54844
763 Weber, A. P. M., & Linka, N. (2011). Connecting the plastid: transporters of the plastid
764   envelope and their role in linking plastidial with cytosolic metabolism. *Annual*
765   *Review of Plant Biology*, *62*, 53–77.
766   http://doi.org/10.1146/annurev-arplant-042110-103903
767 Weber, A. P. M., Schwacke, R., & Flügge, U.-I. (2005). Solute Transporters of the Plastid
768   Envelope Membrane. *Annual Review of Plant Biology*, *56*(1), 133–164.
769   http://doi.org/10.1146/annurev.arplant.56.032604.144228
770 Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag
771   New York.
772

Acknowledgements

Author contributions

M.T. designed the research, performed the molecular dynamics simulation, and analyzed the data. Y.L. performed the liposome transport assay. M.T., R.I. and O.N. wrote the manuscript. R.I. and O.N. directed and supervised all of the research.
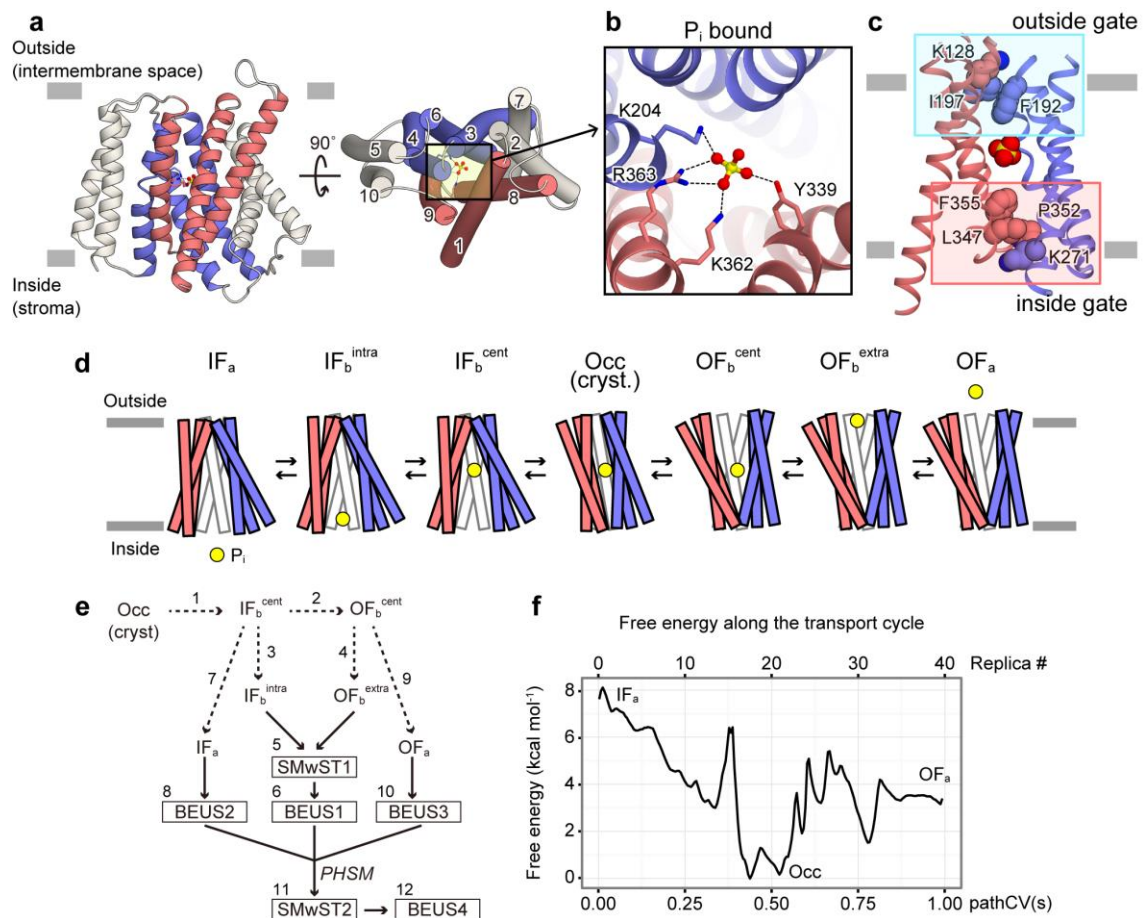
Competing financial interests

The authors declare no competing financial interests.

Figures

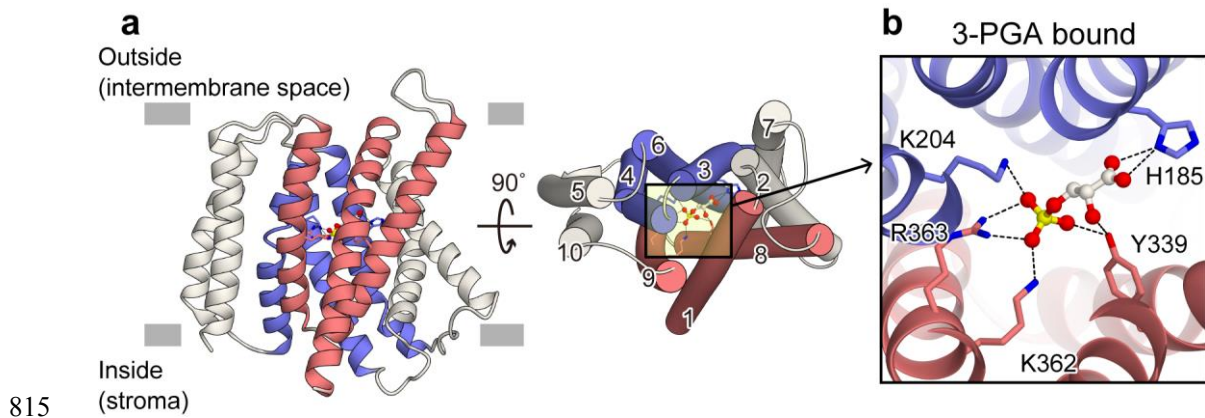**Figure 1 | The crystal structure of GsGPT and simulation overview**

(a) The crystal structure of GsGPT in the P_i-bound state (PDB ID: 5Y78), viewed from the plane of the membrane (left) and from the intermembrane space (right). The numbers in the right panel represent the numbering of the TM helices. Hereafter, TM1, 8 and 9 are colored red, TM3, 4 and 6 are colored blue, and the others are colored white. (b) Close-up view of the central binding site in the P_i-bound state. Key residues involved in substrate binding are shown in stick models. Dotted lines represent polar interactions. (c) Inside and outside gates of GsGPT. The inside and outside gates are

— 36 —

803   highlighted within magenta and cyan rectangles, respectively. The substrate and

804   gate-forming residues are shown in CPK models. TM2, 5, 7 and 10 are not shown. (d)

805   Schematic representation of states, IF_a (IF-*apo*), IF_b^{intra} (IF-bound at intracellular gate),

806   IF_b^{cent} (IF-bound at central binding site), Occ, OF_b^{cent} (OF-bound at central binding site),

807   OF_b^{extra} (OF-bound at extracellular gate) and IF_a (OF-*apo*), involved in the currently

808   studied transport cycle. P_i is represented by the yellow circle. The TM segments are

809   colored in the same manner as in panel a. TM2 and 7 are not shown. (e) Graphical

810   representation of the iterative scheme used for designing the simulations shown in Table

811   1. The number corresponds to the simulation numbering in Table 1. (f) Free energy

812   profile along the simulated transport cycle shown in c, based on the final simulation

813   (simulation set 12). pathCV(s) represents path collective variable s (equation (1)).

814

815

816 **Figure 1—figure supplement 1 | Crystal structure of GsGPT with 3-PGA**

817   (a) The crystal structure of GsGPT in the 3-phosphoglyceric acid

818   (3-PGA)-bound state (PDB ID: 5Y79), viewed from the plane of the membrane (left)

819   and from the intermembrane space (right). The numbers in the right panel represent the

820   numbering of the TM helices. TM1, 8 and 9 are colored red, TM3, 4 and 6 are colored

821   blue, and the others are colored white. (b) Close-up view of the central binding site in

822   the 3-PGA-bound state. Key residues involved in substrate binding are shown in stick

823   models. Dotted lines represent polar interactions.

824

825

826
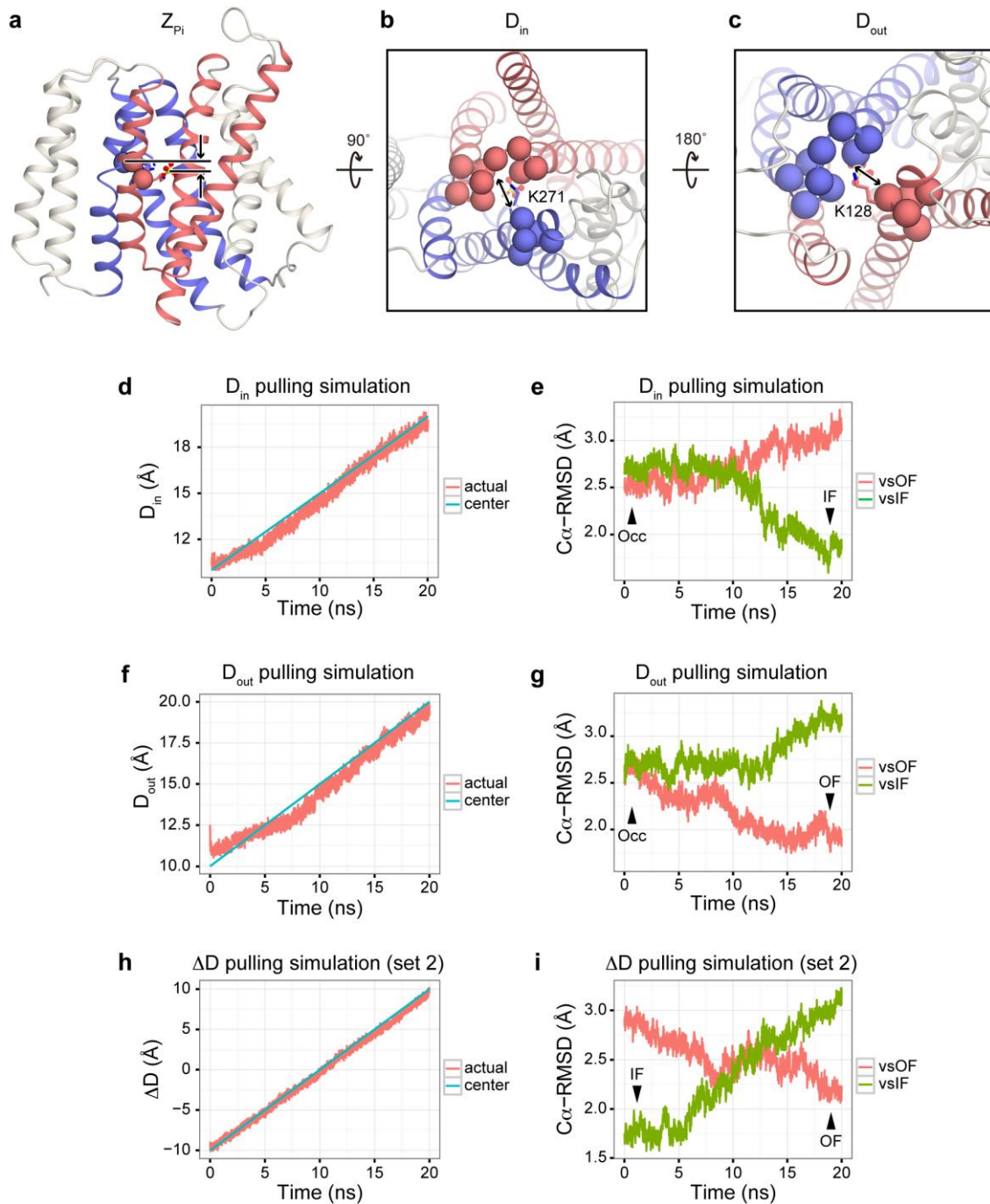
Figure 1—figure supplement 2 | The definitions of collective variables and the steered MD simulations

(a-c) Definitions of collective variables. (a) Definition of Z_{Pi}. P_i is shown with a ball and stick model, and the Cα atoms of Lys204, Lys362 and Arg363 used for the

831   reference points are shown with spheres. (b, c) Definitions of D_{in} and D_{out}. The Cα

832   atoms used for the calculation are shown as spheres. (d, f, h) Time courses of the

833   collective variables, (d) D_{in}, (f) D_{out} and (h) ΔD, for the non-equilibrium pulling (steered

834   MD) simulation of each variable. The blue line represents the center of the harmonic

835   potential, and the red line represents the actual value of the collective variable. (e, g, i)

836   Time courses of the root-mean-square deviation (RMSD) values of all Cα atoms relative

837   to the OF (red) and IF (green) conformations. IF, Occ and OF represent the initial,

838   occluded and final conformations of the protein, respectively.

839

840

841
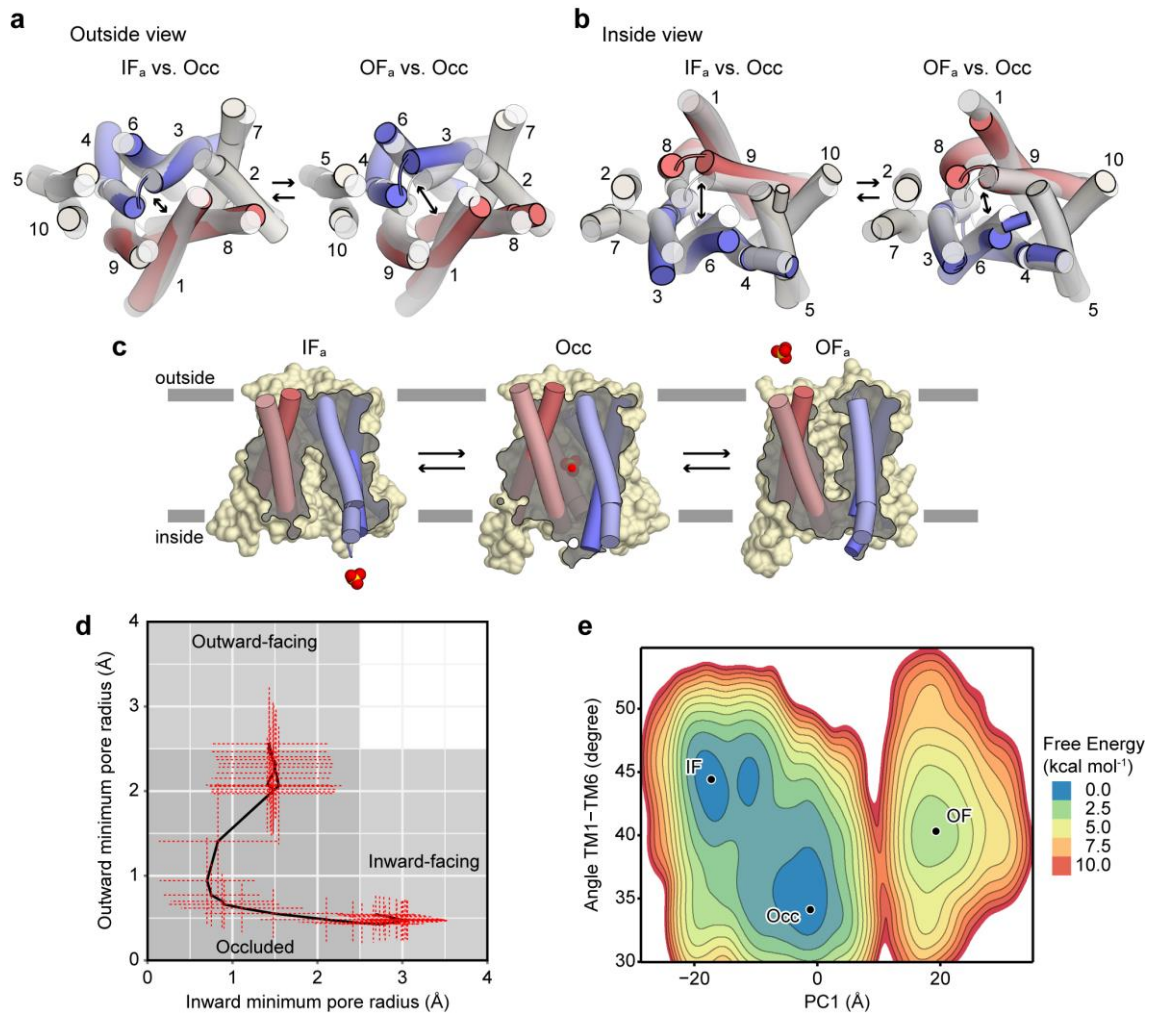
Figure 2 | Overall conformational change

843   (a) Outside and (b) inside views of representative structures in the OF and IF

844   states, shown in cylinder representations. Each structure was superimposed onto the Occ

845   state structure, shown with a transparent cylinder. Bundle1 and bundle2 are colored red

846   and blue, respectively, and the other helices are colored white. The number represents

847   the numbering of the TM helices. (c) Cross-sections of surface representations for the

848   representative structures from the IF, Occ and OF states. P_i is shown in a CPK model.

849   (d) Pore radii at the outside and inside gates. The mean values of the outside and inside

850   radii in each replica were plotted. The error bars represent the standard deviation in each

— 41 —

851   replica. (e) Free energy landscape in terms of (PC1, TM1-TM6 angle) space. PC1

852   represents the first principal component of all Cα atoms, and the inter-helical angle

853   between TM1 and TM6 was calculated using the roll axis of the helix (obtained from

854   the principal axis component analysis for all Cα atoms).

855

856

857
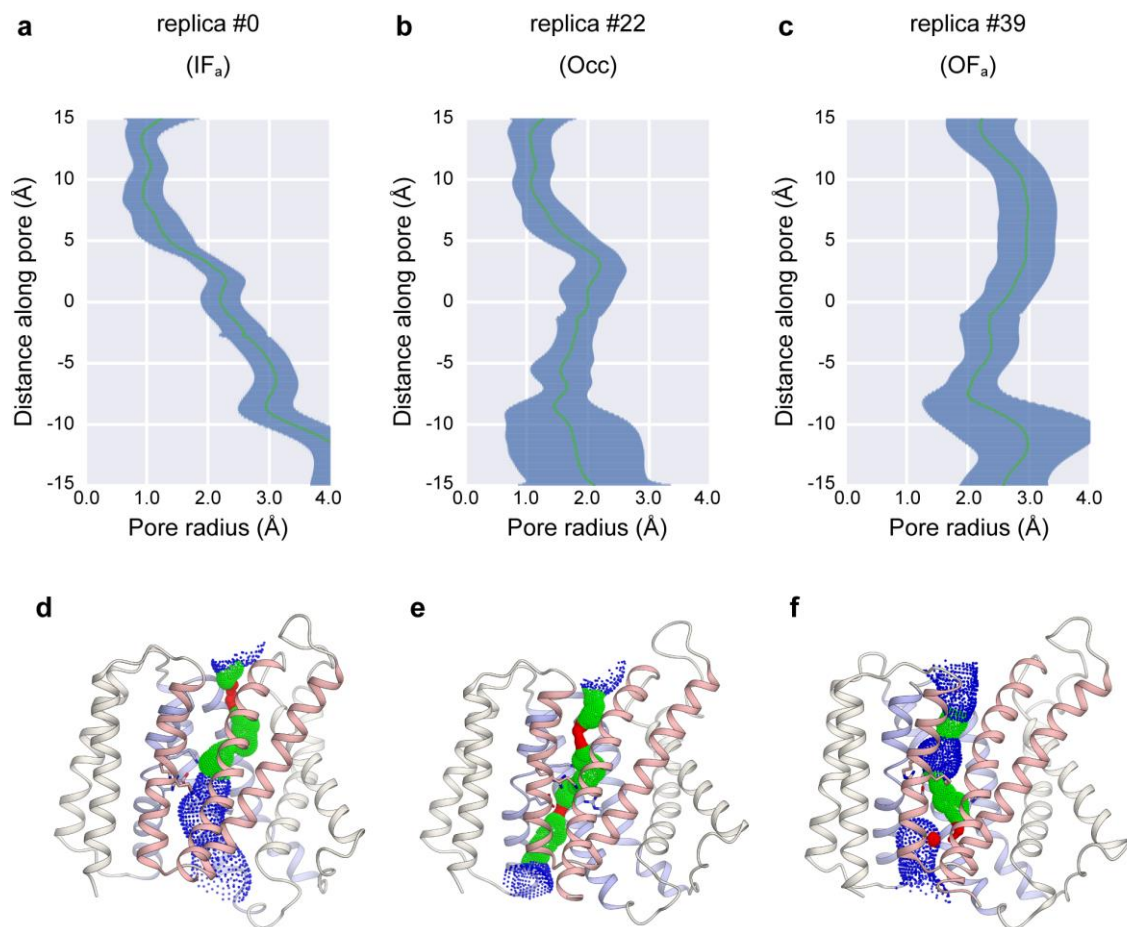
Figure 2—figure supplement 1 | Representative results for pore radius calculation by HOLE

(a-c) Minimum radius along the path through the central binding site for (a) replica#0 (IF), (b) replica#22 (Occ) and (c) replica#39 (OF). The green line represents the mean value and the bl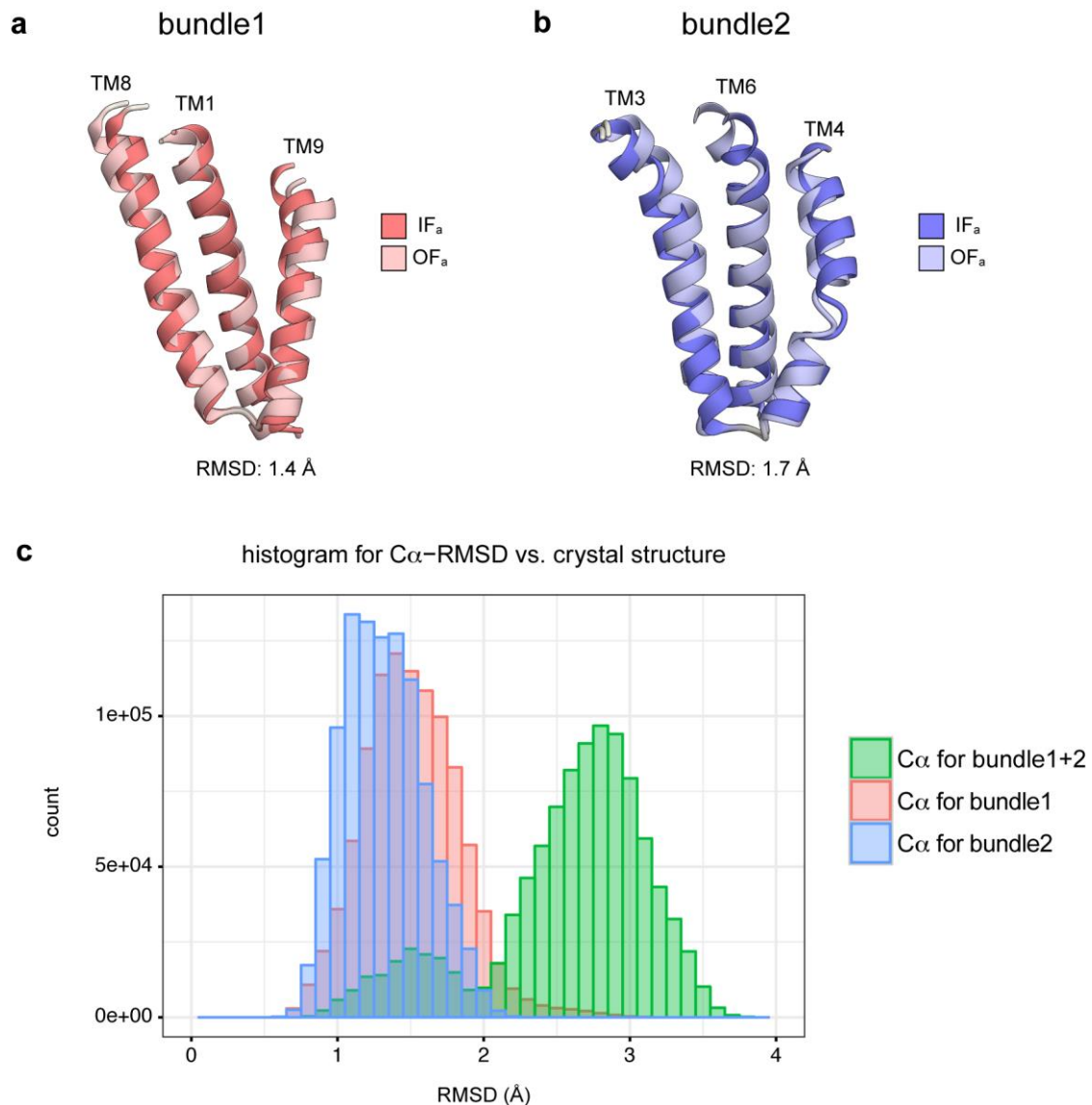ue line represents the mean value and standard deviation for each distance along the pore. (d-f) Representative structures of the (d) IF, (e) Occ and (f) OF states. The red region represents $r < 0.6$ Å, the green region represents $0.6 < r < 1.2$ Å and the blue region represents $r > 1.2$ Å.

866

867

a bundle1

TM8 TM1 TM9

IF$_a$
OF$_a$

RMSD: 1.4 Å

b bundle2

TM3 TM6 TM4

IF$_a$
OF$_a$

RMSD: 1.7 Å

c histogram for Cα−RMSD vs. crystal structure

Cα for bundle1+2
Cα for bundle1
Cα for bundle2

count

RMSD (Å)

868

Figure 2—figure supplement 2 I Intra-domain movement of the core

domain

871   (a, b) Representative conformations of (a) bundle1 and (b) bundle2 from the IF

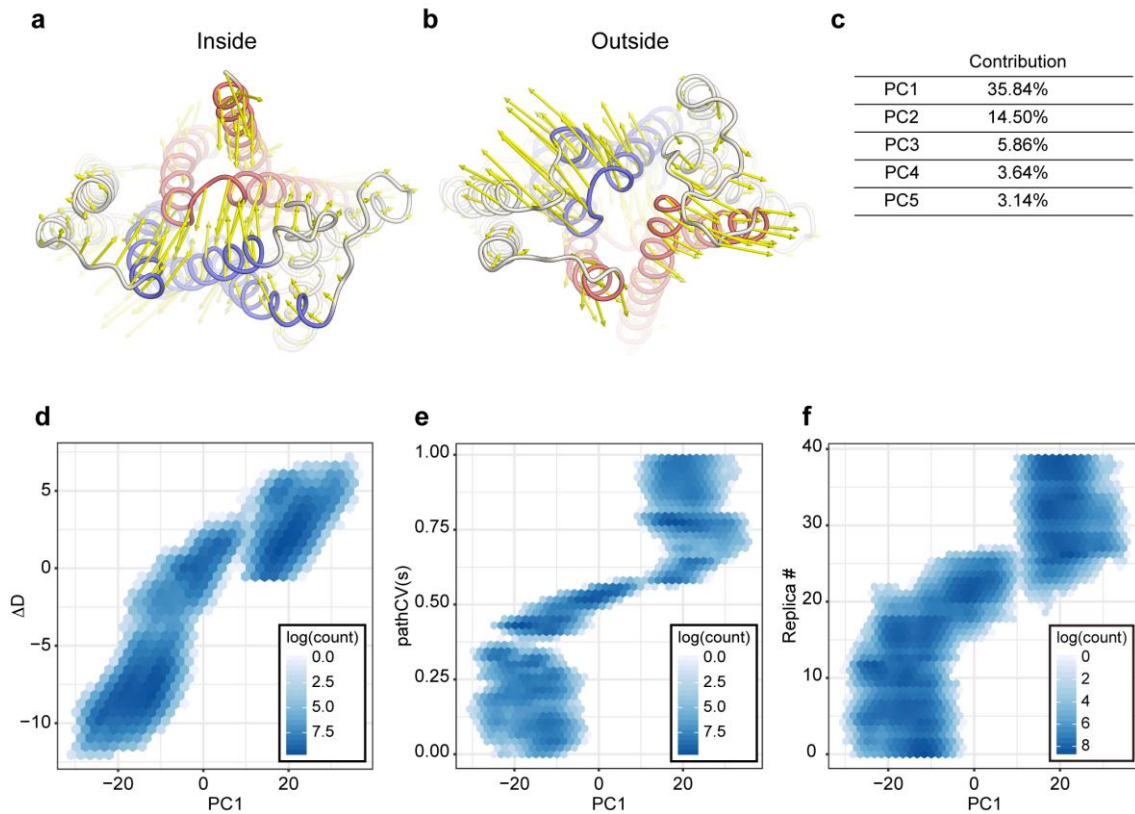872   and OF conformations. The root-mean-square deviation (RMSD) values of the Cα

873   atoms for each structure are shown below. (c) The histogram of RMSD values of Cα

874   atoms in bundles1+2 (green), bundle1 only (red) and bundle2 only (blue) values relative

875   to the crystal structure. The RMSD values were calculated for the α-helix region of each

— 44 —

876   bundle.

877

878

Figure 2—figure supplement 3 | The visualization of PCA results and correlations with other variables

(a, b) Vector representation of the eigenvector of the first principal component (PC1) of GsGPT movement, calculated by principal component analysis (PCA). (c) The contributions of PC1-5 for the overall deviations. (d-f) Correlations of the PC1 *vs.* (d) ΔD, (e) the number of replicas and (f) the pathCV(s) shown in a 2D histogram. The count value is shown in a log scale.

886

887
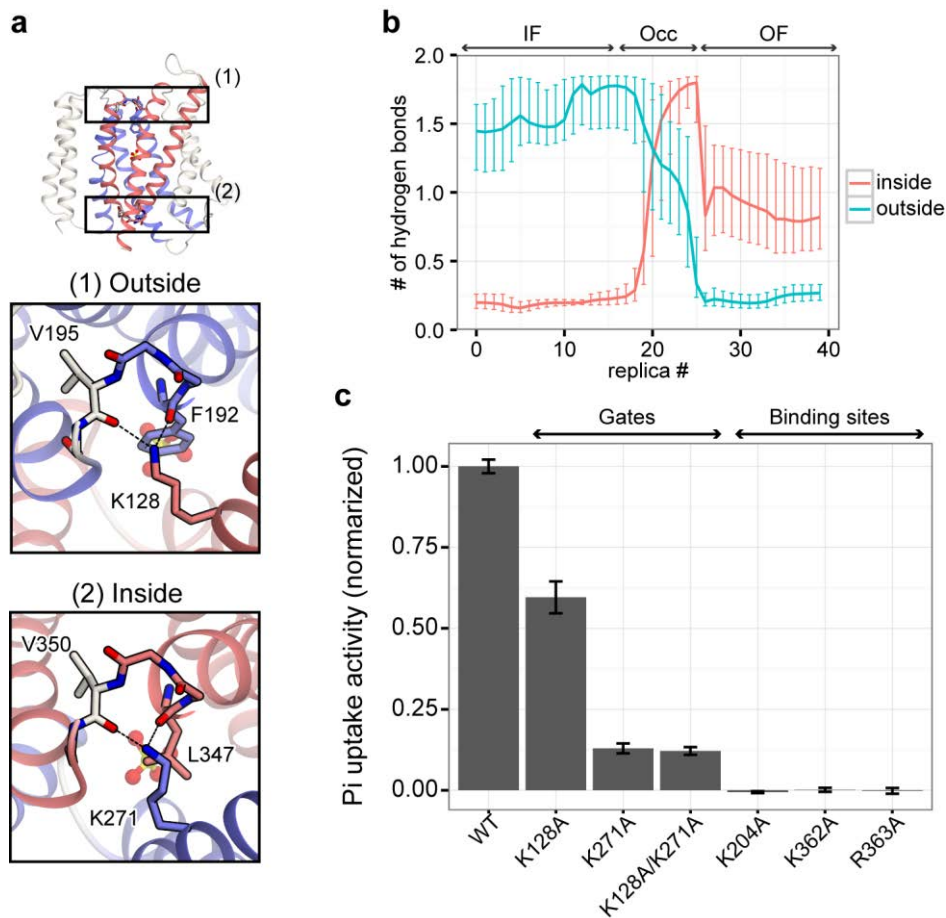
Figure 3 | Hydrogen bonds at the inside and outside gates

(a) Overall structure of GsGPT (upper) and close-up views of the outside (center) and inside (lower) gates. The dotted line represents a hydrogen bond between a Lys residue and a main chain carbonyl atom. (b) The number of hydrogen bonds in the inside and outside gates (magenta and cyan, respectively), shown in panel a. The median value in each replica is plotted. Error bars represent the interquartile range (IQR). (c) Liposome-based mutational analysis. The levels of $[^{32}P]$-P_i uptake by GsGPT mutants were normalized with the mean value of the wild-type. The error bars represent s.e.m. (n=3-9; see Figure 3—Source Data1). The values of K204A, K362A and R363A are from our previous study(Lee et al., 2017).
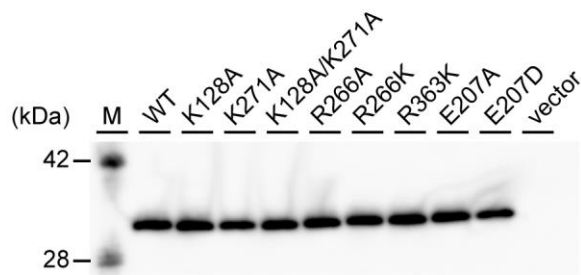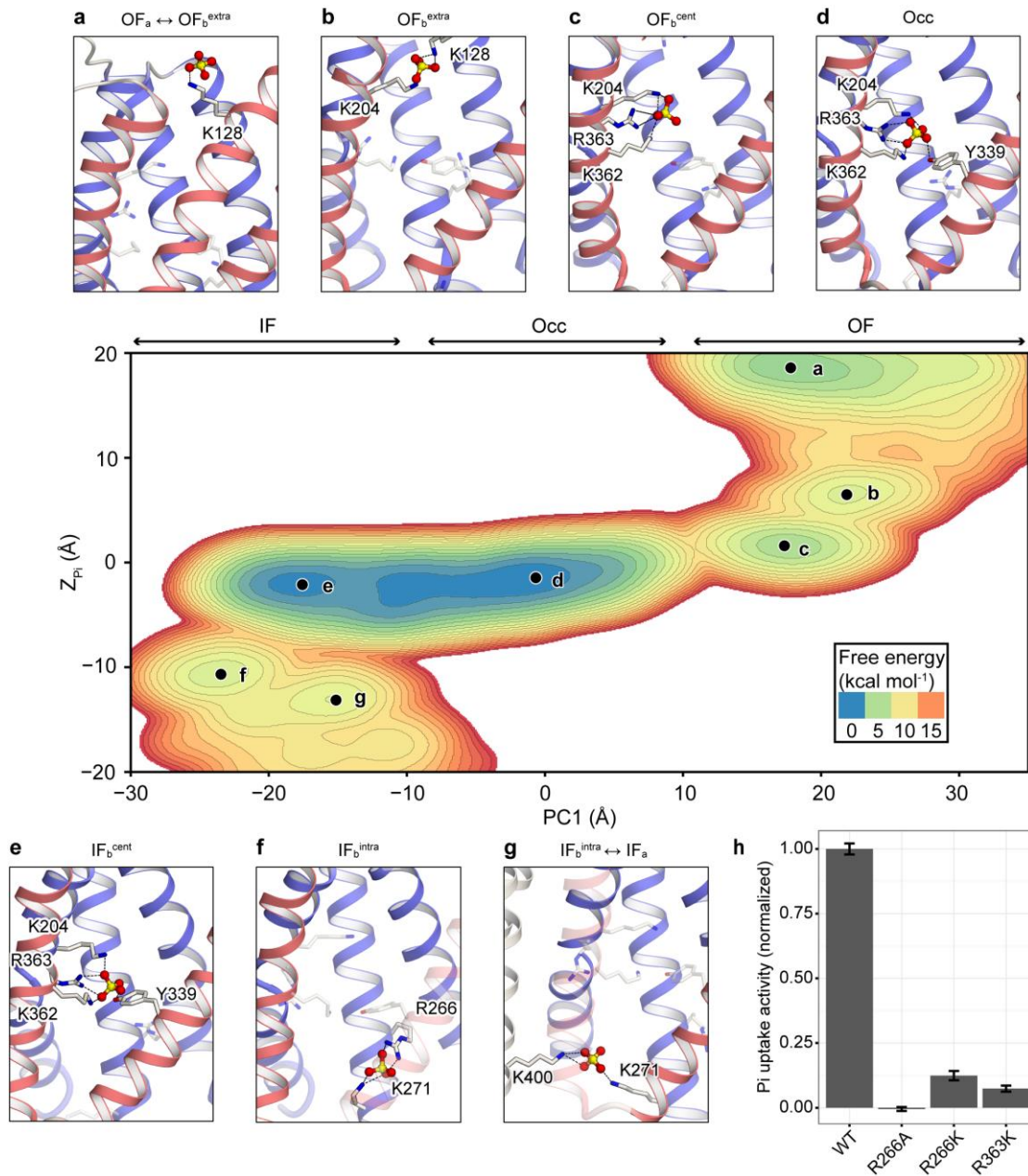
899 **Figure 3—figure supplement 1 | Membrane expression of GsGPT mutants**

900   Western blotting analysis for wild type (WT) and each GsGPT mutant,

901 confirming the comparable expression levels. "M" represents marker lane.

902

903

904

Figure 4 | Substrate binding and global conformational change coupling

Free energy landscape in the (PC1, Z_{Pi}) space is shown for the regions in which P_i is bound to GsGPT. (a-g) Snapshots of the binding site conformation corresponding to each free energy basin. The residues that directly interact with P_i in the transport cycle are shown with stick models. Dotted lines represent strong interactions in which

910 the distance between P_i and each residue is less than 3.0 Å. TM1 is not shown, except in

911 panel a. (h) Liposome-based mutational analysis. The levels of $[^{32}P]$- P_i uptake by

912 GsGPT mutants were normalized with the mean value of the wild-type. The error bars

913 represent s.e.m. (n=3-9; see Figure 3—Source Data1).

914

915

916
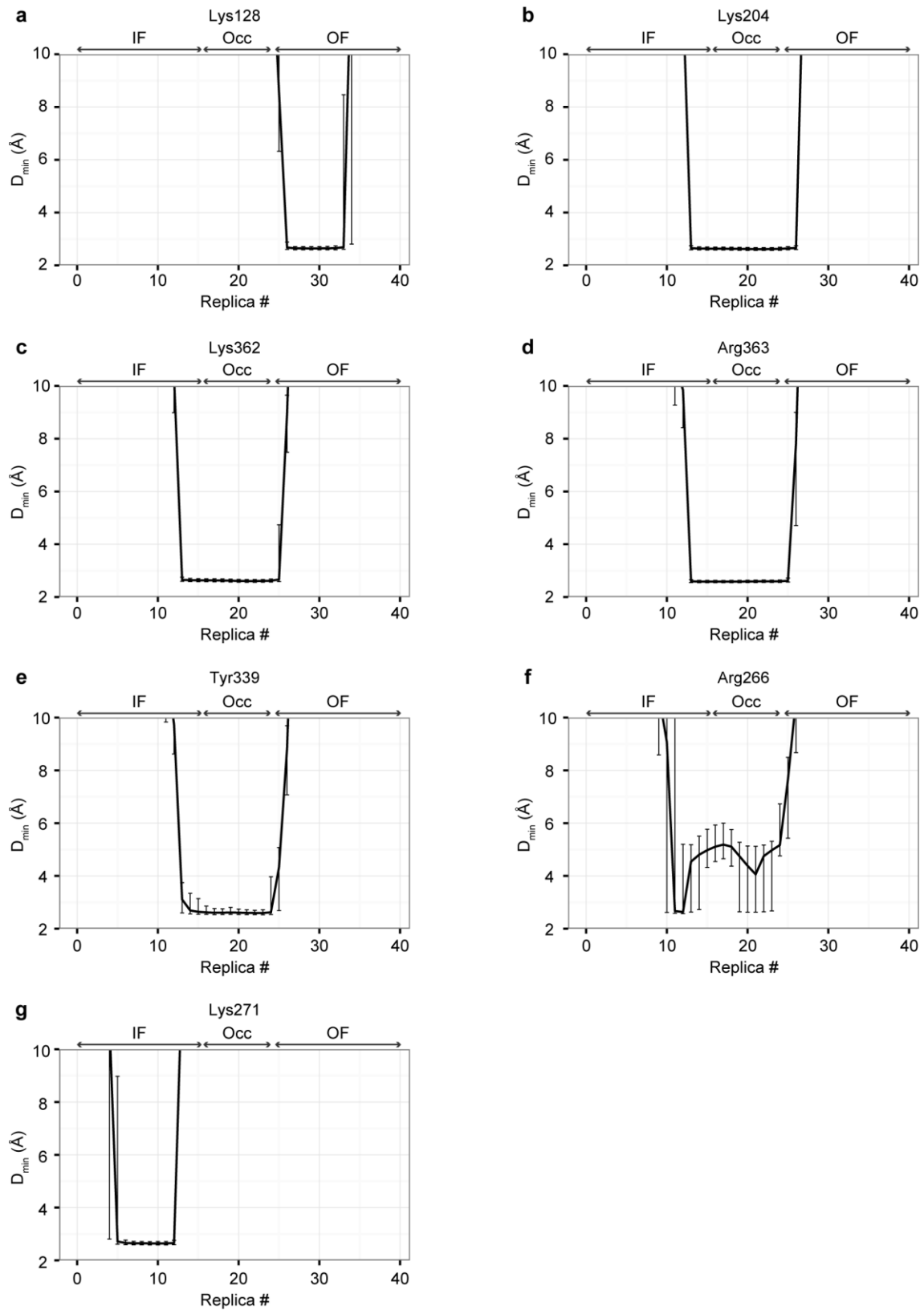
Figure 4—figure supplement 1 I Minimum distances between P_i and the P_i

917

918 **interacting residues**

919   The minimum distances between P_i and the P_i interacting residues, (a) Lys128,

920 (b) Lys204, (c) Lys362, (d) Arg363, (e) Tyr339, (f) Arg266 and (g) Lys271. The median

921 value within each replica is plotted, and the error bars represent the interquartile range

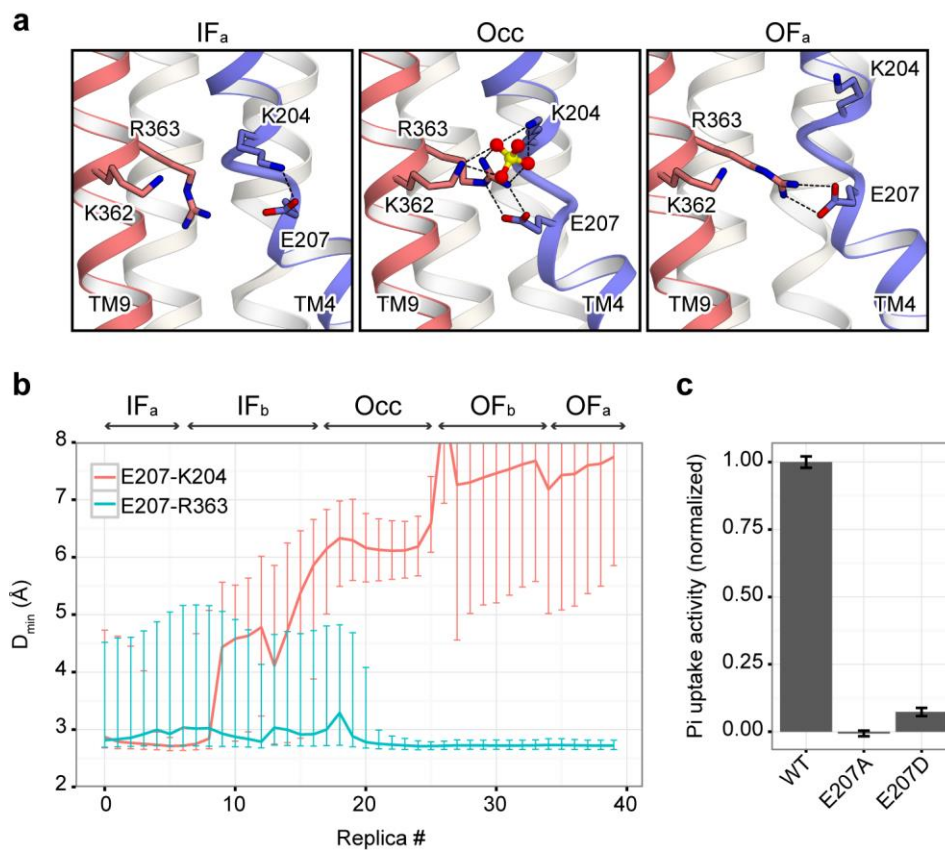922 (IQR).

923

924
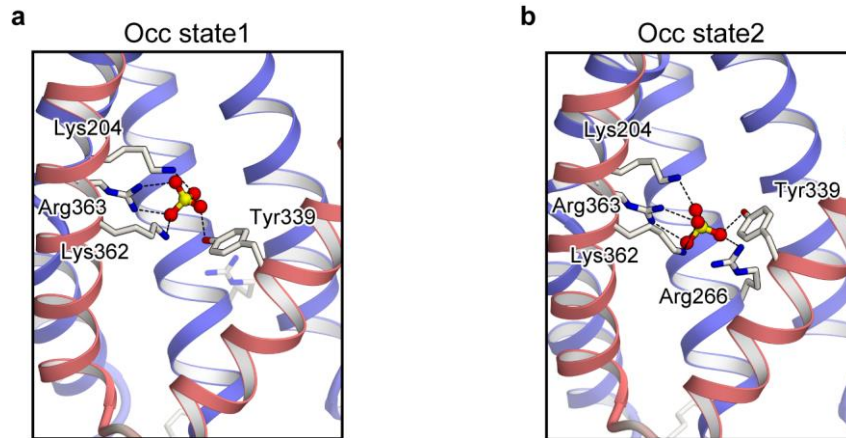
Figure 4—figure supplement 2 | Degenerate states in the Occ conformation

(a, b) Close-up views of the substrate binding site for two conformations corresponding to the free energy basins shown in Figure 4d. TM1 is not shown.
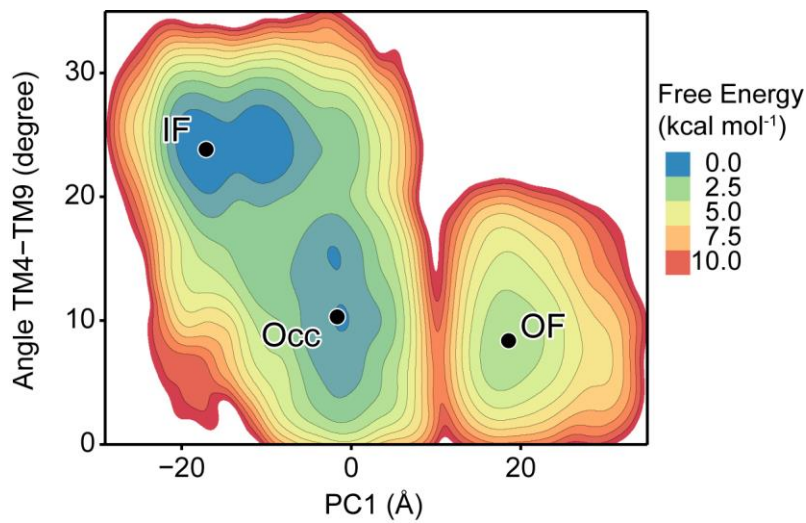
Figure 5 | E207 in the substrate binding site

(a) Close-up view of the substrate binding site for representative structures of IF, Occ and OF. Dotted lines represent strong salt bridge interactions, in which the distance between each atom is less than 3.5 Å. (b) Minimum distance between Glu207 and Lys204 (magenta) and between Glu207 and Arg363 (cyan). The median value of the distance in each replica is plotted, and the error bars represent the interquartile range (IQR). (c) Liposome-based mutational analysis. The levels of $[^{32}P]$-P_i uptake by GsGPT mutants were normalized with the mean value of the wild-type. The error bars represent s.e.m. (n=3-9; see Figure 3—Source Data1).

Figure 5—figure supplement 1 | Free energy landscape for inter-helical

angle between TM4 and TM9

Free energy landscape in terms of (PC1, TM4-TM9 angle) space. PC1 represents

the first principal component of all Cα atoms, and the inter-helical angle between TM4

and TM9 was calculated using the roll axis of the helix (obtained from the principal axis

component analysis for the Cα atoms).

Tables

Table 1 I simulation set list

Sim #	Transition	Method	Collective variable	# of replicas x sim time
1	cryst. $\rightarrow$ IF$_b^{cent}$	non-Eq.	ΔD	10 ns
2	IF$_b^{cent}$ $\rightarrow$ OF$_b^{cent}$	non-Eq.	ΔD	20 ns
3	IF$_b^{cent}$ $\rightarrow$ IF$_b^{intra}$	non-Eq	Z_{Pi}	10 ns
4	OF$_b^{cent}$ $\rightarrow$ OF$_b^{extra}$	non-Eq	Z_{Pi}	10 ns
5	IF$_b^{intra}$ ⇄ Occ ⇄ OF$_b^{extra}$	SMwST	ΔD / Z_{Pi}	40 x 1.2 ns = 48 ns
6	IF$_b^{intra}$ ⇄ Occ ⇄ OF$_b^{extra}$	BEUS	pathCV(s)	20 x 20 ns = 400 ns
7	IF$_b^{cent}$ $\rightarrow$ IF$_a$	non-Eq.	Z_{Pi}	25 ns
8	IF$_b^{cent}$ $\rightarrow$ IF$_a$	BEUS	Z_{Pi}	16 x 20 ns = 320 ns
9	OF$_b^{cent}$ $\rightarrow$ OF$_a$	non-Eq.	Z_{Pi}	25 ns
10	OF$_b^{cent}$ $\rightarrow$ OF$_a$	BEUS	Z_{Pi}	16 x 20 ns = 320 ns
11	IF$_a$ ⇄ Occ ⇄ OF$_a$	SMwST	ΔD / Z_{Pi}	40 x 5.5 ns = 220 ns
12	IF$_a$ ⇄ Occ ⇄ OF$_a$	BEUS	pathCV(s)	40 x 50 ns = 2000 ns
			total	~ 3.4 μs

951

952

Rich media files

Video1

IF-to-OF conformational transition of GsGPT, viewed from the plane of the membrane. The trajectory was reconstructed by the *post‑hoc* string method (PHSM) from the final bias-exchange umbrella sampling (BEUS) simulation (simulation set 12). The P_i molecule is shown with a ball-and-stick model, and the major substrate binding basic residues and Glu207 are highlighted with stick models. The coloring scheme for the TM helices is the same as in Figure 1.

962 **Source Data**

963 **Figure 3—Source Data1 | List of the raw count values of transport assay**

964

965

a

Outside
(intermembrane space)

90°

Inside
(stroma)

b Pᵢ bound

K204

R363

Y339

K362

c outside gate

K128

I197   F192

F355   P352

L347   K271

inside gate

d

Outside

IFₐ   IFᵦⁱⁿᵗʳᵃ   IFᵦᶜᵉⁿᵗ   Occ
(cryst.)   OFᵦᶜᵉⁿᵗ   OFᵦᵉˣᵗʳᵃ   OFₐ

Inside

○ Pᵢ

e

Occ ---1--→ IFᵦᶜᵉⁿᵗ ---2--→ OFᵦᶜᵉⁿᵗ
(cryst)

7   3   4   9

IFᵦⁱⁿᵗʳᵃ   OFᵦᵉˣᵗʳᵃ

IFₐ   5   OFₐ
SMwST1

8   6   10
BEUS2   BEUS1   BEUS3

PHSM

11   12
SMwST2 → BEUS4

f Free energy along the transport cycle

Replica #
0   10   20   30   40

IFₐ

OFₐ

Occ

0.00   0.25   0.50   0.75   1.00 pathCV(s)

a   bundle1

RMSD: 1.4 Å

b   bundle2

RMSD: 1.7 Å

c   histogram for Cα−RMSD vs. crystal structure

a Occ state1

b Occ state2