学位論文

Structural and functional studies of the transporters involved in plant nutrient allocation

(植物の栄養分配に関わる輸送体の構造機能解析)

平成 29 年 12 月博士（理学）申請

東京大学大学院理学系研究科

生物科学専攻

李 勇燦

Abstract

Plants can synthesize organic compounds through photosynthesis and thereby supply almost all organisms on earth with essential building blocks and energy for life. Since plants are multicellular organisms, synthesized compounds must be transported from one cell to another to be distributed across the whole plant body. Such transport process is mediated by various membrane transporters that bind and transport specific solutes across the lipid bilayer membranes. In this thesis, I aimed to understand the molecular mechanisms of two transporters, SWEET and TPT, which play fundamental roles in plant nutrient allocation but are poorly characterized at the structural level.

Sugars Will Eventually be Exported Transporters (SWEETs) are a new family of sugar transporters that play major roles in long-distance sugar transport, seed filling, pollen development and nectar secretion. The SWEET family proteins exhibit no sequence homology to other known sugar transporter families, and thus their structure and mechanism remain elusive. I determined the structures of a bacterial homologue of SWEET, known as SemiSWEET, at two different conformations. The structures revealed a unique three-helix bundle architecture of SemiSWEET, forming a dimer to create a sugar translocation pathway across the membrane. Structural comparison revealed that a kink at transmembrane helix 1 triggers a conformational change that enables the alternating access of a substrate binding site from one side of the membrane to the other.

The triose-phosphate/phosphate translocator (TPT) resides in the inner envelope membrane of chloroplasts and mediates the strict counter-exchange of triose-phosphate, 3-phosphoglycerate, and inorganic phosphate. TPT plays fundamental roles in delivering carbon and energy from chloroplast to cytosol during photosynthesis, but its mechanisms for substrate binding and antiport remain unknown. I determined the structures of TPT in complex with two substrates, 3-phosphoglycerate and inorganic phosphate. TPT exhibits a 10-transmembrane helix topology belonging to the drug-metabolite transporter superfamily. The two substrates exclusively occupy the same central pocket, supporting a competitive binding model. A negatively-charged phosphate group, which is shared by the two substrates, is recognized by the three positively-charged residues. The pocket also accommodates the sugar moiety of 3-phosphoglyceric acid or the water molecules. These structures suggest a model in which competitive binding of substrates and subsequent conformational changes ensure the strict antiport of phosphorylated metabolites and a phosphate ion.

Contents

Figures and Tables

Abbreviations

Abbreviation	Full name
A600	Absorbance (measured at 600 nm)
CHS	Cholesteryl hemisuccinate
DDM	*n*-dodecyl β-D-maltoside
DNA	Deoxyribonucleic acid
FSEC	Fluorescence-detection size-exclusion chromatography
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Luria-Bertani
LCP	Lipidic cubic phase
Ni-NTA	Nickel-nitrilotriacetic acid
MES	2-(N-morpholino)ethanesulfonic acid
PEG	Polyethylene glycol
PMSF	Phenylmethylsulfonyl fluoride
SAD	Single-wavelength anomalous diffraction
SeMet	Selenomethionine
TEV	Tobacco etch virus
Tris	Tris(hydroxymethyl)aminomethane

Chapter 1 General introduction

1.1 Path of carbon in plants

Plants can synthesize organic compounds from CO_2 and thereby supply the Earth's organisms with essential building blocks and energy for life. Since plants are multicellular organisms, synthesized compounds must be transported from one cellular compartment to another and ultimately distributed to the whole plant body (*1*). Such transport process is mediated by various transporters, which bind and transport specific types of solutes across the membrane. Of particular interest are sugar transporters, which play central roles in photosynthetic metabolism, carbon partitioning, sugar signaling, and plant-microbe interactions. Since these sugar transporters can significantly influence plant nutrition and growth, they are important targets of genetic modification for improving crop productivity (*2*).

1.2 Sugar transporters

Major plant sugar transporters in plants comprise sucrose transporters (SUTs) and monosaccharide transporters (MSTs) (*1*). Both these two groups belong to the major facilitator superfamily (MFS) (*3*). Recently, a novel family of sugar transporters belonging to a distinct superfamily were identified, and named SWEETs (*4*). In this section, I briefly discuss these transporter families in terms of their physiological roles, functional properties, structures and unsolved questions.

1.2.1 SUTs

Sucrose is the major form of sugar used for long-distance transport in most plants. An important step in the long-distance transport is the H^+-driven sucrose uptake from the apoplast (extracellular space) to the sieve element-companion cell (SE-CC) complex, a process known as the phloem loading (*5*) (Figure 1a). From 1992 to 1994, molecular genetics studies have identified the transporters responsible for this process, named Sucrose Transporters (SUTs; alternately called SUCs) (*6–8*). Arabidopsis genome contains 9 SUT members (*9*), and rice genome contains 5 SUT members (*10*). According to a recent classification, SUTs are classified into 5 clades (SUT1–SUT5), in which the SUT1 clade is dicot- specific and the SUT3 and SUT5 clades are monocot-specific (*11*).

Functionally, most SUTs function as H^+/sucrose symporters (*12*). In phloem loading, this symport mode drives the import of sucrose from the apoplast (extracellular side) to the intracellular side, by using the H^+ gradient of 1000-fold (from pH 5 to pH 8) (*13*). In sink organs such as potato tubers, SUTs can also operate in a reverse direction, functioning as sucrose exporters (*14*, *15*). In

addition, some SUTs found in nodule-forming legume species are independent of pH gradients and function as sucrose facilitators (SUFs) (*16*), highlighting the diverse modes of transport within the same SUT family. Substrate specificities of SUTs are also diverse; some SUTs transport various glucosides other than sucrose, such as fraxin, exculin and helicin (*17*). Previous studies have shown that the glucose moiety of these sugar-conjugate compounds is important for substrate recognition (*17*, *18*). Transport affinity for sucrose varies among different SUT members, with typical K_m values within the range of 0.1–10 mM (*15*).
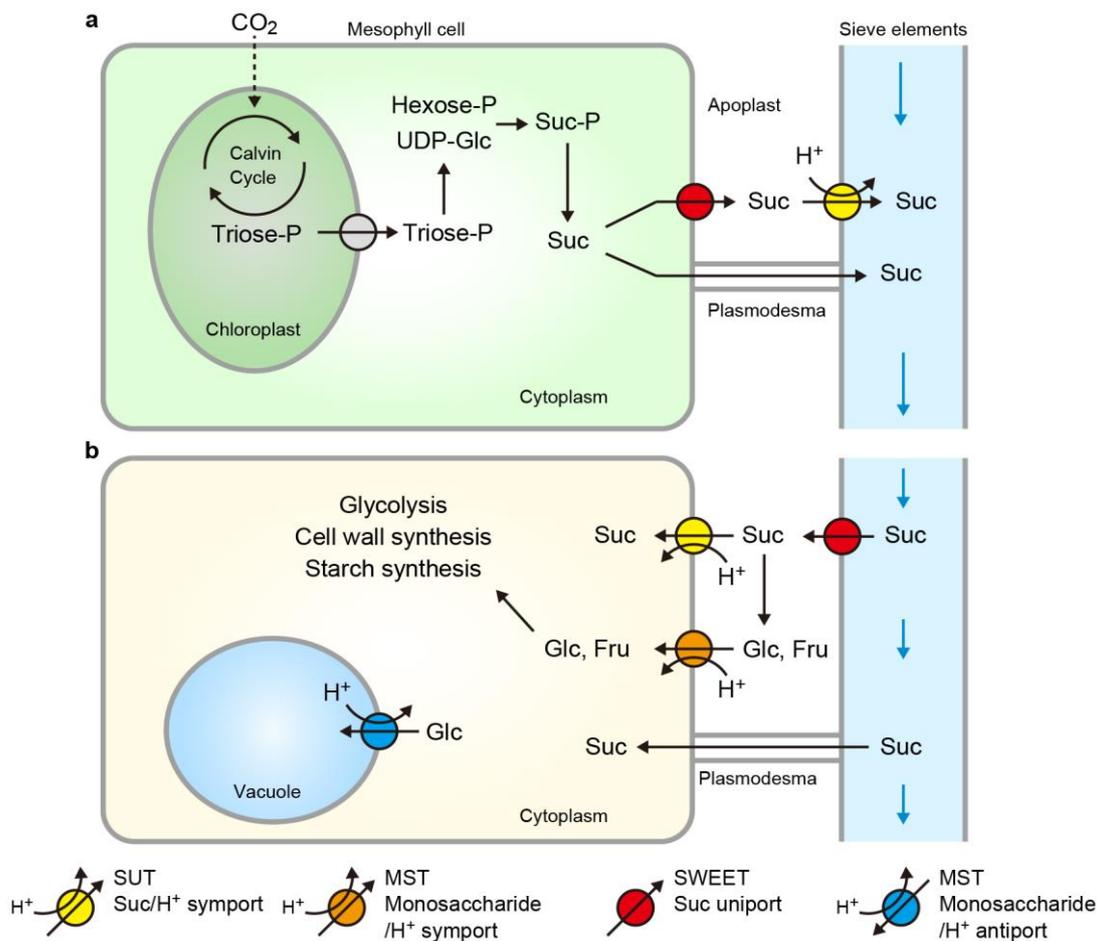
Figure 1 | Sugar transporters in plants

(a) Sucrose (Suc) transport in photosynthesizing leaves (source). Assimilated carbon is converted into Suc within the cytosol, and then loaded to phloem (sieve elements) either through plasmodesmata or by transporters. In the latter, Suc is exported to apoplast by SWEET, and then loaded into the phloem by SUT. (b) Sugar transport in non-green tissues (sink). Suc is unloaded from the phloem either through plasmodesmata or by transporters. In the latter, Suc is taken up by MST after hydrolysis into glucose (Glc) and fructose (Fru), or directly by SUT. Vacuole imports monosaccharides through an MST antiporter. Adapted from ref. (*19*).

1.2.2 MSTs

While disaccharide sucrose is the main sugar used for the long-distance transport, monosaccharides, glucose and fructose, are also widely used for cell-to-cell and subcellular sugar transport in plants. For instance, after long-distance transport, sucrose can be taken up by sink cells as glucose and fructose, after being cleaved by the cell-wall invertase in the apoplast (*20*, *21*) (Figure 1b). Monosaccharide transport is mediated by proteins called Monosaccharide Transporters (MSTs), which were first cloned from a unicellular alga *Chlorella* (*22*). Subsequent analyses have identified MSTs in land plants (*23–26*). MSTs constitute a very large group in most plant species, comprising 53 members in Arabidopsis and 65 members in rice. According to a phylogenetic analysis, MSTs have been classified into 7 distinct subfamilies, namely Sugar Transport Protein (STP), Vacuolar Glucose Transporter (VGT), Tonoplast Membrane Transporter (TMT), Plastidic Glucose Transporter (pGlcT), Polyol/Monosaccharide Transporter (PMT), Inositol Transporter (INT), and Early-Responsive to Dehydration 6-like (ERL) (*25*). Each subfamily contains 4 to 19 members, and the members within the same subfamily show high sequence homology (42%–96%), whereas those between distinct subfamilies show low homology (*25*).

The function of MSTs are diverse. It involves all three types of transport modes: symporter, antiporter and uniporter. The STP subfamily involves H^+/hexose symporters with differing degree of substrate specificities toward glucose, fructose and galactose (*23*, *27*). The PMT and INT subfamilies involve H^+-driven polyol symporters, with some members also transporting monosaccharides (*28–32*). In contrast to these plasma membrane-localized subfamilies, the VGTs and TMT subfamilies are mostly localized at the vacuolar-membranes, and function as H^+/sugar antiporters (*33*, *34*). Since plant vacuole lumens are acidified (pH ~5.5) (*35, 36*), this antiport mode drives the sugar import into the vacuole (*37*, *38*) (Figure 1b). The ESL subfamily is also found at the vacuolar membrane, but its members have been characterized as facilitative uniporters (*39*). The pGlcT subfamily is the only subfamily found at the chloroplast membrane, and involve facilitative uniporters (*40*). With such a diverse range of transport functions, MSTs are involved in various physiological processes (*3*, *41*).

1.2.3 Major facilitator superfamily

SUTs and MSTs described above belong to a large protein superfamily called the Major Facilitator Superfamily (MFS) (*42*). The typical architecture of the MFS protein comprises 12 transmembrane (TM) helices, which can be divided into the N- and C-terminal 6-TM segments (*43*) (Figure 2a). These two 6-TM segments are thought to have occurred through gene duplication (*44*). The first crystal structure of an MFS protein, the bacterial lactose permease LacY, has

revealed that these two 6-TM segments fold into distinct structural units (the N- and C-bundles), forming a substrate-binding site at the interface (*45*) (Figure 2b). Subsequent structural analyses have revealed that this substrate-binding site is alternately exposed inward and outward to transport the substrate, accompanying a large inter-domain conformational change (*45*). These findings have made the MFS sugar transporter a 'text-book' model for all transporters (*46–49*) . Although the structures of plant SUTs and MSTs are yet to be determined, these available structures of MFS sugar transporters help understand the mechanisms of SUTs and MSTs (*50*).
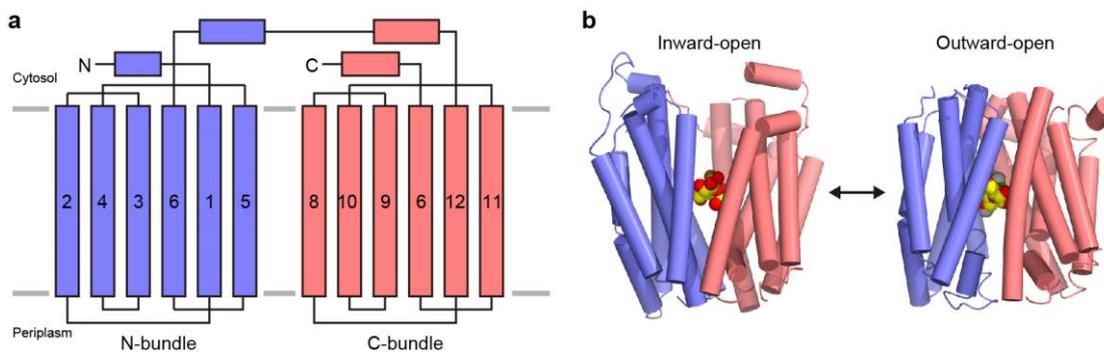
Figure 2 | Major facilitator superfamily (MFS) sugar transporter

(a) Topology of MFS sugar transporters. Adapted from ref. (*44*).

(b) The crystal structures of the lactose permease LacY. In the inward-open structure (left; PDB 1PV7), a substrate analog β-D-galactopyranosyl-1-thio-β-D-galactopyranoside (TDG) is bound at the central cavity, accessible from the cytoplasmic side. In the outward-open structure (right; PDB 4OAA), TDG is bound at the same central cavity, but is now accessible from the periplasmic side.

1.2.4 SWEETs

Although molecular identities of most sugar transporters involved in plant metabolism have been characterized by around the 2000s, some key transporters involved in sugar export have remained elusive. In 2010, a new class of sugar transporters were identified and named SWEETs (Sugars Will Eventually be Exported Transporters) (*51*). SWEETs are prevalent in plants than animals, comprising 17 members in Arabidopsis and 21 in rice (*4*, *52–57*). The plant SWEETs have been classified into four clades (clades I–IV), and each clade shows slightly different preference towards monosaccharides and disaccharides (*4*). The animal SWEETs comprise only a few members in each species (1 in human and 7 in *Caenorhabditis elegans*), and their functions are not well characterized (*4*).

Functionally, SWEETs are uniporters that catalyze facilitative diffusion of sugars down the concentration gradient (*58*, *59*). In physiological contexts, most SWEET members mediate sugar

efflux from sugar-rich compartments to extracellular space (*58*, *59*). For example, AtSWEET11 and AtSWEET12 have been found to mediate sucrose efflux from phloem parenchyma cells to apoplast, supplying sucrose for long-distance transport in phloem (*60*). This discovery has resolved the long-standing questions on the molecular identities of enigmatic sugar exporter during long-distance transport (*61*) (Figure 1a). In angiosperms, flower nectary can secrete high concentration of sugar for attracting pollinators (*61*). AtSWEET9 mediate sucrose export from flower nectary and is thus essential for nectar secretion (*58*, *62*). In maize and rice, ZmSWEET4c and OsSWEET4 mediate hexose import into seeds, playing essential roles in seed filling (*4*, *52*, *53*, *63*, *64*). Some SWEETs are localized in the vacuolar membranes and mediate sugar import to modulate intracellular sugar levels (*4*, *52–54*). Substrates of SWEETs include monosaccharides, glucose, fructose and galactose (*57*, *65*, *66*) and the disaccharide, sucrose (*67*). Some SWEET members have been recently implicated in the transport of gibberellin, a plant hormone without a sugar moiety (*4*, *53*, *57*, *64*). Transport affinities of SWEETs are relatively low, typically within the 1–70 mM range (*68*).

SWEET is predicted to possess 7 transmembrane (TM) helices, which consist of the two '3-TM repeats' at the N- and C-terminal halves (TM1–3 and TM5–7), connected by a 'linker helix' (TM4). This architecture is distinct from any other known transporter families. Recently, a prokaryotic homologue of SWEET has been identified (*68*). This bacterial SWEET homologue possesses only one unit of the 3-TM repeat, and thus appears to be 'half-sized' from its eukaryotic counterpart. Therefore, this bacterial SWEET homologue has been named SemiSWEETs (*68*). Phylogenetic analysis has shown that SemiSWEET is an ancestor of SWEET, probably forming a functional transporter by assembling as an oligomer. However, unlike the MFS proteins, SUTs and MSTs, the mechanism of sugar transport by SWEET and SemiSWEET transporters is not understood.

1.3 Sugar-phosphate transporters

In addition to sugars, sugar phosphates are important metabolites in plant metabolism. For instance, in photosynthesizing leaves, assimilated carbon is exported from the chloroplast in the form of triose phosphates (triose-P; dihydroxyacetone phosphate and glyceraldehyde-3-phosphate) (*69*). This export is achieved in exchange for the import of inorganic phosphate (Pi) from the cytosol (*74*). In non-photosynthetic organs such as roots and seeds, import of glucose-6-phosphate into the plastid is important for starch synthesis (*78*). Sugar phosphates are especially important for connecting the chloroplast-cytosol metabolisms. Since chloroplasts are enclosed by the two lipid bilayer membranes, called the outer and the inner envelopes, specialized transporters are needed to deliver these compounds between chloroplast and cytosol. The inner envelope

membrane is responsible for the specific transport of sugar phosphates (*88*), whereas the outer envelope membrane is non-specifically permeable to a wide variety of solutes (*73*).

1.3.1 pPTs

The plastidic phosphate translocators (pPTs) are a family of transporters found on the plastid inner envelope membrane (*72*, *73*). In land plants, the pPT family consists of four subtypes, the triose-phosphate/phosphate translocator (TPT), the phosphoenolpyruvate/phosphate translocator (PPT), the glucose-6-phosphate/phosphate translocator (GPT) and xylulose-5-phosphate/phosphate translocator (XPT).

TPT is found at the inner envelope membrane of all chloroplasts (Figure 3a). TPT catalyzes either triose-P/Pi or triose-P/3-PGA exchange across the chloroplast inner envelope membrane. The former reaction delivers a carbon skeleton to the cytoplasm and transports Pi back into the chloroplast for ATP regeneration (*74*). The latter reaction, known as the triose-P/3-PGA shuttle, indirectly exports chemical energy (ATP and NADPH) without the net transport of carbon (*75*). PPT is expressed in both chloroplast and non-green plastids, and exchanges phosphoenolpyruvate (PEP) with Pi, playing important roles in amino acid and fatty acid biosynthesis (*76*) (Figure 3b). PPT also serves as a part of the CO_2 concentration mechanism of C4 and CAM photosynthesis (*76*, *77*). GPT is mainly expressed in non-green plastids like amyloplasts and exchanges glucose-6-phosphate (Glc-6-P) with Pi, functioning as a carbon importer (*78*). XPT is expressed in both chloroplasts and non-green plastids and mediate the exchange of xylulose-5-phopshate (Xul-5-P), triose-P, and Pi, thereby supplying precursors for pentose phosphate pathway in the plastids or the cytosol (*79*, *80*). The pPTs are also found in apicomplexan parasites (*81*, *82*), which cause toxoplasmosis and malaria in humans. Since these apicomplexan pPTs are essential for the survival of the parasites (*83*), they are potential drug targets for parasitic infections (*84*).

All pPT proteins are known to catalyze the strict 1:1 exchange reactions, which guarantee the total balance of the plastid and the cytosol while allowing the transport of carbon and energy (*72*, *85*). The affinity for transport by the pPTs are typically in the millimolar range, close to the physiological concentrations of sugar-phosphates and Pi within plastids (*88*) . In addition to sugar-phosphate/Pi hetero-exchange reactions, all functionally characterized pPTs can catalyze the Pi/Pi homo-exchange reaction, which is known as the signature activity of this family (*89*). Previous *in vitro* studies have shown that the direction of import and export reactions can be reversed, although in physiological contexts the net transport direction is dominated by the concentration difference of substrates between the two compartments (*86*).

Despite a wealth of molecular and biochemical studies, the structure and mechanism of the pPTs are poorly understood. The first molecular cloning of the pPT has been achieved for a spinach TPT, which has revealed that TPT possesses seven highly hydrophobic regions, suggestive of a 7-TM architecture (*70*). Later studies have proposed different architectures consisting of 6 to 9 TMs (*86*). All pPTs are nuclear-encoded proteins, and are therefore targeted to chloroplasts after being translated as precursor proteins, which include the target sequences called the chloroplast transit peptides (cTPs) (*91*). The cTP is about 90 residues in each precursor protein (*4*) and is cleaved after membrane insertion, giving rise to a mature translocator protein.
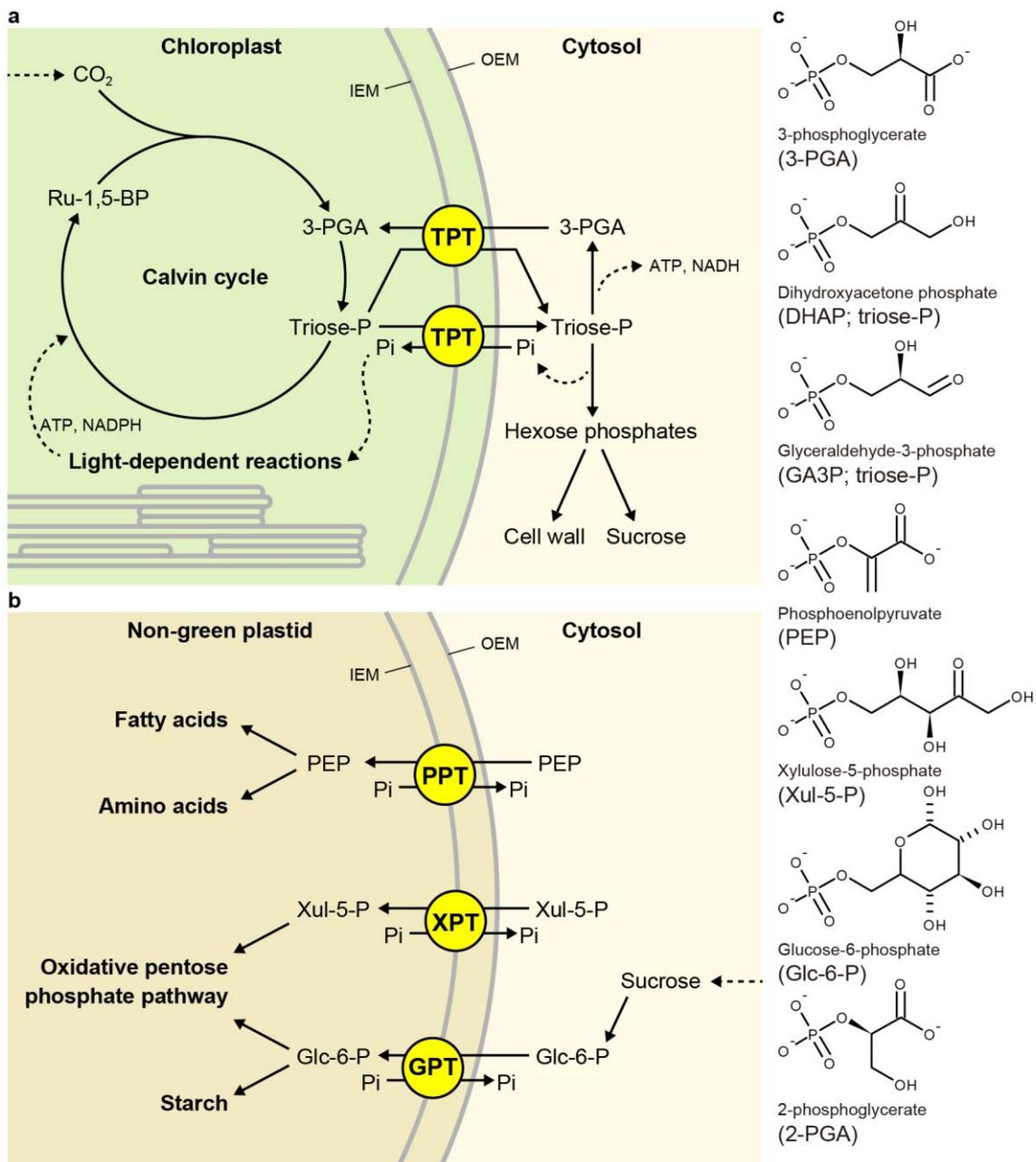
Figure 3 | Roles of the pPTs in plastid metabolism.

(a) The function of TPT in chloroplasts. TPT catalyzes the antiport of triose-P, 3-PGA and Pi across the inner envelope membrane. OEM and IEM denote the outer and inner envelope membranes, respectively.

(b) The functions of PPT, GPT and XPT.

(c) Chemical structures of the pPT substrates.

1.3 Aim of the thesis

In this thesis, I aim to elucidate the molecular mechanisms of two metabolite transporters, SWEET and TPT. These transporters play important roles in photosynthetic carbon fixation, plant nutrition and growth, but are poorly understood at the molecular levels. I therefore undertook structural and functional studies on these transporters, by using X-ray crystallography and in vitro functional characterizations.

Specific goals of this research are:

1) To elucidate the structures of these transporters by using X-ray crystallography.

2) To reveal the molecular basis of substrate recognition and transport with structure-based functional analyses.

Chapter 2 X-ray crystallographic analysis of SemiSWEET

2.1 Introduction

As discussed in Chapter 1, the SWEET family proteins are a new class of sugar transporters that play fundamental roles in plant sugar allocation (*4*). SWEET is a 7TM protein, and its bacterial homologue SemiSWEET shares the common architecture consisting of the 3-TM repeat (*68*). SWEETs and SemiSWEETs are also related to a protein family called the PQ-loop family, which includes lysosomal amino acid exporters implicated in human disease cystinosis (*92–94*). The members of the PQ-loop family share the same '3-TM repeats' architecture and possess a signature motif consisting of Pro-Gln residues (PQ-loop motif), which is located on the first TM helix of each 3-TM repeat. The first proline residue is conserved in both the SemiSWEETs and SWEETs, while the second glutamine residue is conserved in only the SemiSWEETs, suggesting the critical role of the proline residue in their transport mechanisms.

Eukaryotic SWEETs have been shown to form homo- and hetero-oligomers in yeast two-hybrid assays and split ubiquitin assays (*68*). In addition, co-expression of the defective SWEET mutants with wild-type SWEET have exhibited marked decrease in glucose transport activity, indicative of negative dominance. These results have suggested that SWEETs would form a sugar translocation pathway by assembling as an oligomer (*68*). Figure 4 shows a proposed model for oligomerization, which assumes that four 3-TM repeats form a 12-TM architecture, resembling the well-studied MFS structure. However, the precise mechanisms by which SemiSWEETs and SWEETs assemble to form such a sugar translocation pathway is not understood. Notably, SemiSWEET is the smallest sugar transporter identified to date, and is thus a good model system for studying not only the architecture of SWEET family proteins, but also the minimal functional unit of sugar transporters in general. Therefore, I aimed to determine the structure of SemiSWEET for understanding the shared molecular mechanisms of the SWEET family transporters.
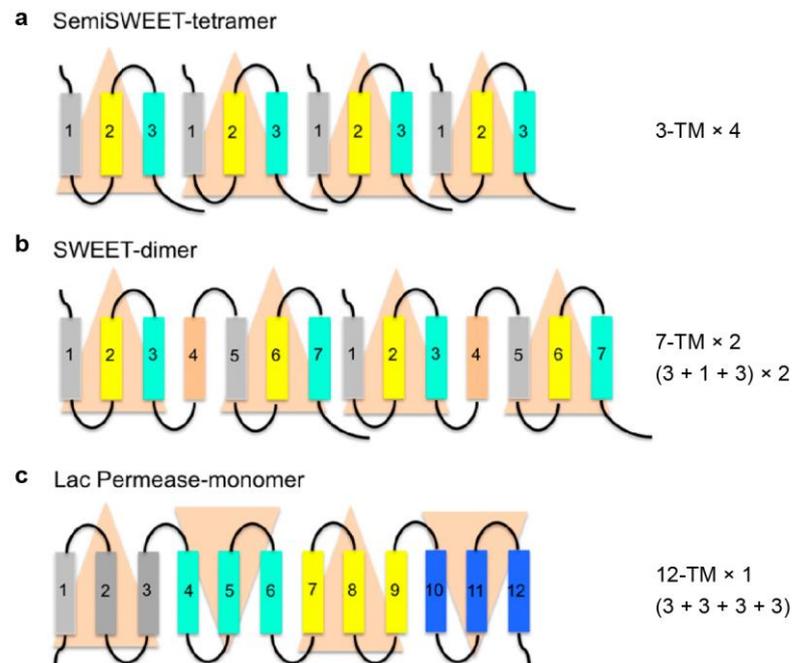
Figure 4 | Proposed model for SWEET and SemiSWEET oligomerization

(a,b,c) Models for the functional oligomers of SemiSWEET (a), SWEET (b) and the lactose permease LacY (c). The functional unit is assumed to be a dimer for SWEET and a tetramer for SemiSWEET. Reproduced from ref. (*68*).

2.2 Materials and methods

2.2.1 Cloning

The SemiSWEET genes were cloned from bacterial and archaeal genomes by using polymerase chain reaction (PCR) with specific primers. These primers were designed to add the XhoI and NdeI cleavage sites at the 5' and 3' termini of the genes, respectively. The cloned gene fragments were cut by XhoI/NdeI double digestion and ligated into the XhoI/NdeI site of a modified pET28a vector. This vector was designed to introduce a tobacco etch virus (TEV) protease-cleavage site and a His_8-tag at the C-terminus of the SemiSWEET protein. The genome accession numbers, the nucleotide accession numbers, and the primer sequences used for cloning are summarized in Table 1.

Table 1 | Primers used for cloning SemiSWEETs

Genome	JCM No.	Nucleotide (Gene ID: start-end)	Primer sequences (5'–3')
Bradyrhizobium japonicum	10833	gi|384213726: c3075640-3075380	GCGCCATATGGACCCGTTCTTGATCAAGCTGATCG*
			GCGCCTCGAGGCCGTATCTCAGCTTCATCACCAG
Magnetospirillum gryphiswaldense	21280	gi|23014828: 16911-17207	GCGCCATATGGACTGGCTTTCACCCACCGACC
			GCGCCTCGAGCTGGTTCAAGCTGGTCCGGGTGCC
Desulfovibrio vulgaris	14930	gi|218885109: 2360689-2360985	GCGCCATATGCCCGCACCCACCGCAGACTCCG
			GCGCCTCGAGCAGCTGGCGGGAATAGCGCAGCTT
Methanosarcina mazei	9314	gi|21226102: 2327857-2328168	GCGCCATATGATCGGCTATATCGCAGGTGCCC
			GCGCCTCGAGCTTCCTGAGAGAAGCATTTTTATT
Methanocaldococcus jannaschii	10045	gi|15668172: c106892-106611	GCGCCATATGGTGATTAATATGGACTTTGATA
			GCGCCTCGAGAGATTTTCTCATCTCATCTCTTAT
Methanocaldococcus fervens	15782T	gi|256809973: 329575-329862	GCGCCATATGGGTATAACGATTATTGGCTACA
			GCGCCTCGAGAGTTATTAATTCTGATTTGCTACC
Methanocaldococcus infernus	15783	gi|296108688: 1158684-1158956	GCGCCATATGGATTTAACCATTATTGGTTACT
			GCGCCTCGAGTCTTTTCTTCCAATTTCTCACCAC
Escherichia coli	20135	gi|510922005: 72678-72947	GCGCCATATGGATACCATTCTTTTAACCGGGC
			GCGCCTCGAGTACGTGTTTTTTCCTGCGGTTGAT
Microlunatus phosphovorus	9379	gi|336115651: c3894738-3894127	GCGCCATATGCTCGTCGTAGCCCTGGGTTGGG
			GCGCCTCGAGAGCGGACATGAGCGCCAGGCGAGC

*XhoI and NdeI sites are underlined.

2.2.2 Protein expression

The plasmids outlined in 2.2.1 was transformed into *E. coli* Rosetta 2 (DE3) cells. The cells were inoculated in Luria-Bertani (LB) medium supplemented with 50 μg/ml ampicillin and 30 μg/ml chloramphenicol at 37ºC. Protein expression was induced with 0.2 mM isopropyl β-D-thiogalactopyranoside (IPTG) when the culture reached $A_{600} = 0.6$. After growth for 20 h at 20ºC, the cells were pelleted and resuspended in a buffer containing 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, and 0.1 mM phenylmethylsulfonyl fluoride (PMSF), and then disrupted using a Microfluidizer processor (Microfluidics) with three passes at 15,000 p.s.i. Cell debris was removed by low-speed centrifugation at 10,000 g for 10 min, and the membrane fraction was collected by ultracentrifugation at 138,000 g for 1 h.

2.2.3 Fluorescence-detection size-exclusion chromatography

The fluorescence-detection size-exclusion chromatography (FSEC) was performed following established protocols (*95*, *96*). To monitor the solution behavior of non-GFP-tagged SemiSWEET constructs, a polyhistidine-tag specific fluorescence probe, known as P3NTA was employed (*90*). The membrane fraction from 4 ml-culture was routinely solubilized in a 300 μl buffer containing

50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.1 mM PMSF, 2% dodecyl-β-D-maltopyranoside (DDM) and 0.4% cholesteryl hemisuccinate (CHS) for 90 min at 4ºC, and insoluble components were removed by ultracentrifugation at 138,000 g, 20 min, in a micro-ultracentrifuge CS100FNX (Hitachi Kohki). To 100 μl of supernatant the P3NTA reagent was added at 100:1 (v/v) and allowed to bind for 30 min. 50 μl of resulting solution was subjected to FSEC analysis on a Superdex 200 10/300 column (GE Healthcare), equilibrated in a buffer containing 10 mM Tris-HCl, pH 8.0, 150 mM NaCl, and 0.03% (w/v) DDM. The elution profile was monitored by a fluorescence detector RF-20Axs (Shimadzu). The excitation and emission wavelengths were set to 482 nm and 520 nm, respectively.

2.2.4 Protein purification

The membrane fractions were solubilized in a buffer containing 20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 10 mM imidazole, 2% dodecyl-β-D-maltopyranoside (DDM) and 0.4% cholesteryl hemisuccinate (CHS), for 90 min at 4 ºC. Insoluble components were removed by ultracentrifugation at 138,000 g for 30 min, and the supernatant was mixed with Ni-NTA Superflow resin (Qiagen) for 90 min. The resin was washed with 20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 50 mM imidazole, 0.05% DDM and 0.01% CHS, and the protein was eluted with the same buffer supplemented with a final concentration of 300 mM imidazole. The eluate was treated with tobacco etch virus (TEV) protease to cleave the His$_8$-tag, and dialyzed overnight against the imidazole-free buffer. The sample was then reloaded onto the Ni-NTA Superflow resin (Qiagen) to remove the cleaved tag and the TEV protease. The flow-through fraction containing SemiSWEET was concentrated to about 5 mg ml^{-1} with a 30 kDa concentrator (Millipore), and was further purified by chromatography on a Superdex 200 Increase 10/300 gel filtration column (GE Healthcare), in a buffer containing 10 mM Tris-HCl, pH 8.0, 100 mM NaCl, 0.05% DDM and 0.01% CHS. The purified protein was concentrated to 15 mg ml^{-1}, flash frozen in liquid nitrogen, and stored at −80ºC until crystallization. The typical protein yield was 0.9 mg per one liter of bacterial culture. The SeMet-labeled SemiSWEET protein was produced in *E. coli* B834 (DE3) cells in a LeMaster medium (Wako), and purified by the same procedure as for the native protein.

2.2.5 Crystallization

Purified samples were thawed and reconstituted into the lipidic cubic phase (LCP) of 1-oleoyl-R-glycerol (monoolein) at a protein to lipid ratio of 2:3 (w/w), using the two-syringe mixing method (*91*). Crystallization experiments were performed in two different setups: the sandwich-drop and the hanging-drop methods. For the sandwich-drop crystallization, aliquots of the protein-LCP mixture were dispensed onto 96-well glass plates and overlaid with the precipitant solution, using

a mosquito LCP (TTP LabTech). For hanging-drop crystallization, the protein-LCP drops were manually spotted onto siliconized glass coverslips and overlaid with the precipitant solutions, and then the coverslips were placed upside down onto 24-well plates, sealing each well containing 300 µl of reservoir solution, with the same composition as that of the precipitant solution. The native protein was crystallized under two similar conditions (Crystal-I and Crystal-II). Crystal-I was grown in sandwich-drop plates, with 50 nl protein-LCP drops overlaid with 700 nl precipitant solutions, which consisted of 28% PEG550MME, 100 mM Tris-HCl, pH 8.0, 350 mM $MgSO_4$ and 3% galactose. Crystal-II was grown in hanging-drop plates, with 50 nl protein-LCP drops overlaid with 800 nl precipitant solution, which consisted of 23% PEG550MME, 100 mM Tris-HCl, pH 8.0, 350 mM NH_4-citrate and 3% dimethyl sulfoxide. The SeMet-labeled protein was crystallized under conditions similar to those for Crystal-I, using a precipitant solution consisting of 23–30% PEG550MME, 100 mM Tris-HCl, pH 8.0 and 150–200 mM $(NH_4)SO_4$. All crystals were harvested and flash-cooled in liquid nitrogen for data collection.

2.2.6 X-ray diffraction data collection

X-ray diffraction data were collected at the SPring-8 beamline BL32XU, using a helical data collection method with a 1×10 µm (width $\times$ height) microbeam. The detector used was MX225HS CCD detector (Rayonix). For Crystal-I, data were collected for 540º from a single crystal, using X-ray wavelength of 1.0000 Å and the camera distance of 200 mm. For Crystal-II, data were collected for 180º from a single crystal, with X-ray wavelength of 1.0000 Å and the camera distance of 250 mm. For SeMet-labeled crystals, data were collected for a total of 1120º from six crystals, with X-ray wavelength of 0.9792 Å and the camera distance of 300 mm. All diffraction data were processed using XDS (*91*). Multiple data sets from SeMet-labeled crystals were scaled and merged using XSCALE (*92*).

2.2.7 Data processing and structure determination

For Crystal-I, the structure was determined by the SAD method, using the merged data from six SeMet-labeled crystals. Selenium sites were first determined using SHELXD (*93*), and those sites were used to calculate initial phases using AutoSHARP (*94*, *95*). The model was automatically built from using PHENIX AutoBuild (*96*). The obtained model was transferred to the native Crystal-I data, and then iteratively rebuilt and refined using COOT (*104*) and PHENIX, respectively. For Crystal-II, the structure was determined by molecular replacement in PHASER (*93*), using the SemiSWEET monomer of Crystal-I as a search model. The resulting model was manually rebuilt using COOT. Initial rounds of refinement were performed using REFMAC (*101*) with 'use jelly-body refinement' option, and subsequent rounds were performed with PHENIX.

2.2.8 Transport assays

The purified SemiSWEET protein was reconstituted into liposomes with the following procedure. *E. coli* Polar Lipid Extract (Avanti) was dissolved in chloroform and dried into a thin film in a glass vial. The film was then resuspended to a final lipid concentration of 20 mg ml^{-1} in a buffer containing 10 mM Tris-HCl, pH 8.0, and 100 mM NaCl, and sonicated for 1 min to obtain a liposome solution. The purified protein was added to the liposome solution at a lipid to protein ratio of 100:1 (w/w), while 0.05% DDM was also added to avoid protein denaturation. The protein-liposome mixture was freeze-thawed three times for full reconstitution, and then sonicated for 1 min for unilamellar vesicle formation. Protein-free liposomes were prepared by a similar procedure, except that the protein solution was replaced with the buffer used for the final purification step.

The time-dependent [14C]-sucrose uptake assay was initiated by mixing the liposome solution with an equal volume of the extra-liposomal solution, consisting of 10 mM Tris-HCl, pH 8.0, 100 mM NaCl, and 10 mM [14C]-sucrose (1 µCi ml^{-1}). After the reaction at 37ºC for indicated times, the liposomes were isolated by gel-filtration with Sephadex G-50 (GE Healthcare), and the radioactivity of the incorporated [14C]-sucrose was measured by liquid scintillation counting. The concentration-dependent [14C]-sucrose uptake was measured by a similar procedure, with different concentrations of the extra-liposomal [14C]-sucrose. For mutational analyses, mutations were introduced by a PCR-based method. The mutant proteins were expressed, purified and reconstituted into liposomes, and the transport activities were measured by a similar procedure to that for the wild-type.

2.3 Results

2.3.1 Identification and screening of SemiSWEET orthologues

At the beginning of the study, the only functionally characterized SemiSWEET protein has been the one from a nodule-forming bacterium *Bradyrhizobium japonicum* (*68*). To identify suitable candidates for structural and functional studies, we first searched for SemiSWEET orthologues in bacteria and archaea. BLAST searches into the National Center for Biotechnology Information (NCBI) database identified more than 90 SemiSWEET orthologues, among which we could clone 9 genes from the available genomes in the laboratory. We heterologously expressed these SemiSWEETs in *Escherichia coli* as poly-histidine-tagged proteins (Figure 5), and subjected them to the fluorescence-detection size exclusion chromatography (FSEC) (*95*, *96*). To avoid possible artefacts stemming from fusing GFP (~290 residues) to smaller SemiSWEET proteins (~100 residues), we labeled the proteins with a fluorescence probe called P3NTA, which binds to a polyhistidine tag (*97*). This method allowed us to rapidly evaluate the expression levels and solution behaviors of the SemiSWEET candidates from unpurified, crude membrane fractions. Figure 6 shows the fluorescent profiles of different SemiSWEET orthologues in the size exclusion chromatography. We observed the highest expression level for a SemiSWEET from *E. coli*, which also exhibited the most homogeneous size distribution. Therefore, we selected this *E. coli* SemiSWEET for subsequent biochemical and structural studies.
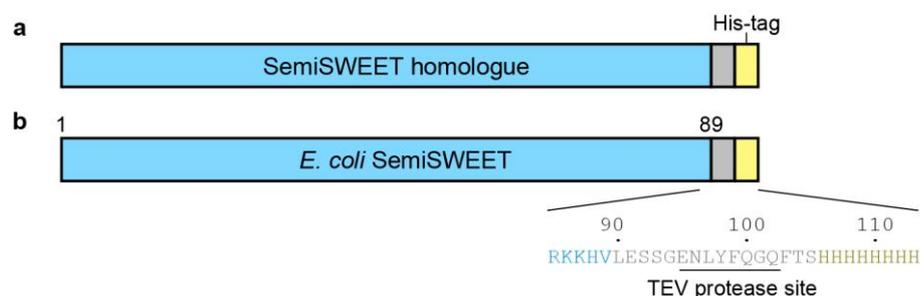
Figure 5 | Expression constructs of SemiSWEET

(a) Construct used for expression screening.

(a) *E. coli* SemiSWEET construct used for purification. His$_8$-tag were fused to the C-terminus with a TEV protease recognition site.
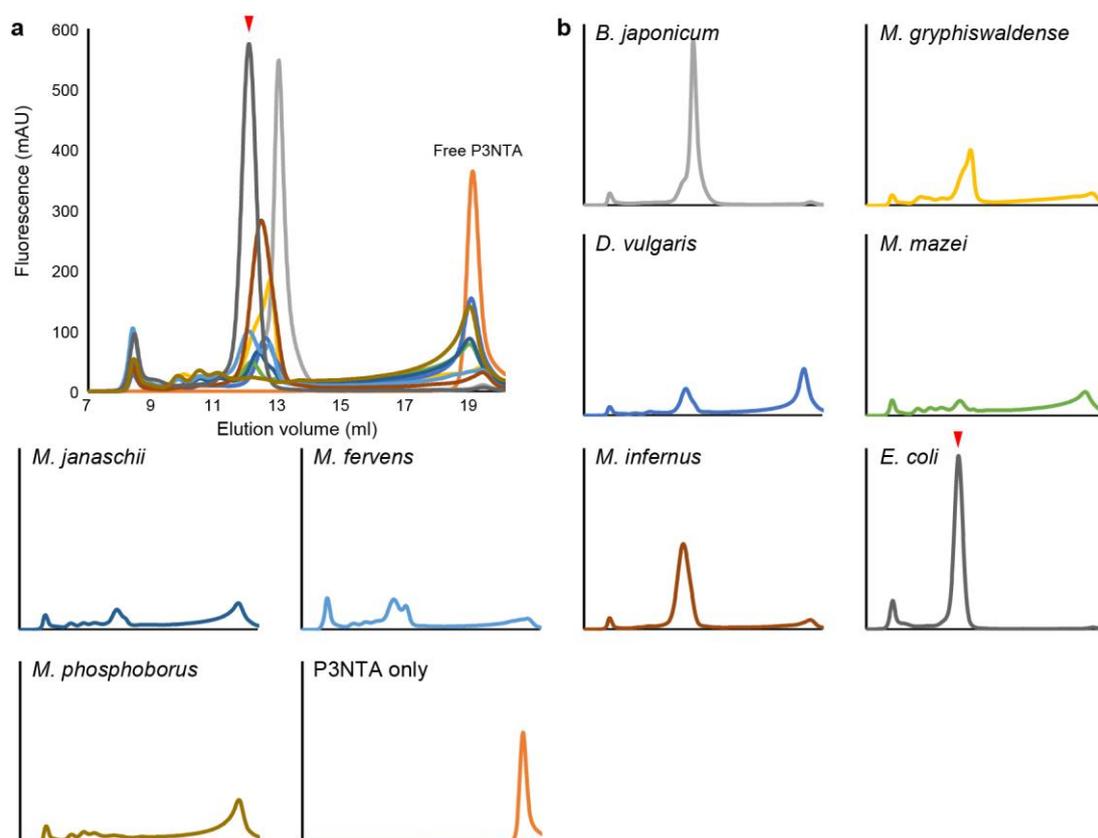
Figure 6 | FSEC profiles of the SemiSWEET orthologues

(a) Superposition of fluorescence traces of different SemiSWEET orthologues in size exclusion chromatography, detected with the histidine tag-specific fluorescence probe P3NTA. mAU!!!

(b) Individual chromatograms of (a). The peaks of *E. coli* SemiSWEET are highlighted by red arrowheads.

2.3.2 Purification of *E. coli* SemiSWEET

We purified *E. coli* SemiSWEET to homogeneity by immobilized metal affinity chromatography and size exclusion chromatography (Figure 7). *E. coli* SemiSWEET shares a 36% sequence identity and a 57% similarity with *B. japonicum* SemiSWEET (Figure 9), which has been characterized as a sucrose uniporter (*4, 57*). To test the transport function of *E. coli* SemiSWEET, we reconstituted the purified proteins into liposomes and measured the uptake of [^{14}C]-labeled sucrose. *E. coli* SemiSWEET showed slow but significant [^{14}C]-sucrose uptake, as compared to the control empty liposomes, confirming its sucrose transport activity (Figure 8a). The rate of [^{14}C]-sucrose was not saturated even up to 300 mM concentration, indicating the low-affinity binding of SemiSWEET to sucrose (Figure 8b). Previous studies have also shown that the plant SWEETs exhibit low affinities for sugars (K_m >70 mM) (*90*), suggesting a common transport mechanism. The unusually slow uptake observed here suggests that sucrose might not be a physiological substrate for *E. coli* SemiSWEET, but these data confirm the sugar transport activity of *E. coli* SemiSWEET.
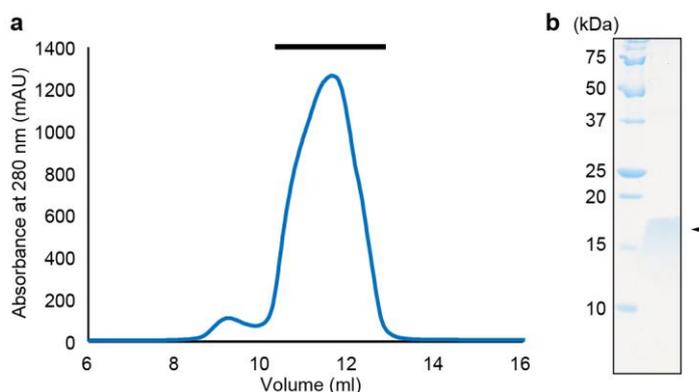
Figure 7 | Purified *E. coli* SemiSWEET

(a) Gel filtration profile of the final sample used for crystallization.

(b) SDS-PAGE gel of (a).
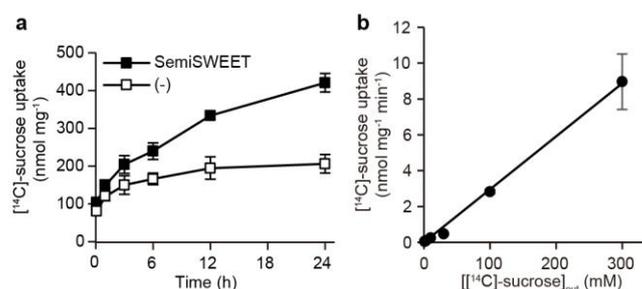
Figure 8 | Functional characterization of *E. coli* SemiSWEET

(a) Time-course of [^{14}C]-sucrose uptake by proteoliposomes containing *E. coli* SemiSWEET (solid black squares) or empty control liposomes (open black squares). Mean ± s.e.m., n = 6.

(b) Plots of the sucrose uptake rate vs. the extra-liposomal sucrose concentration (mean ± s.e.m., n = 3).
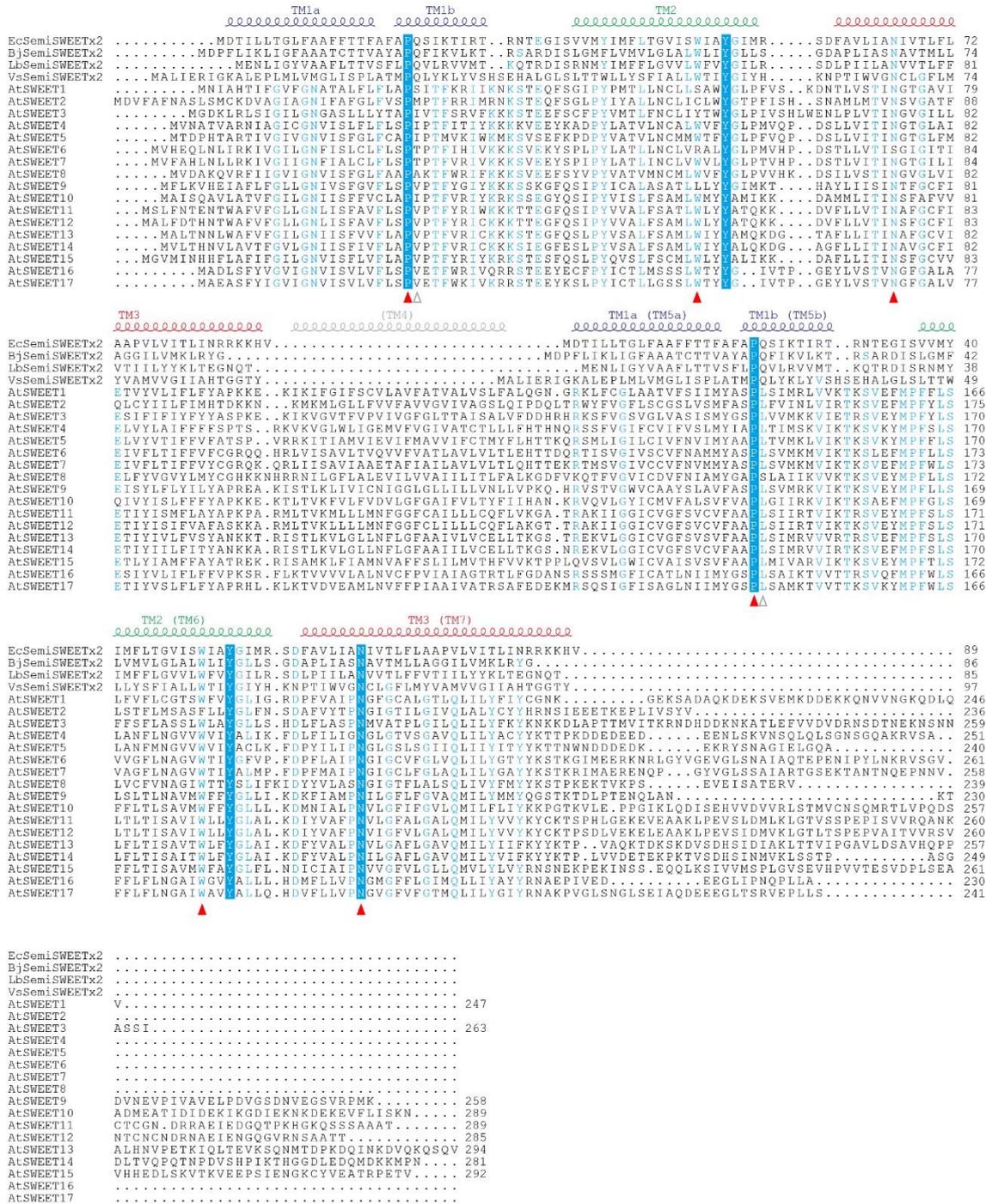
Figure 9 | Sequence alignment of *E. coli* SemiSWEET and other SWEET family proteins

Sequence alignment of *E. coli* (Ec), *B. japonicum* (Bj), *Leptospira biblexa* (Lb), *Vibrio sp.* (Vs) SemiSWEETs and *Arabidopsis thaliana* (At) SWEET1–17, created using ClustalW2 and manually adjusted. SemiSWEET sequences are duplicated to align with two 3-TM domains in SWEETs, and are thus designated as SemiSWEETx2. The conserved functional residues are indicated by red arrowheads, and the glutamine residue in the PQ motif is by a white arrowhead.

2.3.3 Crystallization of *E. coli* SemiSWEET

We performed crystallization trials for purified *E. coli* SemiSWEET by the lipidic cubic phase method (LCP) (*4*). After screening ~600 conditions, we obtained crystals under multiple conditions. Subsequent refinement yielded two types of diffraction-quality crystals, the first one belonging to the space group $P2_12_12$ (designated as Crystal-I) and the second one to $C2$ (Crystal-II) (Figure 10). For Crystal-II, changing the crystallization setup from a typical sandwich-drop method to a hanging-drop method has increased the crystal size dramatically (Figure 10b).

2.3.4 Structure determination

We collected X-ray diffraction data for both crystal forms at the microfocus beamline SPring-8 BL32XU (Figure 11). To determine the initial phases, we prepared the SeMet-labeled crystals in a similar condition to that of Crystal-I (Figure 10c). We collected anomalous diffraction data from these SeMet-labed crystals at a wavelength of 0.9792 Å, which corresponds to the peak wavelength of selenium atoms. By using the SeMet single-wavelength anomalous diffraction (SeMet SAD) method, 9 selenium sites were identified (Figure 12a). The experimental map at 2.6 Å resolution was of sufficient quality for us to build the atomic model *de novo* (Figure 12b). With this model, the native structures of Crystal-I and Crystal-II were built and refined to 2.0 Å and 3.0 Å resolutions, respectively (Figure 12c, d, Figure 13, Figure 14 and Table 2; please also see Section 2.2.7 Data processing and structure determination).
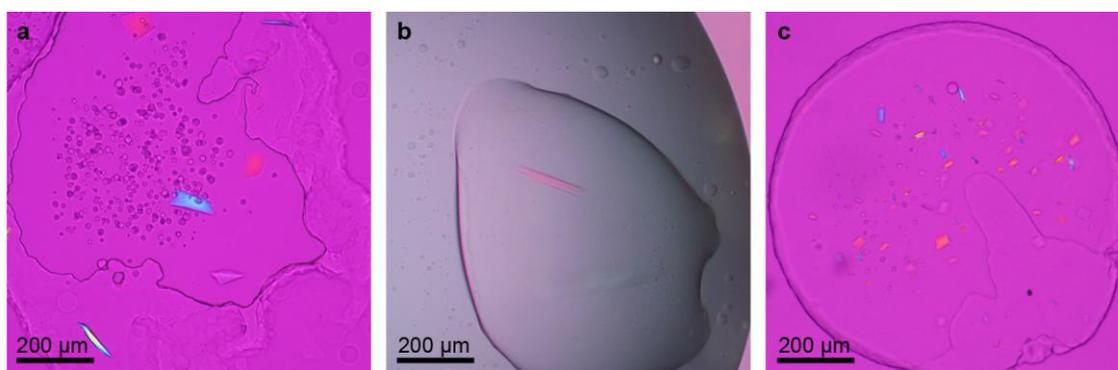
Figure 10 | Crystals of *E. coli* SemiSWEET

(a) Crystal-I. (b) Crystal-II. (c) SeMet-labeled crystal.
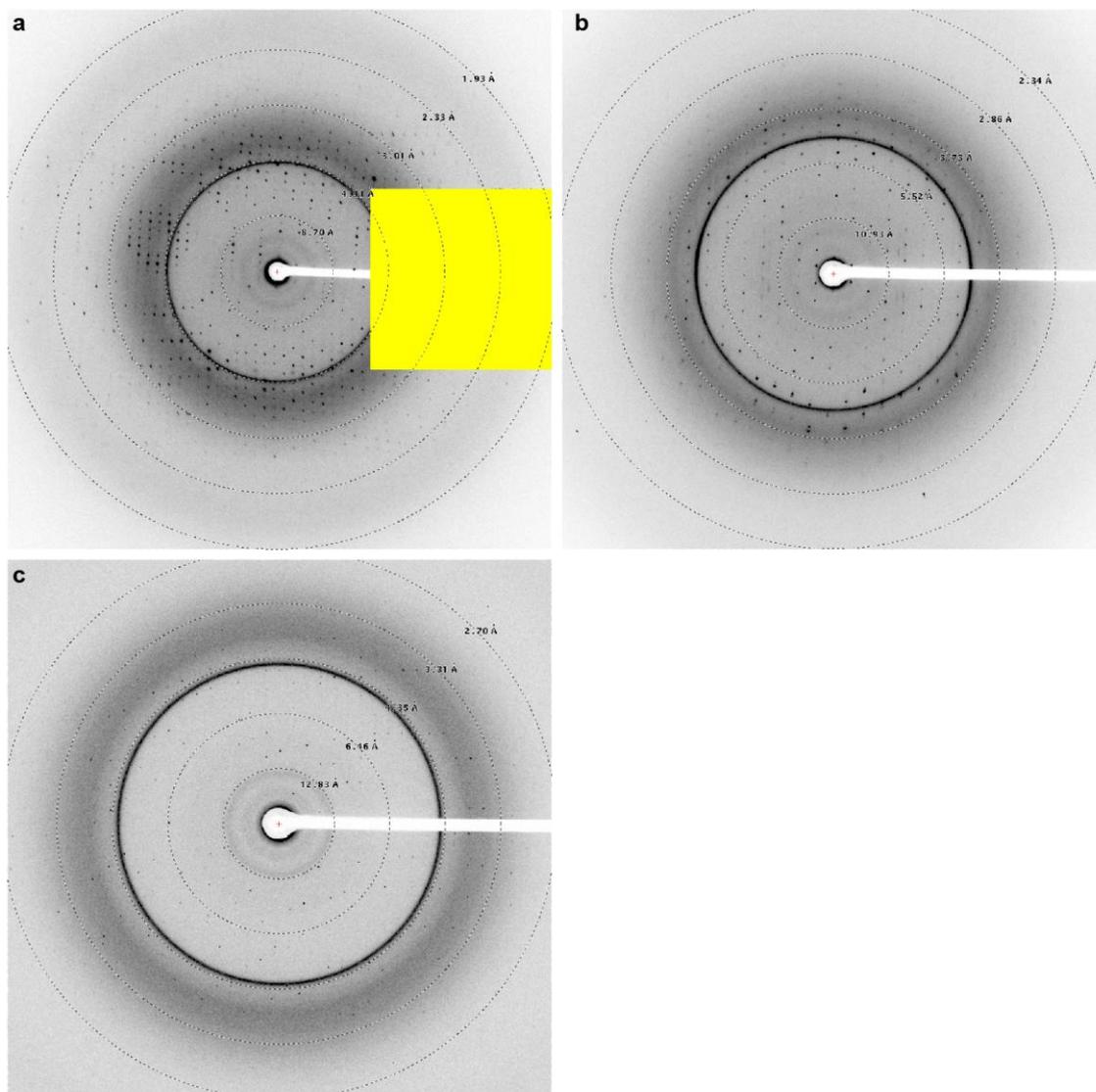
Figure 11 | X-ray diffraction images of SemiSWEET

(a) Diffraction image of Crystal-I. The outermost ring indicates 1.93 Å resolution. Note, a part of the

detector experienced an unexpected failure (yellow). Data from this area were excluded during processing.

(b) Diffraction image of Crystal-II. The outermost ring indicates 2.34 Å resolution.

(c) Diffraction image of the SeMet-labeled crystal. The outermost ring indicates 2.70 Å resolution.
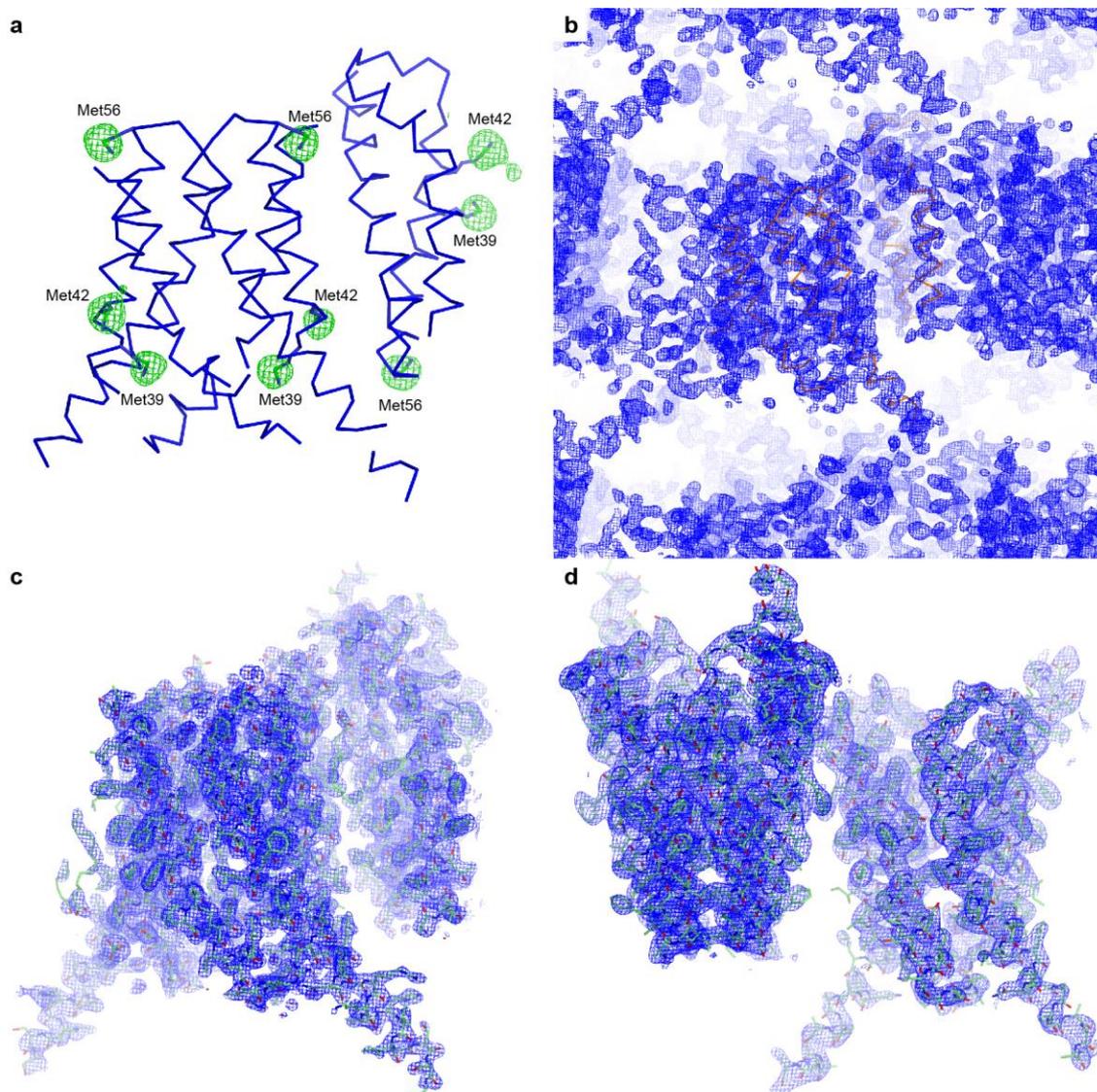
Figure 12 | Structure determination of SemiSWEET by Se-SAD

(a) Anomalous electron density map after Se-SAD in SHELXD, countered at 3.0σ, revealing nine selenium peaks in the crystallographic asymmetric unit. These peaks correspond to the positions of three methionine residues (Met39, Met42 and Met56) in each SemISWEET protomer. The initial model is overlaid as Cα traces.

(b) The experimental electron density map after AutoSHARP, contoured at 1.5σ.

(c, d) The final 2*F*o-*F*c electron density maps for Crystal-I (c) and Crystal-II (d), contoured at 1.0σ.

29
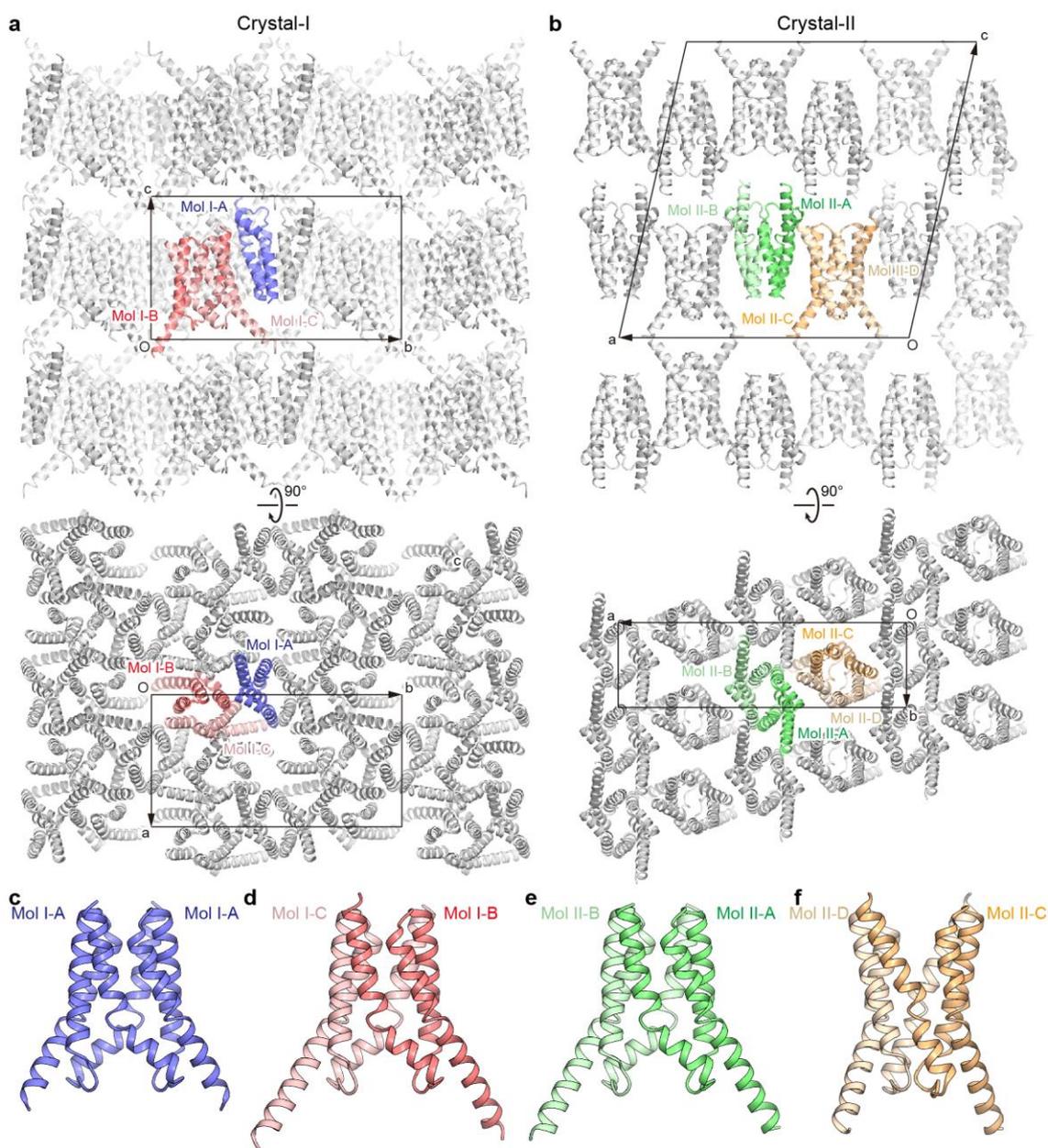
Figure 13 | Crystal packing and dimerization of SemiSWEET

(a) The crystal packing of Crystal-I. Chains A, B and C are labeled as Mol I-A, I-B and I-C, and colored blue, red and pink, respectively. The crystallographic axes are depicted as arrows.

(b) The crystal packing Crystal-II. Chains A, B, C and D are designated as Mol II-A, II-B, II-C and II-D, and colored green, light green, orange and light orange, respectively.

(c–f) Dimeric assemblies of SemiSWEET. The dimers I-A/I-A (c), I-B/I-C (d) and II-A/II-B (e) adopt almost identical conformations (the inward-open state), whereas the dimer II-C/II-D (f) adopts a distinct conformation (the outward-open state).
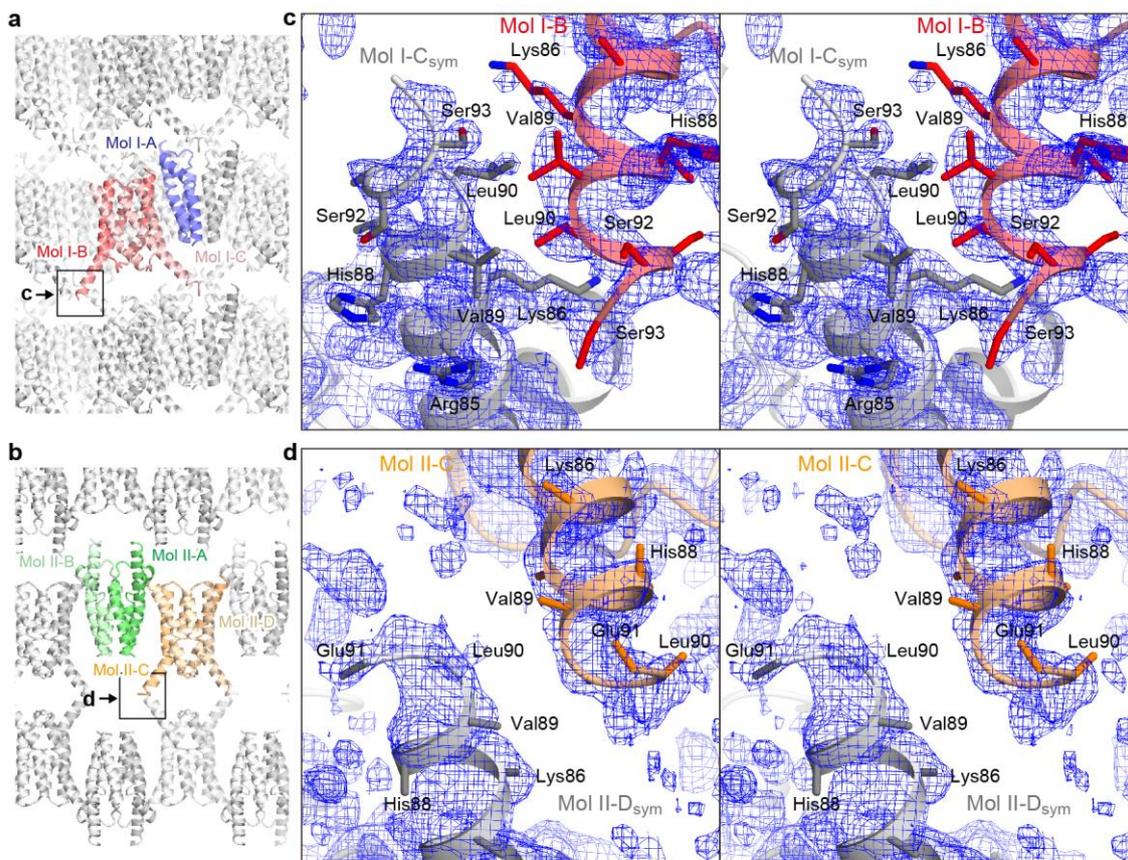
Figure 14 | Crystal packing interactions

(a,b) The crystal packing of Crystal-I (a) and Crystal-II (b).

(c,d) Stereo views of the inter-layer packing interactions observed in Crystal-I (c) and Crystal-II (d). These interactions might have contributed to the formation of two different conformations in our crystals. The residues potentially involved in the interactions are shown as stick models. Residues 90–93 are an artificial linker sequence introduced to the C-terminus as a result of the cloning strategy. The $2F_o$-F_c electron density maps are shown, contoured at 1.0σ. Electron densities for some sidechains are missing, and those sidechains are not modelled.

Table 2 | Data collection and refinement statistics of *E. coli* SemiSWEET

	Native	SeMet[a]	Native
	Crystal-I	Crystal-I	Crystal-II
Data collection			
Space group	*P*2₁2₁2	*P*2₁2₁2	*C*2
Cell dimensions			
a, *b*, *c* (Å)	53.7, 102.1, 59.0	53.8, 101.0, 58.65	118.0, 34.6, 123.2
σ, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 102.9, 90.0
		Peak	
Wavelength (Å)	1.0000	0.9792	1.0000
Resolution (Å)	50.0–2.00 (2.07–2.00)[b]	50.0–2.60 (2.69–2.60)	50.0–3.00 (3.11–3.00)
R_{pim}	0.032 (0.431)	0.049 (0.693)	0.072 (0.480)
$CC_{1/2}$	0.999 (0.688)	0.986 (0.094)	0.993 (0.655)
$I / \sigma I$	17.7 (2.0)	15.1 (1.4)	8.2 (1.7)
Completeness (%)	99.9 (100.0)	99.9 (100.0)	98.0 (98.4)
Redundancy	19.5 (18.0)	46.8 (31.3)	3.3 (3.2)
Refinement			
Resolution (Å)	50–2.00		50–3.00
No. reflections	22,575		9,943
R_{work} / R_{free} (%)	19.7 / 22.4		28.1 / 32.8
No. atoms			
Protein	2,077		2,679
Lipid/ion	259		
Water	66		
B-factors			
Protein	36.02		67.17
Lipid/ion	55.07		
Water	47.37		
R.m.s deviations			
Bond lengths (Å)	0.0024		0.0035
Bond angles (°)	0.639		0.713
Ramachandran plot			
Favored (%)	100.0		97.8
Allowed (%)	0.0		2.2
Outliers (%)	0.0		0.0

[a]SeMet data were collected from six crystals and others were from one crystal.

[b]Values in parentheses are for highest-resolution shell.

2.3.5 Overall structure of SemiSWEET

The crystallographic asymmetric unit of Crystal-I contains three SemiSWEET molecules. Two molecules assemble to form a dimer, and the other molecule forms a similar dimer with its crystallographic symmetry molecule related by a two-fold axis (Figure 13). As the two dimers are structurally almost identical, with a root mean square deviation (r.m.s.d.) value of 0.55 Å over all Cα atoms, the non-crystallographic dimer of SemiSWEET will be discussed here. The two protomers are arranged with identical membrane topologies, with the central two-fold axis perpendicular to the membrane (Figure 15a). Each protomer comprises three transmembrane (TM) helices (TM1, TM2 and TM3). TM1 is largely kinked in the middle and separated into two segments (TM1a and TM1b). The short TM2 helix is almost entirely buried inside the membrane, while the long TM3 helix protrudes into the aqueous environment on the intracellular side. A total of 6 TM helices from two protomers create a central cavity that is widely open to the intracellular side (Figure 15a). This cavity penetrates into the dimer core, but is completely occluded from the extracellular side and the lipid environment by the tight association of the extracellular regions and the surrounding TM helices. Hence, we designated this structure as the inward-open conformation.

In Crystal-II, the asymmetric unit contains four SemiSWEET molecules organized as two separate dimers (Figure 13). Whereas the conformation of one dimer is almost identical (r.m.s.d.) to that of the inward-open state in Crystal-I, the other dimer adopts a markedly different conformation. In this alternative conformation, the extracellular halves of the protomers are separated from each other, while the intracellular halves approach the central axis (Figure 15b). Consequently, the central cavity formed by the dimer is open toward the extracellular side. Hence, we designated this structure as the outward-open state.
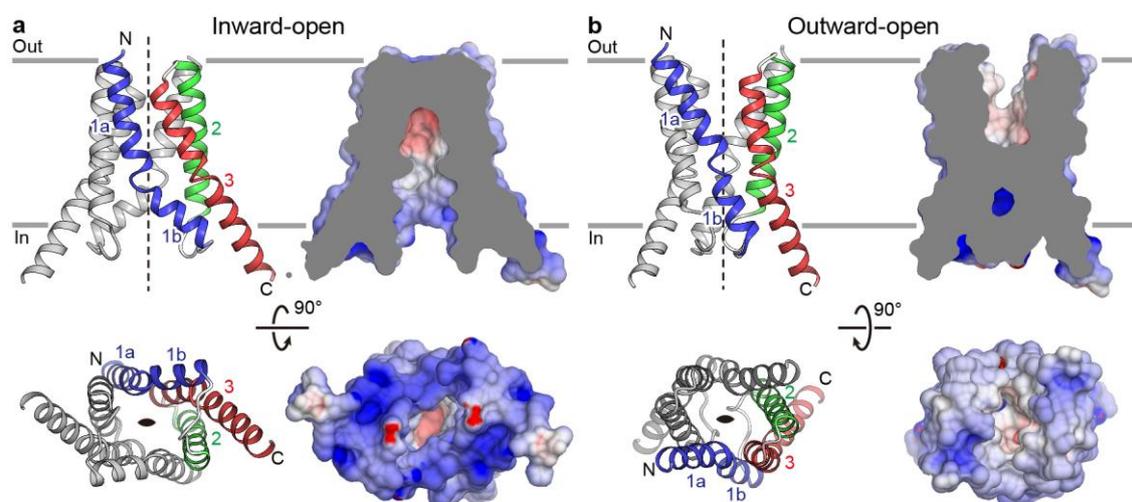
Figure 15 | Structure of SemiSWEET

(a) Overall structure of the inward-open SemiSWEET, viewed parallel to the membrane (upper) or from the intracellular side (lower). TM1, TM2 to TM3 of one protomer is colored blue, green and red from, and the other protomer is in gray. Surface representations are colored according to electrostatic potential. The two-fold axis is indicated by dashed lines and an almond-shaped symbol.

(d) Overall structure of the outward-open SemiSWEET dimer, viewed parallel to the membrane (upper) and from the extracellular side (lower), colored as in (c).

2.3.6 Substrate-binding site

The largely open cavities on the intracellular and extracellular sides in the inward-open and outward-open conformations, respectively, suggested that the substrate sugar is translocated along the central axis of the SemiSWEET dimer, accompanied by a conformational change. The high resolution structure of Crystal-I revealed the presence of a monoolein molecule occupying the cavity of the inward-open state (Figure 16 and Figure 17). The glycerol head group of the monoolein molecule is located in the pocket formed by symmetrically arranged residues in the dimer core. Within the pocket, the hydroxyl of the glycerol head group forms direct and water-mediated hydrogen bonds with the side chains of Asn66 (Figure 16c,e). In addition, the glycerol head group is sandwiched by the aromatic rings of the Trp50 side chains, which are stabilized by the side chains of Thr15 (Figure 16c,e). Given that glycerol and sugar share similar polyol moieties, the glycerol head group of the bound monoolein molecule is likely to be mimicking the sugar substrate. This pocket has a width of approximately 8–9 Å and a length of 11 Å along the central axis, and is suitable for accommodating a monosaccharide or disaccharide (Figure 16a,b). Furthermore, in the outward-open state, this pocket is exposed to the extracellular environment, while the arrangement of the lining residues is preserved (Figure 16b,d,f). Together, these observations implicated this central pocket as the binding site for transport sugars.

To investigate the functional role of this putative substrate-binding pocket, we introduced mutations to Trp50 and Asn66, and measured the sucrose uptake activities of these mutants (Figure 16f). The N66A mutation significantly decreased the sucrose uptake activity, suggesting that the hydrophilic moiety of the Asn66 side chain is important for sucrose binding (Figure 16g). In contrast, the W50F mutation did not affect the sucrose uptake, and the W50A mutation actually greatly increased the sucrose uptake activity, suggesting that Trp50 is not essential for sucrose binding (Figure 16g). This result is inconsistent with the recent report showing the crucial role of the equivalent tryptophan residue of *Arabidopsis thaliana* SWEET1 in glucose transport (*98*). This might be due to the different size of the substrates used for the transport assays. As the disaccharide sucrose has a larger molecular size than the monosaccharide glucose, the increased sucrose uptake activity of the SemiSWEET W50A mutant might be attributed to the enlargement of the pocket, by the replacement of the bulky Trp residue with a smaller Ala residue, which likely promotes sucrose entry into the pocket. Taken together, both our results and the previous studies suggest that the transported sugars are accommodated in the central pocket lined by Trp50 and Asn66.

As compared to other sugar transporters such as human GLUT1 (*99*, *100*), SemiSWEET has fewer hydrogen-bonding residues within the substrate-binding pocket. This feature could reflect the low-affinity transport by SemiSWEET as observed in our liposome assay (Figure 8). The Trp50 and Asn66 residues are highly conserved among the SemiSWEET and SWEET families (Figure 9), suggesting that the two families share substrate-binding pockets with similar architectures. Therefore, our structural and functional analyses provide insight into the low-affinity transport by the SWEET transporters.
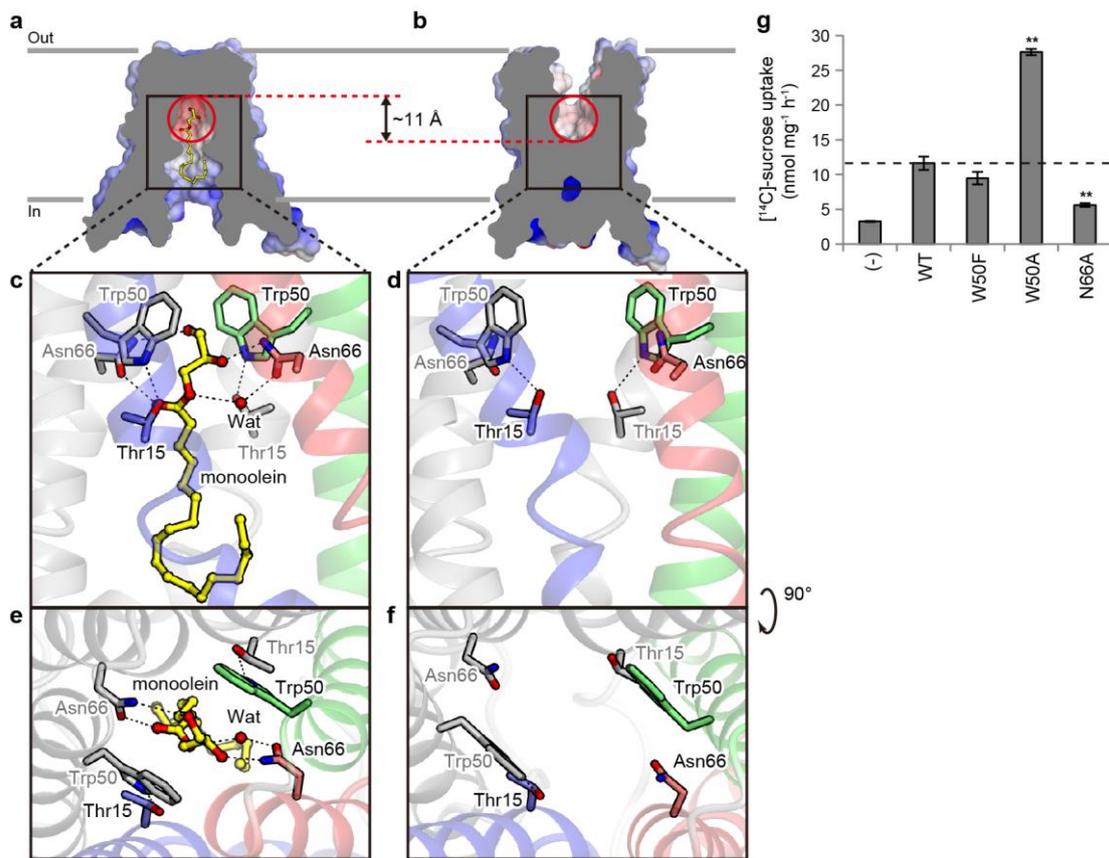
Figure 16 | Putative substrate-binding site of SemiSWEET

(a, b) Cut-away surface representations of the inward- (a) and the outward-open (b) structures. The position of the putative substrate-binding site is indicated by red circles. The monoolein molecule is shown as a ball-and-stick model.

(c–f) Close-up views of the substrate-binding site in the inward- (c,e) and the outward-open (d,f) structures. Hydrogen bonds are depicted as black dotted lines.

(g) Sucrose uptake by the SemiSWEET mutants in the liposome assay (mean ± s.e.m., n = 3). (-) is for empty control liposomes. Significant differences from the wild-type value (WT) are indicated by asterisks (**P < 0.01, Student's t-test).
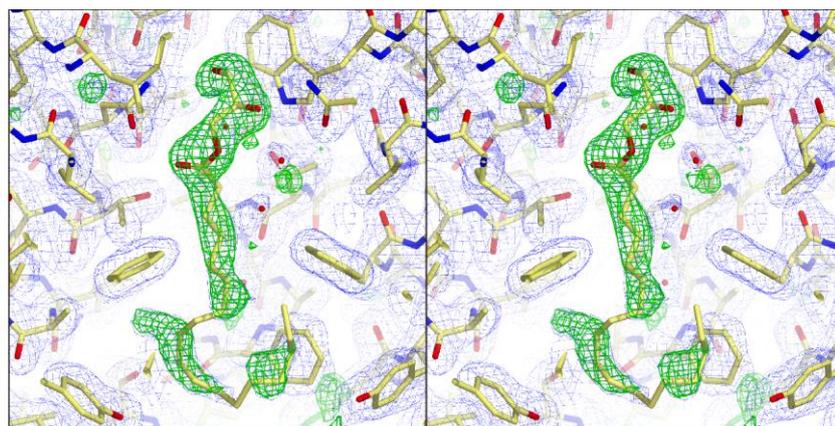
Figure 17 | Electron density of lipid monoolein

Stereo view of the omit electron density map, calculated without the monoolein molecule in the pocket. The Fo-Fc map is shown in green, contoured at 3.0σ, and the 2Fo-Fc map is in blue, contoured at 1.0σ.

2.3.7 The role of conserved PQ-loop motif

A structural comparison of the individual protomers in the inward-open and outward-open conformations revealed an approximately 30-degree kink between TM1a and TM1b, at Pro21 in the PQ-loop motif (Figure 18a). Furthermore, a structural comparison of the dimers in the two distinct conformations revealed the relative rotational motion of the two symmetrical helix bundles, consisting of TM1b, TM2, and TM3 from one protomer and TM1a' from the other protomer (Figure 18b). Each helix bundle from the two conformations can be superimposed well, suggesting their rigid-body movement (Figure 18c). This conformational change resembles a 'binder clip', in which the opening of the extracellular cavity is coupled with the closing of the intracellular cavity, with the kink at the PQ-loop motif serving as a molecular hinge.

In addition, Gln22 in the PQ-loop motif interacts with the main chain atoms of the residues on the intracellular loop connecting TM1 and TM2, at the position immediately next to TM2 (Figure 18d,e). In the outward-open state, the oxygen atom of the Gln22 side chain hydrogen bonds with the backbone amide group of Ser36 (Figure 18e). In the inward-open conformation, although the side chain of Gln22 is not within hydrogen-bonding distance with the equivalent residue, it still points toward the backbone carbonyl of Gly34 on the same loop and retains within a close distance (Figure 18d). In contrast, the cytoplasmic end of TM1b moves largely apart from the adjacent protomer. The Cα distance between Asn31 of one protomer and Gly34 of the adjacent protomer is about 10 Å longer in the inward-open conformation than in the outward-open conformation. These observations suggest that Gln22 may stabilize the hinge by bridging the two helix bundles to allow for a dynamic structural change in TM1b and the following cytoplasmic loop.

To verify the functional importance of the PQ-loop motif, we introduced alanine mutations to Pro21 and Gln22, and then measured the sucrose uptake by these mutants (Figure 18f). The P21A mutation, which was expected to decrease the conformational flexibility of TM1, significantly decreased the sucrose uptake activity almost to the level of the control empty liposomes, thus demonstrating the essential role of this conserved proline residue in sucrose transport. In contrast, the Q22A mutant showed slightly decreased sucrose uptake activity as compared to that of the wild type, indicating the less important role of Gln22 in sucrose transport, which is consistent with the fact that this glutamine residue is not conserved in the SWEETs (Figure 9). Together, our structural and functional analyses indicated that the proline and glutamine residues in the PQ-loop motif serve as a flexible hinge, thereby enabling the binder clip-like motion of SemiSWEET. Because this proline residue is highly conserved in the SWEETs and the PQ-loop family transporters (*68*, *101*) and its significance for the transport function in vivo has been demonstrated (*102*, *103*), the binder clip-like motion of SemiSWEET is likely to be conserved in all of these

transporters.
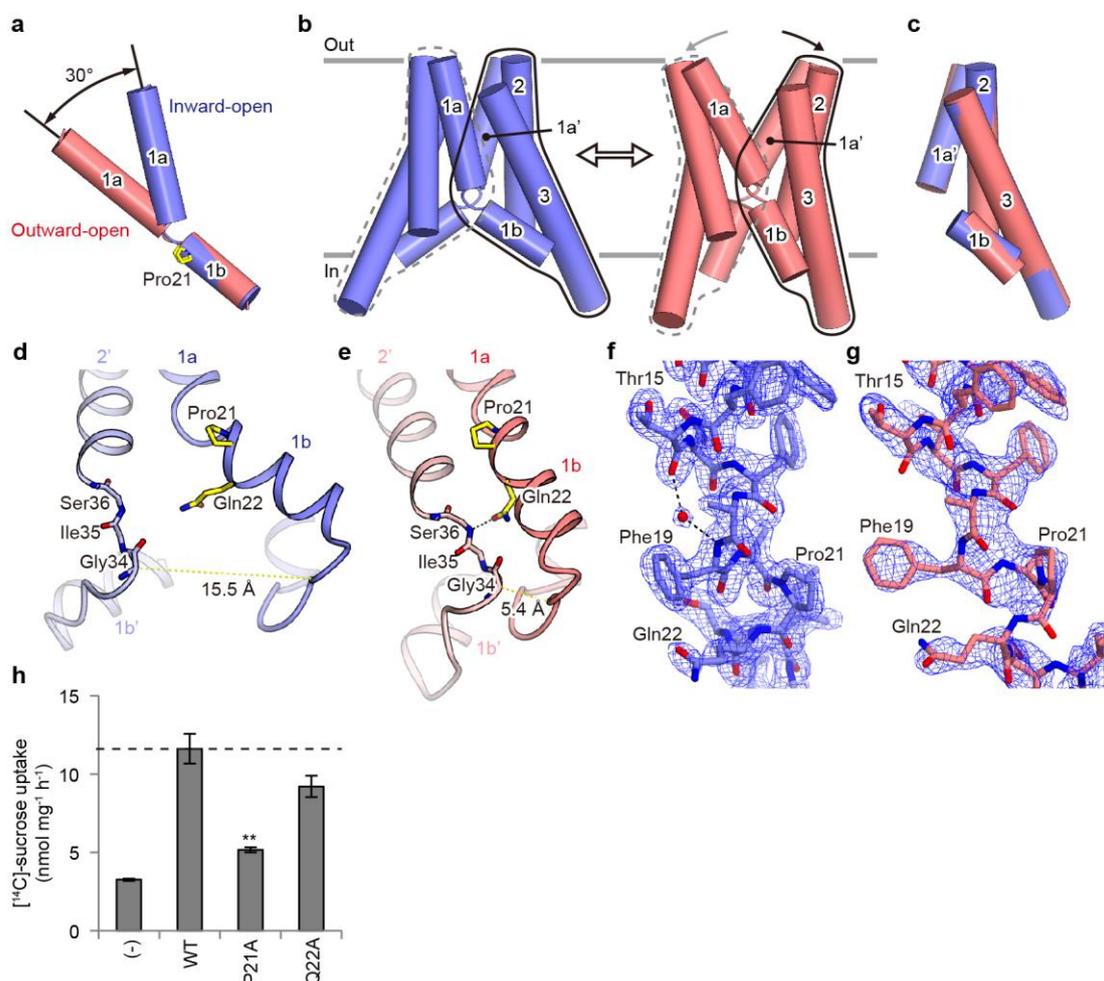
Figure 18 | The PQ-loop motif

(a) Superimposition of TM1 between the two conformations. The inward- and outward-open conformations are colored blue and pink, respectively.

(b) Structural comparison of the two conformations. The structural unit consisting of TM1b, TM2, TM3 and TM1a' (from the adjacent protomer) is enclosed by black lines, and its counterpart by gray dotted lines.

(c) Superimposition of the structural units between the two states, colored as in (a).

(d, e) Close-up views of the PQ-loop motif. The hydrogen bond is depicted as a black dotted line.

(f, g) The $2F_o$-F_c electron density maps around the PQ-loop motif, contoured at 1.5σ. In the inward-open conformation, a water molecule mediates a hydrogen-bonding interaction in the backbone.

(h) Sucrose uptake by SemiSWEET mutants in the liposome assay (mean ± s.e.m., n = 3). (-) is for empty control liposomes. A significant difference from the wild-type value (WT) is indicated by asterisks (**P < 0.01, Student's t-test).

2.3.8 Extracellular and intracellular gates

The structures of the inward-open and outward-open conformations revealed the formation of two distinct gates on the extracellular and intracellular sides, which restrict the accessibility of the substrate-binding pocket (Figure 19a,b). The extracellular gate is formed by the amino acid residues on a loop connecting TM2 and TM3. In the inward-open conformation, Tyr53, Arg57 and Asp59 interact with the equivalent residues of the adjacent protomer; Asp59 forms a salt bridge with Arg57 and hydrogen bonds with Tyr53, and the Arg57 side chain hydrogen bonds with the Arg57 main-chain carbonyl group (Figure 19c). These interactions completely seal off the substrate-binding pocket from the extracellular environment. In contrast, in the outward-open conformation, these interactions are not observed, due to the outward movements of TM2 and TM3 (Figure 19d).

On the opposite side of the membrane, the intracellular gate is formed by hydrophobic residues on TM1 and TM2 (Figure 19e,f). In the outward-open conformation, Phe19, Met39, Tyr40 and Phe43 of the two protomers form a cluster of aromatic and hydrophobic residues through van der Waals interactions (Figure 19f). These interactions seal off the substrate-binding pocket from the intracellular environment. In contrast, in the inward-open state, these hydrophobic residues are separated from each other, thereby creating a cytoplasmic cavity that allows substrate access to the substrate-binding pocket (Figure 19e). Overall, these observations suggested that the extracellular and intracellular gates restrict the central translocation pathway, and determine the accessibility of the substrate-binding pocket.

The binder clip-like motion of SemiSWEET implied that the opening and closing of the intracellular and extracellular gates are closely related. To investigate the functional significance of these gates, we created various mutants of the residues constituting these gates, and measured their sucrose transport activities (Figure 19g). Y53F and R57A, which would disrupt the hydrogen-bonding or salt bridge interactions in the extracellular gate, showed significantly increased activities. In contrast, F19A and Y40A, which would weaken the hydrophobic interactions in the intracellular gate, showed significantly decreased activities. These results revealed that the defects in the extracellular and intracellular gates have the opposite effects on the sucrose transport. These opposite effects could be explained on the basis of the different preferences of the two conformations of SemiSWEET, as follows. The temperature factors for the two conformations in Crystal-II implied that the outward-open conformation is inherently less stable than the inward-open conformation, suggesting that SemiSWEET probably prefers the inward-open conformation in the lipid environment. Therefore, the mutation in the extracellular gate may result in the release of the extracellular 'lock', which preferentially captures

SemiSWEET in the inward-open state. The observed increased sucrose uptake activities of the mutants are likely due to the facilitated state transition between the inward-open and outward-open conformations, by disrupting the interactions in the extracellular gate. In contrast, the defects in the intracellular gate may arrest SemiSWEET in the rather stable inward-open state, thus resulting in the lower sucrose uptake activities.
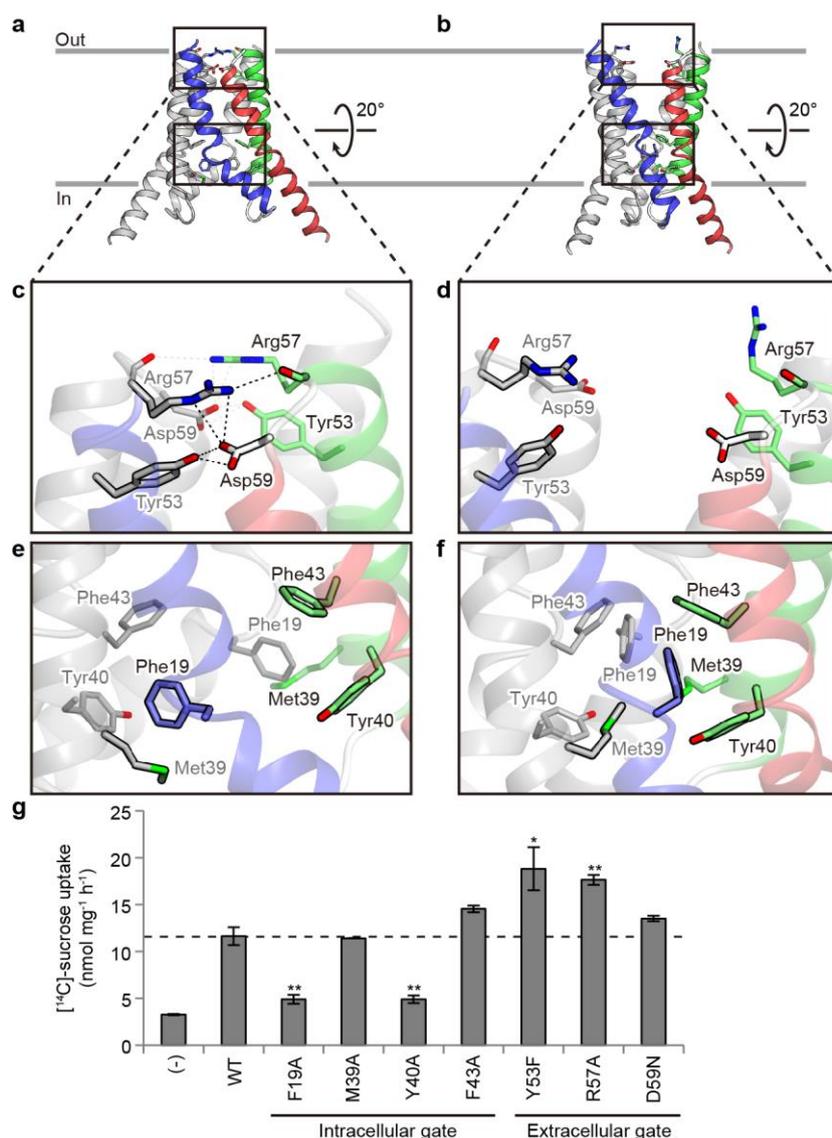
Figure 19 | Extracellular and intracellular Gates

(a, b) Ribbon representations of the inward- (a) and outward-open (b) states. The residues constituting the extracellular and intracellular gates are shown as stick models.

(c, d) Close-up views of the extracellular gate in the inward- (c) and outward-open (d) states. The hydrogen-bonding and salt bridge interactions are depicted as black dotted lines.

(e, f) Close-up views of the intracellular gate in the inward- (e) and outward-open (f) states. (g) Sucrose uptake by SemiSWEET mutants in the liposome assay (mean ± s.e.m., n = 3). (-) is for empty control liposomes. Significant differences from the wild-type value (WT) are indicated by asterisks (*P < 0.05, **P < 0.01, Student's t-test).

2.3.8 Occluded SemiSWEET structures from different species

During the course of this study, we noticed two papers that reported the structures of SemiSWEETs from different species, namely from *Leptospira biflexa*, *Vibrio sp.* and *Thermodesulfovibrio yellowstonii* (*109*, *110*) . These structures were resolved in the novel occluded conformation, as well as in the outward-open conformation that resembles our structure. Notably, in the occluded conformation of *L. biflexa* SemiSWEET, both the extracellular and intracellular gates are 'closed' (Figure 20). A structural comparison of this occluded conformation of the *L. biflexa* SemiSWEET with the present inward- and outward-open conformations of *E. coli* SemiSWEET and revealed a slight bend at TM2 in the *L. biflexa* SemiSWEET (Figure 20a). This bend allows the closing of the both extracellular and intracellular gates in the occluded conformation (Figure 20b). This observation suggested that the helix bundle constituting each piece of the 'binder clip' does not move as an exact rigid body, but it may allow a slight bend in TM2. The extracellular and intracellular gates might associate with each other through this elasticity.
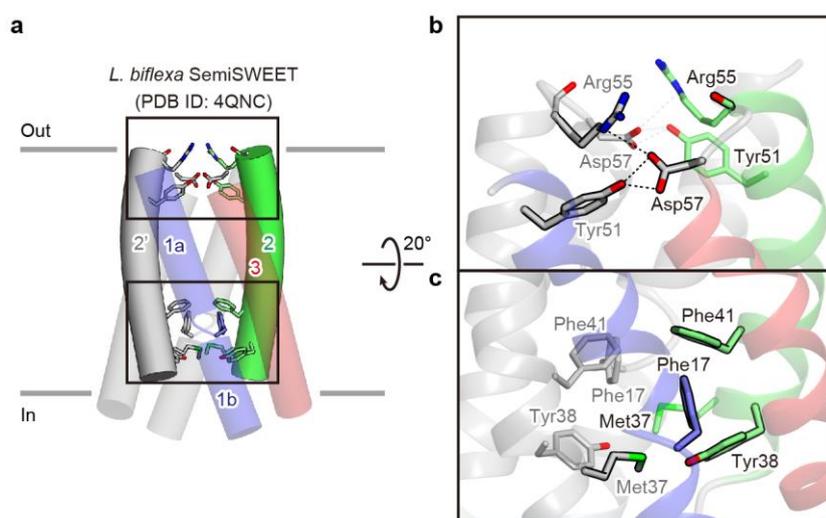
Figure 20 | Occluded conformation of *L. biflexa* SemiSWEET.

(a) Crystal structure of *L. biflexa* SemiSWEET (PDB 4QNC), showing the slight bending of TM2.

(b) Close-up view of the extracellular gate of *L. biflexa* SemiSWEET. The hydrogen-bonding and salt bridge interactions are depicted as black dotted lines.

(c) Close-up view of the intracellular gate of *L. biflexa* SemiSWEET.

2.4 Discussion

The structural and functional analyses performed here allowed us to propose a possible sugar transport mechanism of SemiSWEET (Figure 21). In the present structure in the inward-open conformation, the putative substrate-binding pocket is occupied by the monoolein molecule mimicking the substrate sugar, indicating that the observed structure represents an inward-open, substrate-binding state (Figure 21a). The PQ-loop motif serves as a hinge that enables the binder clip-like motion of SemiSWEET to transit toward the outward-open state. Along with this transition, the slight bend of TM2 might close the intracellular gate, leading to the occluded state (Figure 21b), while the closely associated opening/closing of the gates would prevent the formation of an open channel. Subsequently, the opening of the extracellular gate allows the substrate to exit (Figure 21c). Since the inward-open and outward-open conformations were simultaneously captured in Crystal-II in the absence of any sugars (Figure 13), the transition between these two conformations could spontaneously occur even without any substrates (Figure 21d). These sequential conformational changes allow the small amphipathic molecules that can be accommodated in the central pocket to permeate through the membrane, while preventing the leakage of the larger molecules or ions.

In summary, the current structures and functional analyses revealed the molecular detail of "alternating-access" by the SWEET and PQ-loop transporters mediated by the binder clip-like motion. The mechanistic insight presented here will aid further experiments towards understanding the substrate selectivity, transport kinetics, and regulatory mechanism of the plant SWEETs, and the malfunction of the human PQ-loop transporters.
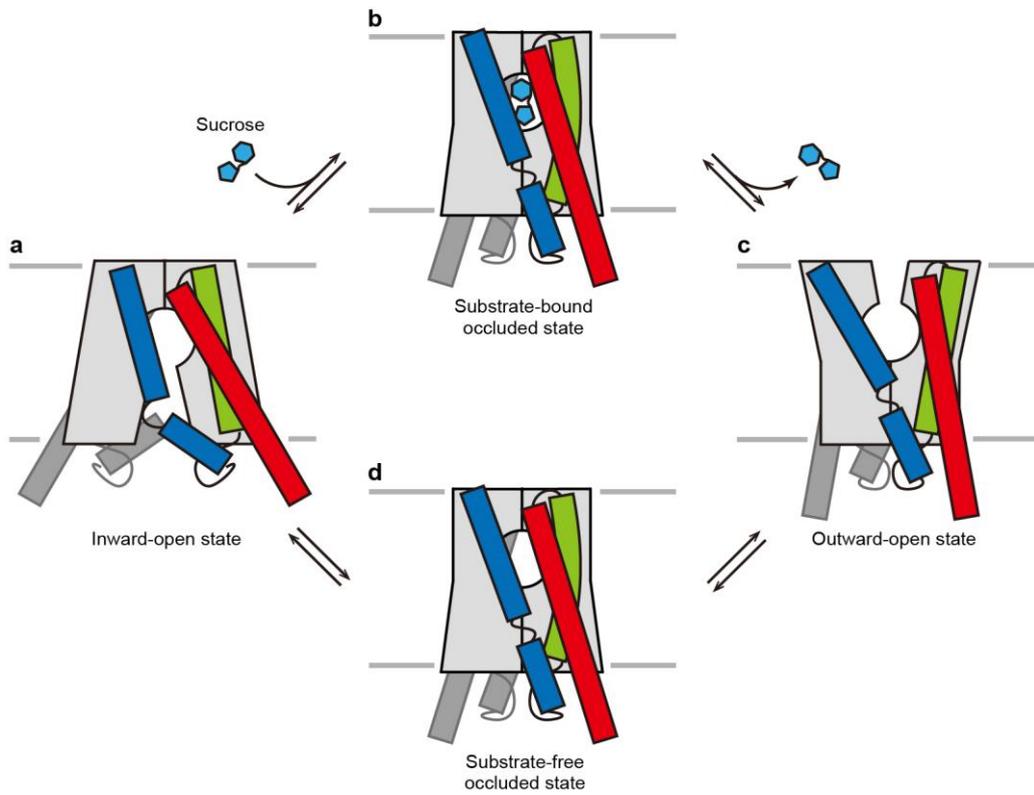
Figure 21 | Transport mechanism of SemiSWEET

(a–d) Schematics of the SemiSWEET dimer in the inward-open (a), substrate-bound occluded (b), outward-open (c) and substrate-free occluded (d) states, viewed parallel to the membrane. (a) and (c) are drawn based on the present crystal structures.

Chapter 3 X-ray crystallographic analysis of TPT

3.1 Introduction

As discussed in Chapter 1, the triose-phosphate/phosphate translocator (TPT) plays a central role in the transport of Calvin-Benson cycle intermediates from chloroplasts to cytosol during photosynthesis. Discovered more than 40 years ago (*70*), TPT is the most rigorously characterized transporter of the chloroplast at both genetic and biochemical levels (*111*, *112*). TPT is also the most abundant protein on the chloroplast envelope membranes, constituting up to 15% of total proteins, highlighting its important role in plastid function (*113*).

There has been a long-standing debate regarding the structure of TPT. Previously proposed models include the 6-TM, 7-TM, 8-TM and 9-TM topology models, which are structurally diverse (*75*). A biophysical study has suggested that TPT forms a dimer within the membrane, and has one substrate-binding site per dimer (*114*, *115*). Figure 22 shows a proposed structural model of spinach TPT (*116*, *117*) , in which two 6-TM protomers are arranged as a dimer with the two-fold symmetry, creating a substrate-translocation pore. However, without the three-dimensional structure, it has been difficult to understand the mechanisms for substrate recognition and transport. Therefore, to fully understand the molecular mechanisms of TPT, I undertook structural analyses on TPT.
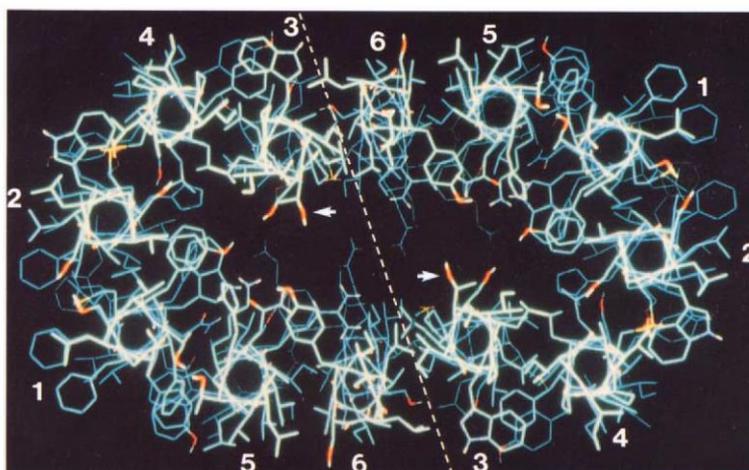
Figure 22 | Proposed structural model of TPT

Top view of the predicted structural model of spinach TPT. Numbers refer to the TM helices of each protomer. The positions of Thr139 are indicated by arrows. Reproduced from ref. (*117*).

3.2 Materials and methods

3.2.1 Cloning

The total cDNAs from plants were purchased from ZYAGEN. The cDNAs of *Cyanidioschyzon merolae* and *Galdieria sulphuraria* were provided by Dr. Ayumi Minoda at Tsukuba University. The genes encoding pPT family proteins were amplified from these plant and algal cDNAs with specific primers. The nucleotide regions coding for putative chloroplast transit peptides (cTPs; about 80–100 amino acid residues) were removed during cloning. The amplified gene fragments were ligated into a modified pYES2 vector with In-Fusion method following manufacturer's protocols. This vector was designed to add a tobacco etch virus (TEV) protease cleavage site, an enhanced green fluorescent protein (EGFP) and a His$_8$-tag at the C-terminus of the target protein. The species names, gene accessions, amplified target regions and the specific primers used for cloning were summarized in Table 3.

Table 3 | Primers used for cloning pPTs

No.	Species	Protein	Region	Primer sequences (5'–3') *
1	*Arabidopsis thaliana*	At5g46110 (AtTPT)	88–410	AACGAGGAATTCATGGGTGGAGATACCGCTGGGG
				CAAGTTTTCGGTACCTGCTTTCTTTCCTTGCCGTTTCT
2	*Arabidopsis thaliana*	At5g33320 (AtPPT1)	87–408	AACGAGGAATTCATGACTGCAGTTCCTGAAAGTGCTGAG
				CAAGTTTTCGGTACCAGCAGTCTTTGGCTTTGGCTTAATA
3	*Arabidopsis thaliana*	At3g01550 (AtPPT2)	56–383	AACGAGGAATTCATGTCAGCTACAGTTCCGGAGAATGTG
				CAAGTTTTCGGTACCAGACATTTTTGGATTTGGTTTGACTTG
4	*Arabidopsis thaliana*	At5g54800 (AtGPT1)	73–388	AACGAGGAATTCATGTCAGAGCCTCATCCGATCGG
				CAAGTTTTCGGTACCGAGCTTTGCCTGGGAATACAAGAAG
5	*Arabidopsis thaliana*	At1g61800 (AtGPT2)	77–388	AACGAGGAATTCATGTCCCGTCCACTGGACATCAAC
				CAAGTTTTCGGTACCCTGCTTCGCCTGTGAGTAGAGGA
6	*Arabidopsis thaliana*	At5g17630 (AtXPT)	89–417	AACGAGGAATTCATGTCAAATCCAGACGAAAAATCCGAT
				CAAGTTTTCGGTACCGTTCTTCTTATCACCTCCCACTTCA
7	*Oryza sativa*	Os01g13770 (OsTPT1)	100–417	AACGAGGAATTCATGGGGGGAAGCGAAGCCAGCG
				CAAGTTTTCGGTACCTGCACTCTTCGCCCTTTTCTCCTC
8	*Oryza sativa*	Os05g15160 (OsTPT2)	86–404	AACGAGGAATTCATGGGGGGAAGCAAAGCCTGTGGG
				CAAGTTTTCGGTACCTGCGCTCTTCATTTGCGTTTTCTC
9	*Oryza sativa*	Os09g12600 (OsPPT1)	92–408	AACGAGGAATTCATGGGTGAGGCAGGGGCGGAG
				CAAGTTTTCGGTACCGGCAGTCTTGGGCTTGGGTTTAAG
10	*Oryza sativa*	Os08g25624 (OsPPT2)	90–407	AACGAGGAATTCATGGGGGACGCCAAGGCGG
				CAAGTTTTCGGTACCCGCAGCCTTGGGCTTGG
11	*Oryza sativa*	Os01g07730 (OsPPT3)	83–393	AACGAGGAATTCATGGGGCGGGAGAGAGGCG
				CAAGTTTTCGGTACCGGCATTCTTTGGTTTTGTTCTCTTC
12	*Oryza sativa*	Os05g07870 (OsPPT4)	101–404	AACGAGGAATTCATGGGCGGCGCCGTCG
				CAAGTTTTCGGTACCTGCAGTCTTAGCCTTTGGTTTAGCT
13	*Oryza sativa*	Os08g08840 (OsGPT1)	79–387	AACGAGGAATTCATGTCTAAGACCGAGGTGGTGCCC
				CAAGTTTTCGGTACCCTGCTTTGCCTGAGAATACAGGAAT
14	*Oryza sativa*	Os07g34006 (OsGPT2)	79–392	AACGAGGAATTCATGGGCGCCCGGCCGGT
				CAAGTTTTCGGTACCCTGCTTCGCCTGAGAGTAGATGAAA
15	*Solanum lycopesicum*	Solyc01g081390	96–422	AACGAGGAATTCATGGGAACCCCAGAAGAAGTGAGTCC
				CAAGTTTTCGGTACCGTCCTTCTTTTCCACTGCCTCTTTC
16	*Solanum lycopesicum*	Solyc02g086650	77–397	AACGAGGAATTCATGGGTGAATCTGAGGTTGTGAAACCC
				CAAGTTTTCGGTACCCGCATCTTTTGGCTTCCCC
17	*Solanum lycopesicum*	Solyc03g112870	94–410	AACGAGGAATTCATGAGCGCTGGAGAGGCACCC
				CAAGTTTTCGGTACCTTCTGTTTTTGCCTTGGCCTTAATG
18	*Solanum lycopesicum*	Solyc07g064270	88–395	AACGAGGAATTCATGCCGATGGAAGGACCTGAATCG
				CAAGTTTTCGGTACCTTGTTTTGCCTGTGAGTACAAGAAAGTT

*The overlap regions used for In-Fusion reactions are underlined.

Table 3 | Continued

No.	Species	Protein	Region	Primer sequences (5'–3')
19	*Nicotiana tobacum*	CAA52979	80–401	AACGAGGAATTCATGGGCAGCGATTCCGCCG
				CAAGTTTTCGGTACCGGCAGCTTTCTTTTGCCTTTTCTCT
20	*Nicotiana tobacum*	AAB40648	77–385	AACGAGGAATTCATGACATCCGTGCCCGAGAGTGC
				CAAGTTTTCGGTACCTTCCGTTTTTGCCTTAGGTTTAATG
21	*Nicotiana tobacum*	AAB40647	79–398	AACGAGGAATTCATGACGTCGTCGGAGCCCGA
				CAAGTTTTCGGTACCAGCTGTCTTTGGCTTAGGTTTGATG
22	*Triticum aestivum*	AAK01174	80–402	AACGAGGAATTCATGTCCGCCGAGCCCGCC
				CAAGTTTTCGGTACCCGCGGCCTTCGCCCTT
23	*Triticum aestivum*	AAP80864	77–385	AACGAGGAATTCATGTCAAAGGCTGAGGTGTTGCCAG
				CAAGTTTTCGGTACCCTGCTTTGCCTGAGAGTACAGGAAT
24	*Triticum aestivum*	CDM81988	79–398	AACGAGGAATTCATGGCACCGGCGTCGGACG
				CAAGTTTTCGGTACCTGAGCTCTTTGGCTTTGTTTTCTTG
25	*Glycine max*	XP_003542721	81–406	AACGAGGAATTCATGGGCAGTGATTCCGCCGGAG
				CAAGTTTTCGGTACCTGCTGCTTTTGCTTGTCGCTTTTC
26	*Glycine max*	XP_006594434	96–391	AACGAGGAATTCATGAACATTGAGCTGCCGGC
				CAAGTTTTCGGTACCCTGTTTAGCCTGTGAATAGAGGAAG
27	*Glycine max*	XP_003519000	87–395	AACGAGGAATTCATGTCAGAGGTTGAGGGTGCAAGC
				CAAGTTTTCGGTACCTAGTTTCGCCTGTGAATACAAGAAG
28	*Glycine max*	XP_006600314	83–392	AACGAGGAATTCATGAACATTGAGCTCCCTGATGAAGAAG
				CAAGTTTTCGGTACCTTGTTTCGCCTGTGAATACAAGAAG
29	*Glycine max*	XP_003545984	95–418	AACGAGGAATTCATGAATCCTGAAGGAGAAAATGTAGCCC
				CAAGTTTTCGGTACCACTACTCTTTTCACCTTCAATTTTCATT
30	*Glycine max*	XP_003537155	96–419	AACGAGGAATTCATGAATCCTGAAGGAGAAAATGTAACCC
				CAAGTTTTCGGTACCACTAGTCTTTTCATCTTCAATTTTCTGTG
31	*Glycine max*	XP_003545023	88–395	AACGAGGAATTCATGTCAAAGGTTGGAGGGGCTG
				CAAGTTTTCGGTACCCTGATTTGCCTGTGAATACAAGAAG
32	*Glycine max*	XP_003519563	90–395	AACGAGGAATTCATGTCAGAGGTTGGAGGGGCTCC
				CAAGTTTTCGGTACCCTGATTTGCCTGTGAATACAAGAAG
33	*Glycine max*	XP_003527013	77–406	AACGAGGAATTCATGCCTCCCCGCGCCG
				CAAGTTTTCGGTACCAGCTGTTTTTGGCTTTGCCTTAATA
34	*Glycine max*	NP_001242209	83–408	AACGAGGAATTCATGCCTTCTTCACCTCCTCGCGC
				CAAGTTTTCGGTACCAGCTGTTTTTGGCTTTGCCTTAATA
35	*Glycine max*	XP_003550731	67–382	AACGAGGAATTCATGTCCTCTATACCTGATGCTAGAAGTGAT
				CAAGTTTTCGGTACCATTTGTCTTTTGGACTGACTTGATC
36	*Glycine max*	XP_003538441	124–447	AACGAGGAATTCATGACCTCCGTGCCGGAGAGC
				CAAGTTTTCGGTACCAGTTGTCTTTGGCTTAAGTCGCTTC
37	*Glycine max*	XP_006591369	85–408	AACGAGGAATTCATGTCAGCACAAAAGGAAGAAGAAGAAG
				CAAGTTTTCGGTACCCTTGCTTCTCTTTTTGGCTTCACTG

Table 3 | Continued

No.	Species	Protein	Region	Primer sequences (5'–3')
38	*Cyanidioschyzon merolae*	CYME_CMK114C	101–422	AACGAGGAATTCATGTCAGGTGGTGTGCCAAAACC
				CAAGTTTTCGGTACCCCTCTGCTGTTTTGAACCCCG
39	*Cyanidioschyzon merolae*	CYME_CMO269C	90–409	AACGAGGAATTCATGAGCGCGGGTGACGCG
				CAAGTTTTCGGTACCGGCGGCCTTTTTCTTGGCT
40	*Cyanidioschyzon merolae*	CYME_CMN328C	75–394	AACGAGGAATTCATGAGCGGCACGAGTGCGC
				CAAGTTTTCGGTACCGGCTTTTTTCACCTTACCGCCTC
41	*Galdieria sulfuraria*	Gs21660 (GsTPT)	82–406	AACGAGGAATTCATGTCCATCAAAGTGACGGAAGCT
				CAAGTTTTCGGTACCCTTTATTTTTTGACTGTAATAATATTTGGTTAA
42	*Galdieria sulfuraria*	Gs53050 (GsPPT)	79–407	AACGAGGAATTCATGTCTTCTACTGGGACCTCATCTTCG
				CAAGTTTTCGGTACCTTCTATTTTCTCCTTCTTCTTGGTGGATATAT
43	*Galdieria sulfuraria*	Gs48050 (GsGPT)	91–410	AACGAGGAATTCATGTCTCCTCAAAAGTCCAGTGTAGGAGTG
				CAAGTTTTCGGTACCCTGCTTTTCTCGCTTCGATGGTAG

3.2.2 Fluorescence-detection size-exclusion chromatography

The GFP-based screening and optimization were performed based on published protocols (*95*, *96*, *118*). The plasmids outlined in Section 3.2.1 were transformed into *Saccharomyces cerevisiae* cells (strain BY4742) using Frozen-EZ Yeast Transformation II Kit (ZYAGEN) following manufacturer's protocol. Transformed cells were grown in a CSM −URA medium containing 2% raffinose, and protein expression was induced by adding 2% galactose when the culture reached an A600 = 0.6. After allowing protein expression for 22 h at 30ºC, cells were harvested by centrifugation at 8,000 g for 1 min and disrupted in a buffer containing 50 mM Tris-HCl, pH 8.0, 150 mM NaCl and protease inhibitors, using acid-washed glass beads (200–400 μm; Sigma). Cell debris were removed by centrifugation at 1,000 g for 1 min. The supernatants were ultracentrifuged at 138,000 g for 60 min with a micro-ultracentrifuge CS100FNX (Hitachi Kohki) to prepare the membrane fraction.

Expression levels and solution behaviors of candidate proteins were evaluated by the following procedure. Membrane fractions from 2 ml-culture were solubilized in a 300 μl buffer containing 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5% glycerol, 1 mM β-mercaptoethanol (β-ME), protease inhibitors (phenylmethylsulfonyl fluoride, aprotinin, leupeptin and pepstatin) and 1–2% of specified detergents. After solubilization for 90 min at 4ºC, insoluble materials were removed by ultracentrifugation. 50 μl of supernatants were subjected to a fluorescence-detection size-exclusion chromatography on a Superdex 200 10/300 Increase column (GE Healthcare), equilibrated with a buffer containing 10 mM Tris-HCl, pH 8.0, 150 mM NaCl, and 0.03% DDM. The elution profile was monitored by a fluorescence detector RF-20Axs (Shimadzu) with the excitation and emission wavelengths at 480 nm and 512 nm, respectively. To evaluate thermal stability, protein samples were heated at various temperature for 10 min. After removing protein aggregates by ultracentrifugation, 50 μl supernatants were subjected to FSEC analysis as described above.

3.2.3 Protein purification

For large-scale purification, proteins were overexpressed in insect-baculovirus system. The gene regions encoding selected pPT proteins were transferred to a modified pFastbac vector, with a C-terminal TEV cleavage site, EGFP and a His10-tag. Suspension cultures of *Spodoptera frugiperda* Sf9 cells were maintained in SF900 II SFM medium (Thermo Fisher Scientific). Recombinant baculoviruses were produced with the Bac-to-Bac system (Invitrogen), and were used to infect sf9 cells at a density of 2–3×10^6 cells ml^{-1}. After growth for 48 h at 27ºC, the cells were harvested and sonicated in a lysis buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl and protease inhibitors). The cell debris was removed by low-speed centrifugation (10,000 g, 10 min), and the membrane

fraction was collected by ultracentrifugation (138,000 g, 1 h). The membrane fraction was solubilized in a solubilization buffer (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 1% (w/v) lauryl maltoside neopentyl glycol (LMNG) (*119*) and 1 mM β-mercaptoethanol (β-ME)) for 3 h at 4ºC. The supernatant was isolated by ultracentrifugation (138,000 g, 30 min) and subjected to immobilized metal ion affinity chromatography (IMAC) with Ni-NTA resin (Qiagen). The resin was washed with an IMAC buffer (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 0.05% LMNG, 1 mM β-ME and 30 mM imidazole), and the protein was eluted with an IMAC buffer supplemented with 300 mM imidazole. The eluate was treated with TEV protease and dialyzed overnight against a dialysis buffer (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 0.01% LMNG and 1 mM β-ME). The cleaved EGFP-His10 and TEV protease were removed by reverse IMAC with Ni-NTA. The protein was concentrated to 2–3 mg ml^{-1} using a 50 kDa MWCO concentrator (Millipore), and further purified by size-exclusion chromatography (SEC) on a Superdex 200 Increase 10/300 column (GE Healthcare) in a SEC buffer (10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.01% LMNG and 1 mM β-ME). The peak fractions were collected, concentrated to 10–20 mg ml^{-1}, flash-frozen in liquid nitrogen and stored at −80ºC until crystallization.

3.2.4 Size exclusion chromatography coupled to multi-angle laser light scattering (SEC-MALLS)

The instrument setup for the SEC-MALLS experiment consisted of an Agilent 1100 Series HPLC system connected in series with a Shimadzu SPD-10Avp UV absorbance detector, a Wyatt DAWN HELEOS 8+ light scattering detector and a Shodex RI 101 refractive index detector. Analytical size-exclusion chromatography was performed at 25ºC on a Superdex 200 10/300 column equilibrated with a buffer containing 10 mM Tris-HCl, pH 8.0, 150 mM NaCl and 0.01% LMNG. A 90 µl portion of the purified GsGPT sample (1.5 mg ml^{-1}) was injected into the column and eluted at 0.5 ml min-1. Elution was monitored in line with the three detectors, which simultaneously measured UV absorption, light scattering and refractive index. A 658 nm laser was used in the light scattering measurement. Molecular masses were calculated using the three-detector method (*120*), as implemented in the ASTRA software package (Wyatt Technology).

3.2.5 Crystallization

Purified samples were thawed and mixed with 1-oleoyl-R-glycerol (monoolein), at a protein to lipid ratio of 2:3 (w/w) to prepare the lipidic cubic phase (LCP) as previously described (*98*). Crystallization experiments were performed with 96-well glass sandwich plates (Molecular Dimensions), using a Gryphon LCP robot (Art Robbins Instruments). Typically, 50 nl of protein-laden LCP drops were overlaid with 800 nl of precipitant solution. After extensive co-crystallization screening, needle-shaped crystals appeared under conditions containing high

concentrations of 3-PGA or Pi. Optimized crystals of the 3-PGA-bound state were obtained in 35–40% PEG200, 100 mM Na-citrate, pH 6.0, 50–100 mM citrate 3K and 50–100 mM 3-PGA·2Na. Optimized crystals of the Pi-bound state were obtained in 43–48% PEG200, 50–100 mM MES-NaOH, pH 6.0 and 200–250 mM $(NH_4)_2HPO_4$. Crystals were harvested and flash-cooled in liquid nitrogen for data collection.

3.2.6 X-ray diffraction data collection

X-ray diffraction experiments were performed at the micro-focus beamline BL32XU at SPring-8. The locations of well-diffracting crystals were identified by raster scanning, and data were collected for a 5–30º wedge from each crystal. For the Pi-bound datasets, the detectors used were MX225HS CCD detector (Rayonix) and EIGER X 9M (Dectris), with the camera distances of 220 mm and 180 mm, respectively. For the 3-PGA-bound crystals, the detector used was EIGER X 9M (Dectris) with the camera distance of 250 mm. All diffraction data were processed with XDS (*99*), and merged with XSCALE based on the hierarchical clustering analysis with BLEND (*121*) or with the cross-correlation method as implemented in the KAMO software (https://github.com/keitaroyam/yamtbx).

3.2.7 Data processing and structure determination

For the determination of the Pi-bound structure, molecular replacement trials were performed on Phaser (*105*) using various truncated structures of DMT proteins as search models. Initial solutions were obtained with a full-length polyalanine model of the YddG monomer (PDB 5I20). After initial refinement in PHENIX (*96*), the resulting map (R_{free} value 53%) showed poor or no electron density for substantial portions of the structure, particularly for TM5, TM10 and all loop regions. These invisible segments were deleted from the model and rebuilt by multiple trials of manual modelling of new polyalanine helices using COOT (*104*) and refinement with phenix.refine, to find the correct helix assignment. After subsequent rounds of model building and refinement, we could build ten helix backbones and several sidechains into the visible electron density. However, at this point, further model building did not improve the R_{free} value or the quality of the electron density. We then noticed that the model was of a 'swapped' form of the protein, where the N-terminal repeat (TM1–5) and the C- terminal repeat (TM6–10) were inversely assigned to each other. We corrected this swapping by renumbering the residues in COOT and proceeded with further model building. After building the protein regions (residues 100–404), strong electron densities were observed within the central cavities of the two monomers, which were unambiguously assigned as bound Pi molecules. During the later stages of refinement, electron densities for water and lipid molecules were also identified. The structure was iteratively rebuilt and refined with COOT and PHENIX to achieve good stereochemistry and R_{free} values.

The 3-PGA-bound structure was determined using the Pi-bound GsGPT dimer as the starting model, and iteratively rebuilt and refined with COOT and PHENIX. All molecular graphics were illustrated using CueMol (http://www.cuemol.org).

3.2.8 Transport assays

Yeast membranes expressing recombinant proteins were prepared as previously described (*122*), with slight modifications. The region of GsGPT encoding residues 91–410 was cloned into a modified pYES2 vector, with a C-terminal His_6 tag. For mutant assays, mutations were introduced by a PCR-based method. The plasmids were transformed into *Saccharomyces cerevisiae* cells (strain BY4742). Transformed cells were grown in CSM−URA medium containing 2% raffinose, and protein expression was induced with 2% galactose when the culture reached an A600 = 0.6. After growth for 22 h at 30ºC, the cells were harvested and disrupted in a lysis buffer (50 mM Tricine-KOH, pH 7.5), 0.1 mM phenylmethylsulfonyl fluoride and 5% glycerol), using acid-washed glass beads (200–400 μm; Sigma). Glass beads and cell debris were removed by low-speed centrifugation (4,000 g, 2 min), and the membrane fraction was collected by ultracentrifugation (138,000 g, 1 h). The membrane pellet was resuspended in 50 mM Tricine-KOH, pH 7.5, flash-frozen in liquid nitrogen and stored at −80ºC until use. Aliquots of the resuspended membranes were subjected to SDS-PAGE, and the His_6-tagged recombinant proteins were detected by a western-blot analysis, using an anti-His-tag polyclonal antibody (code PM032; MBL).

Soybean L-α-phosphatidylcholine (Avanti) in chloroform was dried into a thin film under a stream of nitrogen gas, and further dried under vacuum. Dried lipids were resuspended at 20 mg ml^{-1} in intra-liposomal solution (120 mM Tricine-KOH, pH 7.5, and 30 mM NaH_2PO_4) or Pi-free intra-liposomal solution (150 mM Tricine-KOH, pH 7.5), and sonicated for 5 min at 4ºC to form unilamellar vesicles. This unilamellar vesicle solution was reconstituted with the yeast membranes at 19:1 (v/v), by the freeze-thaw procedure. The reconstituted liposomes were sonicated again for 5 min at 4ºC, to form unilamellar vesicles. The extra-liposomal solution was exchanged by gel-filtration on Sephadex G-50 (GE Healthcare) pre-equilibrated with 150 mM Tricine-KOH, pH 7.5.

For assays in the purified system, the unilamellar liposome solution was reconstituted with purified GsGPT at a lipid-to-protein ratio of 100:1 (w/w), by the freeze-thaw procedure. The resulting proteoliposomes were sonicated and buffer-exchanged, as described above. Protein-free liposomes were prepared by the same procedure, except that the protein solution was replaced with the SEC buffer used in the purification.

The liposome assays were performed as previously described (*122*), with slight modifications. For the time-dependent uptake assay, the reaction was started by mixing the reconstituted liposome solution with an equal volume of extra-liposomal solution (150 mM Tricine-KOH, pH 7.5, and 1 mM [^{32}P]-NaH$_2$PO$_4$ (0.1 mCi ml^{-1})). At different time points, liposomes were isolated by anion exchange on AG-1 X8 resin (acetate form, 200–400 dry mesh size; Bio-Rad), pre-equilibrated with 150 mM sodium acetate. The radioactivity of the incorporated [^{32}P]-Pi was quantified by liquid scintillation counting. Mutant assays were performed with a similar procedure, and the total amounts of incorporated [^{32}P]-Pi were compared at 30 min. For the counter-flow assay, the liposomes containing 30 mM Pi, 30 mM 3-PGA or no substrate were mixed with extra-liposomal solution containing 0.25 mM [^{32}P]-Pi. The total amounts of incorporated [^{32}P]-Pi were compared at 3 min. For the competitive inhibition assay, the liposomes containing 30 mM Pi were mixed with extra-liposomal solution containing 0.25 mM [^{32}P]-Pi and 40 mM of the indicated competitive inhibitor. The total amounts of incorporated [^{32}P]-Pi were compared at 3 min. For the determination of kinetic constants, the Michaelis constant (K_m) for Pi was analyzed using various external concentrations of [^{32}P]-Pi (0.3125–10 mM) and a fixed internal concentration of Pi (30 mM). Inhibitor constants (K_i) were assessed with two different external concentrations of [^{32}P]-Pi (0.5–2.5 mM) and four different concentrations of the indicated inhibitors (0–10 mM). To assess the background uptake, control experiments were performed with membranes from yeast cells transformed with empty vector. Enzyme kinetic data were analyzed by non-linear regression fitting, as implemented in the GraphPad Prism 7 software.

3.2.9 Molecular dynamics simulation

All molecular dynamics simulations were performed by Mr. Mizuki Takemoto at the University of Tokyo. The simulation system included the GsGPT dimer, 1-phosphoryl-2-oleoylphosphatidylcholine (POPC), TIP3P water and 150 mM NaCl. The disordered sidechains in the GsGPT crystal structure were modelled by COOT. To embed the protein within the POPC bilayer, we used the protocol described by Javanainen (*123*). One POPC molecule was placed in the GsGPT dimerization interface, corresponding to the two monoolein molecules in the crystal structure. Finally, the periodic boundary systems, including 136,668 (with Pi) and 136,652 (without Pi) atoms, with the size of 90.7×147.9×100.0 Å, were prepared. The net charge of the solute was neutralized with sodium and chloride ions. The molecular topologies and force field parameters from CHARMM36 (*124*) were used. Molecular dynamics simulations were performed by the program Gromacs, version 5.0.5 (*125*). First, energy minimization was performed using the steepest descent, with a cut-off of 1,000.0 kJ mol^{-1} nm^{-1}. Next, random velocities were assigned according to a Maxwell distribution, at a temperature of 310 K for each atom, and an

equilibration run (eq1) was performed for 100 ps in the canonical (NVT) ensemble (310 K, $90.7 \times 147.9 \times 100.0$ Å volume). Finally, an equilibration run (eq2) was performed for 1,000 ps in the isothermal-isobaric (NPT) ensemble (310 K, 1 bar). The positions of non-hydrogen atoms in the protein and phosphates were restrained with a force constant of 1,000 kJ mol^{-1} nm^{-2}, in the minimization and equilibration runs. Production runs were performed for 100 ns in the NPT ensemble (310 K, 1 bar). The same simulation was performed twice with different initial velocities, and similar results were obtained. Constant temperature was maintained by using V-rescaling ([126]) with a time constant of 0.1 ps in eq1, and a Nosé-Hoover thermostat ([127], [128]) with a time constant of 0.5 ps in eq2 and the production runs. Pressure was controlled with semiisotropic coupling to a Parrinello-Rahman barostat ([129]), with a time constant of 5.0 ps and a compressibility of $4.5 \times 10\text{-}5$ bar^{-1}. The LINCS algorithm ([130]) was used for bond constraints. Long range electrostatic interactions were calculated with the particle mesh Ewald method ([131]).

3.3 Results

3.3.1 Expression screening of TPT homologues

To elucidate the structure and the molecular mechanism of TPT, we first performed systematic screening of dozens of plant pPTs for identifying candidates with good expression and stability. Since previous studies have used yeast cells for heterologous expression of the pPTs (*122*, *132–135*), we selected yeasts as the primary expression host in initial trials. Our expression construct includes a C-terminal GFP fusion and a His_8 tag for purification, and an N-terminal truncation that essentially omits the chloroplast transit peptide (~90 residues), giving rise to a mature translocator protein (Figure 23). Expression trials in the fluorescent-detection size-exclusion chromatography showed that most plant pPTs could be readily overexpressed and solubilized (Figure 24). Among those candidates, two GPT orthologues from *Oryza sativa* (rice) and *Glycine max* (soybean), designated as OsGPT1 and GmGPT1, exhibited good solution behavior, with apparent melting temperatures (T_m) of >70ºC (Figure 25). We purified these proteins from insect-baculovirus expression system and proceeded with crystallization trials. However, even after extensive screening, both of them yielded no crystals.

To identify more suitable candidates for crystallization, we extended our expression trials to the pPTs from other species. We reasoned that the organisms that inhabit high-temperature environments could possess more stable proteins. We thus focused on the two unicellular thermophilic red algae, *Cyanidioschyzon merolae* and *Galdieria sulphuraria*, which naturally inhabit hot springs and can grow at temperatures up to 45ºC and 56ºC (*136*, *137*). We found that, among the six pPTs coded within these red algal genomes, GsGPT (*122*) exhibits the best solution behavior, with an apparent T_m of ~60ºC when solubilized in lauryl maltose neopentyl glycol (LMNG), a newly-developed detergent (*119*) (Figure 25). After optimization, we were able to purify GsGPT to high purity and homogeneity (Figure 26), which were then subjected to further functional and structural studies.
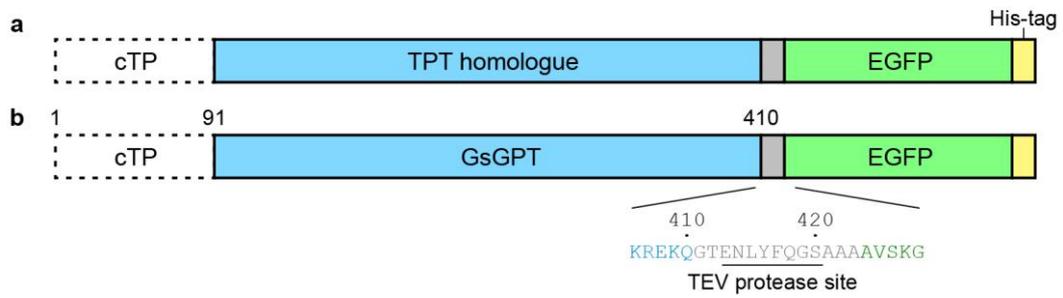
Figure 23 | Expression construct for pPTs

(a) Construct used for expression screening. The chloroplast transit peptide (cTP) was truncated.

(a) GsGPT construct used for purification. The TEV protease cleavage site, EGFP and His$_8$-tag were fused to the C-terminus.
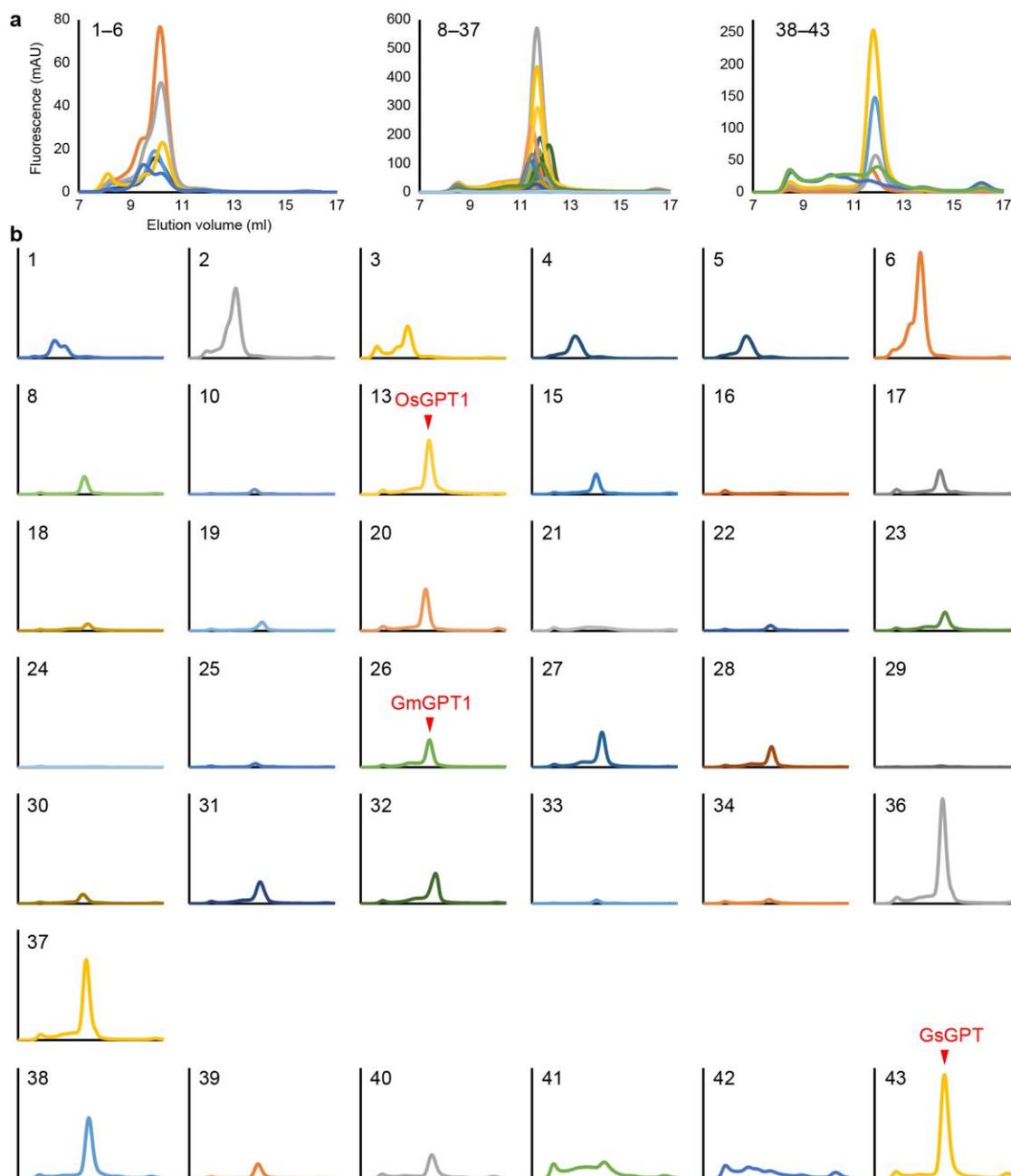
Figure 24 | FSEC profiles of pPTs

(a) Superposition of fluorescence traces of different pPTs in size exclusion chromatography. The numbers correspond to the proteins summarized in Table 3. Note, experiments were performed in three separate batches (1–6, 8–37 and 38–43), and thus those chromatograms are superposed separately.

(b) Individual chromatograms of (a). Candidates that exhibit good solution behavior are highlighted by red arrows. Note, some proteins (No. 7, 9, 11, 12, 14 and 35) showed no expression, and thus they are not displayed here.
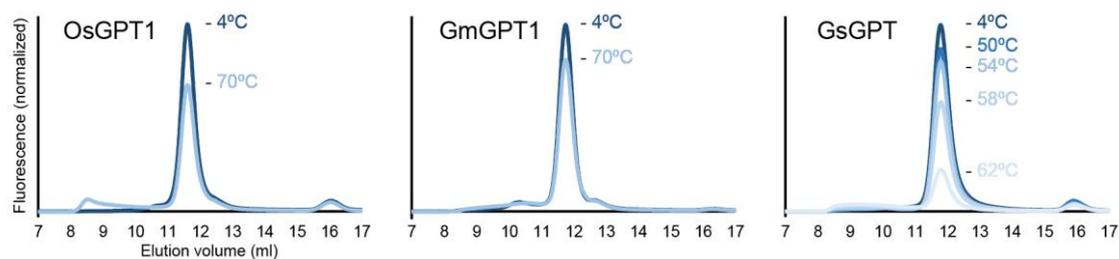
Figure 25 | FSEC-based thermal stability assay on pPTs

FSEC traces of OsGPT1, GmGPT1 and GsGPT after being kept on ice (4ºC) or heated at indicated

temperatures (50–70ºC, 10 min). OsGPT1 and GmGPT1 were solubilized in 2% DDM + 0.4% CHS, and

GsGPT was in 1% LMNG.
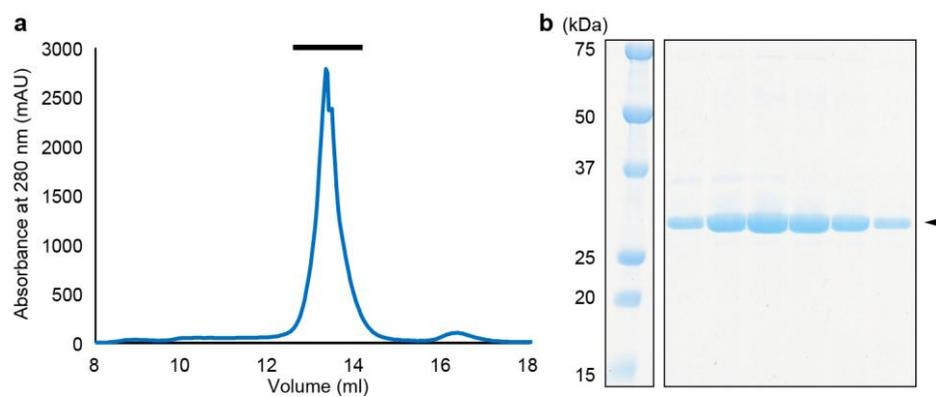
Figure 26 | Purified GsGPT

(a) Final gel filtration profile of the crystallized sample.

(b) SDS-PAGE gel of (a).

3.3.2 Functional characterization of TPT from red alga

The function of GsGPT has not been fully characterized (*122*). To characterize the transport function of GsGPT, we performed liposome-based assays on its crystallization construct (residues 91–410). Counter-flow assays with radiolabeled substrates confirmed the 'signature' Pi/Pi homo-exchange activity of GsGPT (Figure 27a), as well as the 3-PGA/Pi hetero-exchange activity (Figure 27b,c). Competitive inhibition assays by various phosphorylated metabolites suggested that GsGPT transports phosphorylated C3, C4 and C5 compounds with linear sugar moieties, but not phosphorylated C5 and C6 compounds with ring sugar moieties (Figure 27d), reminiscent of the function of TPT (*70*, *72*). Further measurement of the kinetic constants confirmed that the substrate specificity of GsGPT is comparable to that of higher plant TPTs: the Michaelis constant (K_m) for Pi is about 1.3 mM, and the inhibition constants (Ki) for triose-P, 3-PGA and phosphoenolpyruvate (PEP) are about 1.0, 0.6 and 8.3 mM, respectively (Figure 27e,f). The affinities for phosphorylated C4 and C5 compounds were slightly lower than those for triose-P and 3-PGA. These biochemical data show that GsGPT is functionally similar to TPT, although it has been named GPT based on the sequence similarity (*122*).
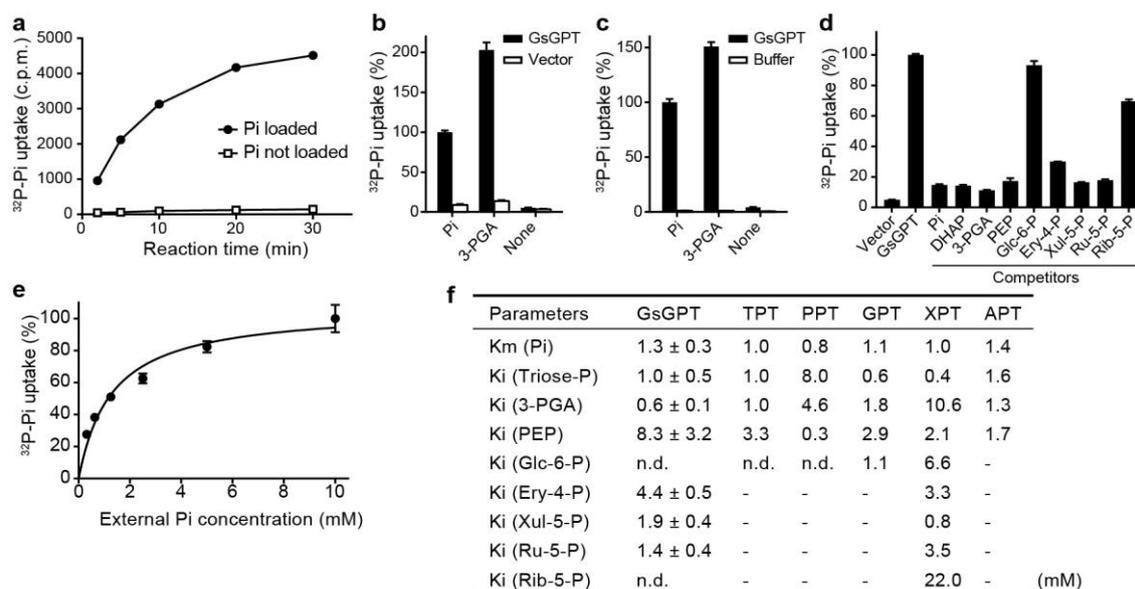
Figure 27 | Functional characterization of GsGPT

(a) Pi/Pi homo-exchange activity of GsGPT. Liposomes were reconstituted with yeast membranes expressing GsGPT, and the time-dependent uptake of [^{32}P]-Pi was measured in the presence (filled circles) or absence (open squares) of internal Pi (30 mM). Error bars are s.e.m. (n=3). All transport experiments were performed using similar yeast-based assays, unless otherwise stated.

(b) Counter-flow assay. The uptake of [^{32}P]-Pi was measured in the presence of the indicated internal substrates (30 mM). Control experiments were performed with membranes from yeast cells harboring empty vector. Error bars are s.e.m. (n=3).

(c) Transport activity of the purified protein. Liposomes were reconstituted with purified GsGPT, and counter-flow assays were performed as in (b). Control experiments were performed with protein-free liposomes. Error bars are s.e.m. (n=3).

(d) Competitive inhibition assay. The uptake of [^{32}P]-Pi was assayed in the presence of the indicated competitive inhibitors (40 mM) in the external solution. Error bars are s.e.m. (n=3).

(e) Concentration-dependent uptake of [^{32}P]-Pi. Error bars are s.e.m. (n=3).

(f) Kinetic constants of GsGPT and other pPTs. The Michaelis constant (K_m) of GsGPT was calculated from the experiment shown in Figure 1e. Inhibitor constants (K_i) of GsGPT were evaluated at two different Pi concentrations with increasing inhibitor concentrations. Data are mean ± s.e.m. (n=3); n.d., not detectable. Values for the plant and apicomplexan pPTs were adopted from refs. (*79*, *81*, *84*, *132*).

3.3.3 Crystallization and structure determination

Toward high-resolution structure determination of GsGPT, we performed crystallization trials by using the lipidic cubic phase method (*138*). After extensive screening, co-crystallization with high concentrations (50–250 mM) of 3-PGA or Pi yielded diffraction-quality crystals belonging to the $P2_12_12$ space group (Figure 28). We collected diffraction data from several hundred crystals using the microfocus X-ray beam (Figure 29), and then merged those good data based on the hierarchical clustering method (*121*) to obtain the complete, high-quality datasets (Table 4).

A previous bioinformatics analysis suggested the classification of TPT into the drug/metabolite transporter (DMT) superfamily (*139*), which involves a diverse range of membrane proteins possessing 4, 5, 9 or 10 transmembrane (TM) helices. Therefore, we tried calculating the initial phases by molecular replacement by using available structures of DMT proteins. After extensive molecular replacement trials with full-length or truncated search models, an initial solution was obtained from a poly-alanine model of the 10-TM DMT transporter SnYddG (*140*) (Figure 30). The final structures were determined at 2.2 and 2.1 Å resolutions for the 3-PGA- and Pi-bound states, respectively (Figure 31 and Figure 32). Except for the bound ligands, the two co-crystal structures are almost identical, with an r.m.s.d. value of 0.18 Å over 608 Cα atoms.
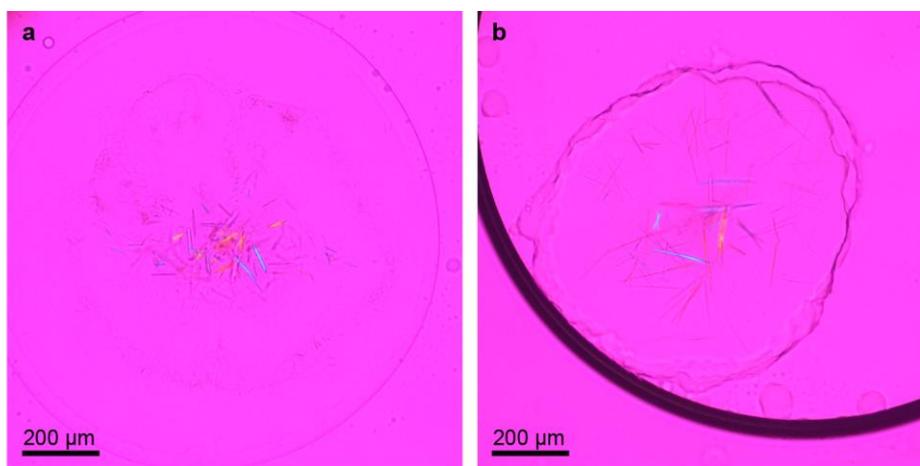
Figure 28 | Crystals of GsGPT

(a) 3-PGA co-crystals. (b) Pi co-crystals.
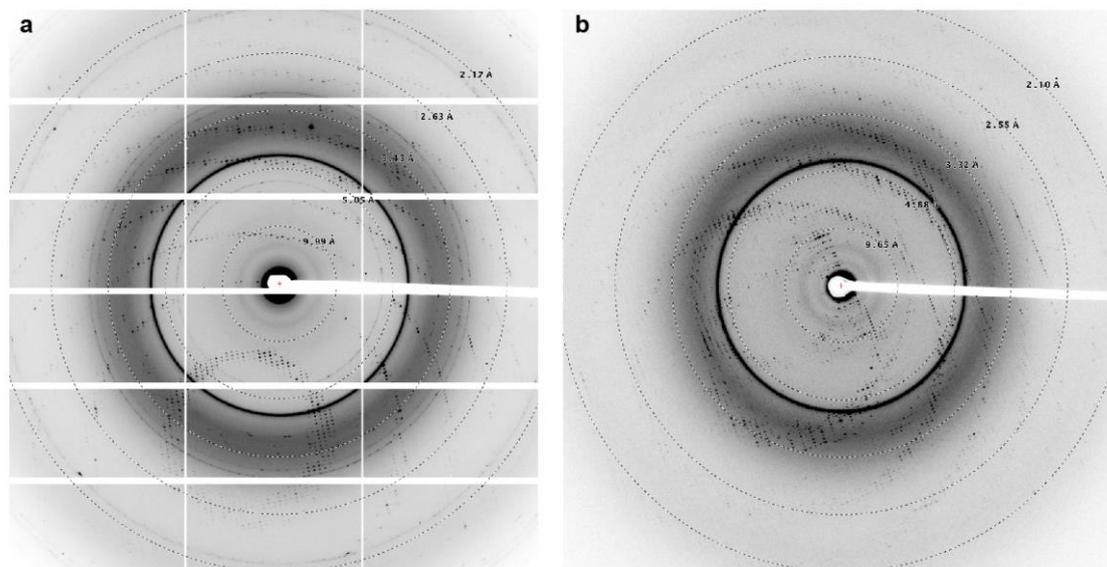
Figure 29 | X-ray diffraction images of GsGPT

(a) Diffraction image of the 3-PGA-bound crystal. The outermost ring indicates 2.17 Å resolution.

(b) Diffraction image of the Pi-bound crystal. The outermost ring indicates 2.10 Å resolution.
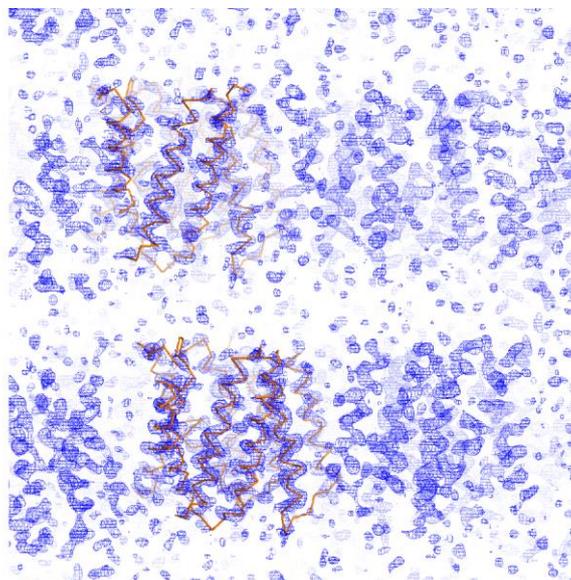
Figure 30 | Structure determination of GsGPT by molecular replacement

The 2Fo-Fc electron density map after molecular replacement, contoured at 1.8σ. The SnYddG structure used as the search model is shown as Cα traces. We identified two molecules in the asymmetric unit.
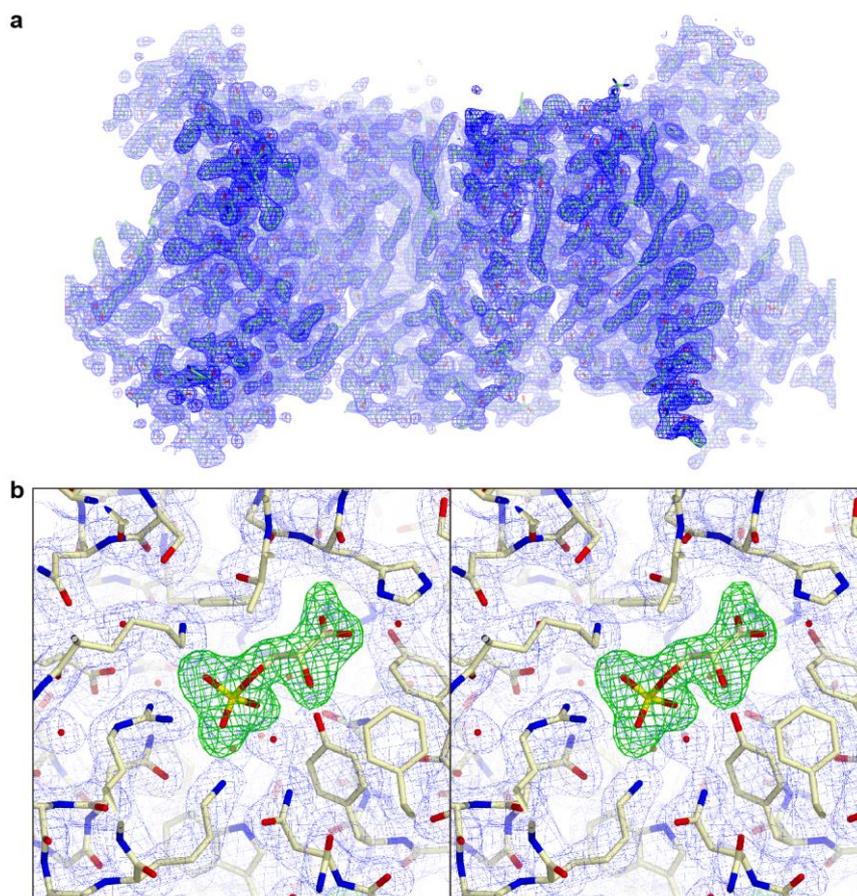
Figure 31 | Electron density maps of the 3-PGA-bound GsGPT

(a) The 2Fo-Fc electron density map after refinement, contoured at 1.0σ.

(b) Stereo view of the omit electron density map, calculated without 3-PGA. The Fo-Fc map is shown in green, contoured at 3.0σ, and the 2Fo-Fc map is in blue, contoured at 1.0σ.
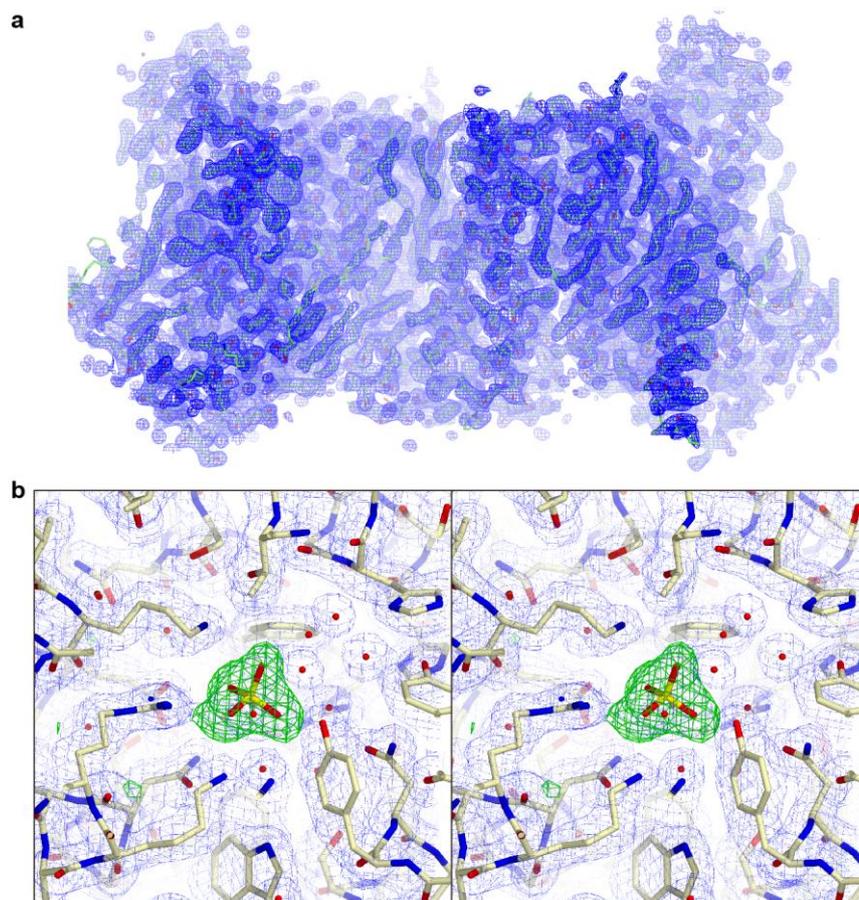
Figure 32 | Electron density maps of the Pi-bound GsGPT

(a) The 2Fo-Fc electron density map after refinement, contoured at 1.0σ.

(b) Stereo view of the omit electron density map, calculated without Pi. The Fo-Fc map is shown in green, contoured at 3.0σ, and the 2Fo-Fc map is in blue, contoured at 1.0σ.

Table 4 | Data collection and refinement statistics of GsGPT

	3-PGA-bound[a]	Pi-bound[b]
Data collection		
Space group	$P2_12_12$	$P2_12_12$
Cell dimensions		
a, b, c (Å)	107.05, 165.33, 41.14	106.55, 164,97, 41.42
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	50.0–2.20 (2.28–2.20)[c]	50.0–2.10 (2.18–2.10)
R_{pim}	0.087 (1.035)	0.138 (4.357)
$CC_{1/2}$	0.994 (0.698)	0.995 (0.653)
$I / \sigma I$	13.3 (1.3)	11.3 (1.1)
Completeness (%)	100.0 (99.9)	99.8 (100.0)
Redundancy	39.2 (27.0)	90.3 (87.0)
Refinement		
Resolution (Å)	50.0–2.20	50.0–2.10
No. reflections	38,123	43,639
R_{work} / R_{free} (%)	19.0 / 22.9	19.5 / 22.6
No. atoms		
Protein	4,713	4,721
Substrate	22	10
Lipid	426	558
Other	235	283
B-factors		
Protein	28.70	24.82
Substrate	23.45	20.09
Lipid	41.58	43.90
Other	36.85	35.94
Ramachandran plot		
Favored (%)	99.3	99.5
Allowed (%)	0.7	0.5
Outliers (%)	0	0

[a]Diffraction data were collected from 199 crystals.

[b]Diffraction data were collected from 319 crystals.

[c]Values in parentheses are for highest-resolution shell.

3.3.4 Overall structure

The overall structure of GsGPT reveals a 10-transmembrane (TM) helix topology with both the N- and C-termini on the stromal side (inside), rather than the previously predicted 6–9 TM helix topologies (*75*) (Figure 33). GsGPT contains two 'inverted' structural repeats, comprising the N- and C-halves. Viewed from the intermembrane space side (outside), the five helices within the N- and C- halves (i.e., TM1–5 and TM6–10) are arranged in counter-clockwise and clockwise manners, respectively (Figure 34). This fold is essentially similar to that of the bacterial DMT superfamily transporter SnYddG (*140*), despite the low sequence identity (13.9%), suggesting that this '10-TM DMT fold' could be conserved across all putative 10-TM members of the DMT superfamily (*141*). In contrast to the 'outward-open' conformation of SnYddG, the current structure of GsGPT shows that its substrate-binding site is occluded from both sides of the membrane, revealing the 'occluded' conformation of a DMT protein for the first time.

Although the purified GsGPT protein is monomeric in solution, GsGPT forms a dimer in the crystallographic asymmetric unit (Figure 35). The inter-protomer interaction involves polar interactions at TM5, TM10 and a short β-strand connecting TM4 and TM5, and hydrophobic contacts through the lipid molecules bound at the interface. The same topological orientation of the monomers within the membrane suggests that this dimeric assembly could be physiologically relevant (*114*), although we cannot exclude the possibility that this is a crystallization artifact.

3.3.5 3-PGA and Pi recognition

The electron density maps clearly showed that 3-PGA and Pi are bound to the same site located halfway across the membrane, as if trapped in a central 'cage' formed by TM1–4 and TM6–9 (Figure 31 and Figure 32). The phosphate moiety of both ligands is identically recognized by ionic bonds with Lys204, Lys362 and Arg363 and a hydrogen bond with Tyr339 (Figure 33c,d). While the three oxygen atoms of the phosphate (P-O2, O3 and O4) are directly recognized by these sidechains, the remaining oxygen atom (P-O1) is not directly recognized. In the 3-PGA-bound structure, this P-O1 is attached to the glycerate group, which extends into the space on the opposite side of the phosphate moiety and forms specific interactions with protein sidechains (Figure 33c). The carboxyl group on the C1 atom forms an ionic bond with the sidechain of His185, and the hydroxyl group on the C2 atom hydrogen bonds with the sidechain of Tyr339. In addition, the C2 and C3 atoms form hydrophobic contacts with the sidechains of Thr188, Phe192 and Phe263 (Figure 36).

In the Pi-bound structure, the corresponding space near P-O1 is occupied by three water molecules (Figure 33d). These water molecules form polar interactions with His185 and Tyr339,

contributing to the indirect recognition of P-O1. Notably, the positions of these water molecules roughly correspond to those of the three oxygen atoms of the glycerate moiety of 3-PGA, mimicking the organic carbon structure. This water-mediated hydrogen-bonding network is likely to lower the energy of the Pi-bound state and could explain why Pi, which lacks a sugar moiety, is transported with a similar affinity to those of other sugar phosphates (*76*). To examine the functional importance of the observed interactions, we performed mutational assays of the five residues involved in the phosphate recognition (Figure 33f). All of the tested mutations exhibited greatly reduced Pi/Pi homo-exchange activity, confirming their essential roles in transport.

The observed binding mode of 3-PGA suggests that triose-P can be recognized in a similar manner. Indeed, modelling of triose-P into the crystal structure indicates a good fit, with the oxygen atoms at the C1 and C2 positions forming similar polar interactions with His185 and Tyr339 (Figure 33e). Therefore, the structures suggest that triose-P, 3-PGA and Pi, the three major counter-substrates of TPT, are similarly recognized in a single pocket. The structures also reveal that this substrate-binding pocket could not accommodate two or more phosphate moieties at a time, explaining why pyrophosphate or bisphosphate compounds are not readily transported across the chloroplast envelope membrane (*69*, *70*).
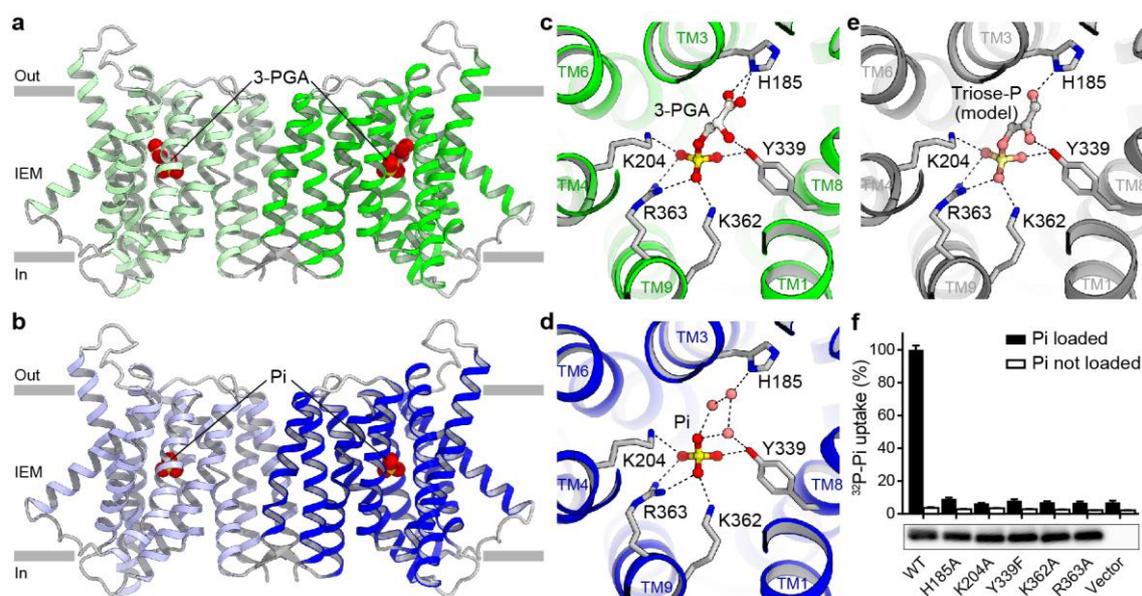
Figure 33 | Overall structure of GsGPT

(a, b) Ribbon representations of the 3-PGA- (a) and Pi-bound (b) structures. IEM denotes the chloroplast inner envelope membrane.

(c, d) Close-up views of the 3-PGA- (c) and Pi-binding site (d). Dotted lines indicate polar interactions.

(e) Model of triose-P (dihydroxyacetone phosphate) binding.

(f) Liposome-based mutational analysis. The levels of [^{32}P]-Pi uptake by GsGPT mutants were compared to that of the wild-type. Error bars are s.e.m. (n=3). Western blotting confirmed the comparable expression levels of the wild type and mutant proteins (small inset).
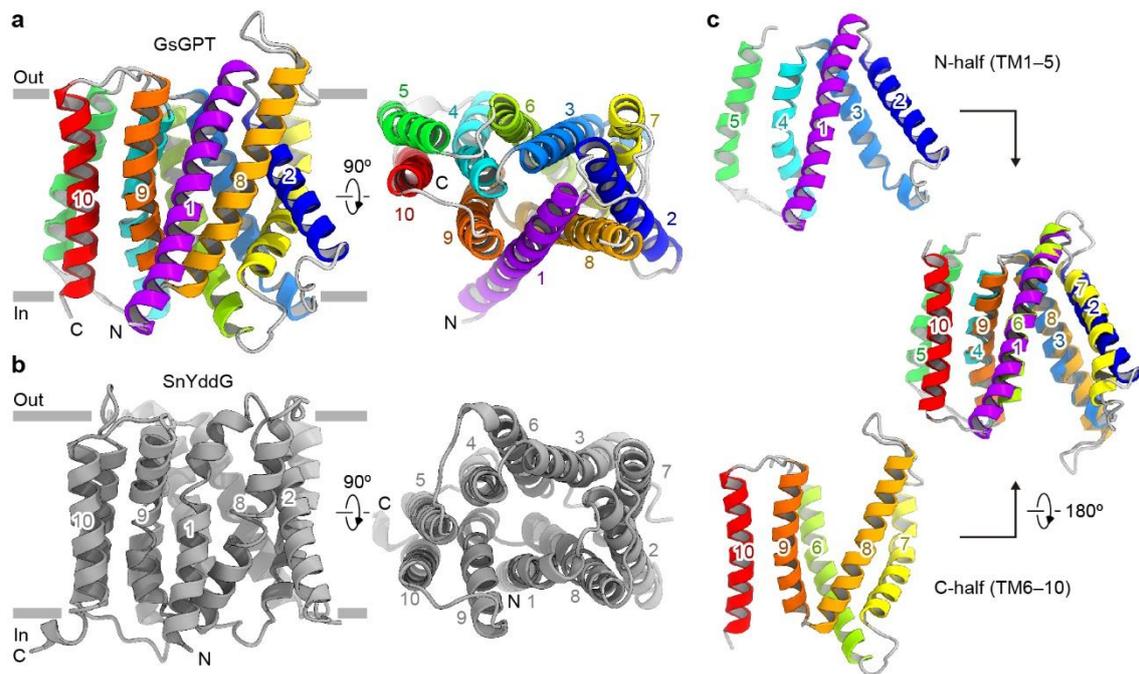
Figure 34 | Structural pseudo-symmetry of GsGPT.

(a) Transmembrane topology of GsGPT. TM1 to TM10 are color-coded in rainbow.

(b) Transmembrane topology of SnYddG, showing the same helix topology and overall fold.

(c) Superimposition of the two inverted structural repeats of GsGPT. The N-terminal (TM1–TM5; residues 100–245) and the C-terminal halves (TM6–TM10; residues 246–404) could be superimposed with an approximately 180⁰ rotation. Structures were superimposed with the SSM algorithm (*142*).
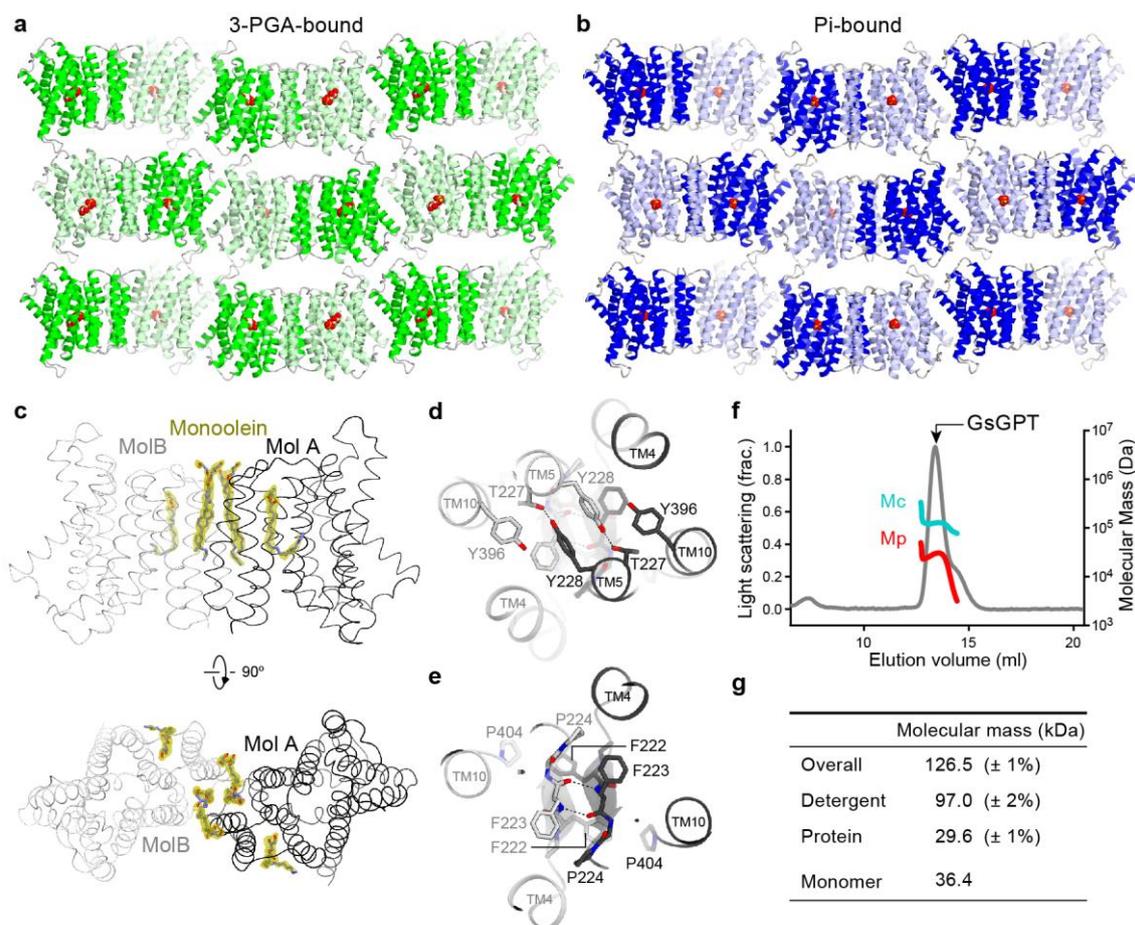
Figure 35 | Dimerization of GsGPT within the lipid bilayer

(a, b) Crystal packing of the 3-PGA- (a) and Pi-bound (b) structures.

(c) Monoolein molecules identified near the dimer interface. The 2*F*o-*F*c electron density maps of the six monoolein molecules, contoured at 1.0 σ, are overlaid onto the structure.

(d, e) Interactions at the dimer interface. Polar sidechains form hydrogen bonding interactions (d) and the loops connecting TM4 and TM5 form a short, two-stranded β-sheet between the two protomers (e).

(f) SEC-MALLS analysis of GsGPT. The gray chromatogram represents the reading of the light scattering detector, with values given on the left axis. The cyan and red curves indicate the calculated molecular masses of the protein-detergent complex (Mc) and the protein only (Mp), respectively, with values given on the right axis. The black arrow highlights the position of the elution peak of GsGPT. The refractive index increments (dn/dc) of the protein and the detergent were assumed to be 0.185 and 0.132, respectively (*143*).

(g) Molecular mass values determined by the SEC-MALLS experiment. The protein mass was determined to be about 29.6 kDa, corresponding to the theoretical mass of the GsGPT monomer, 36.4 kDa.
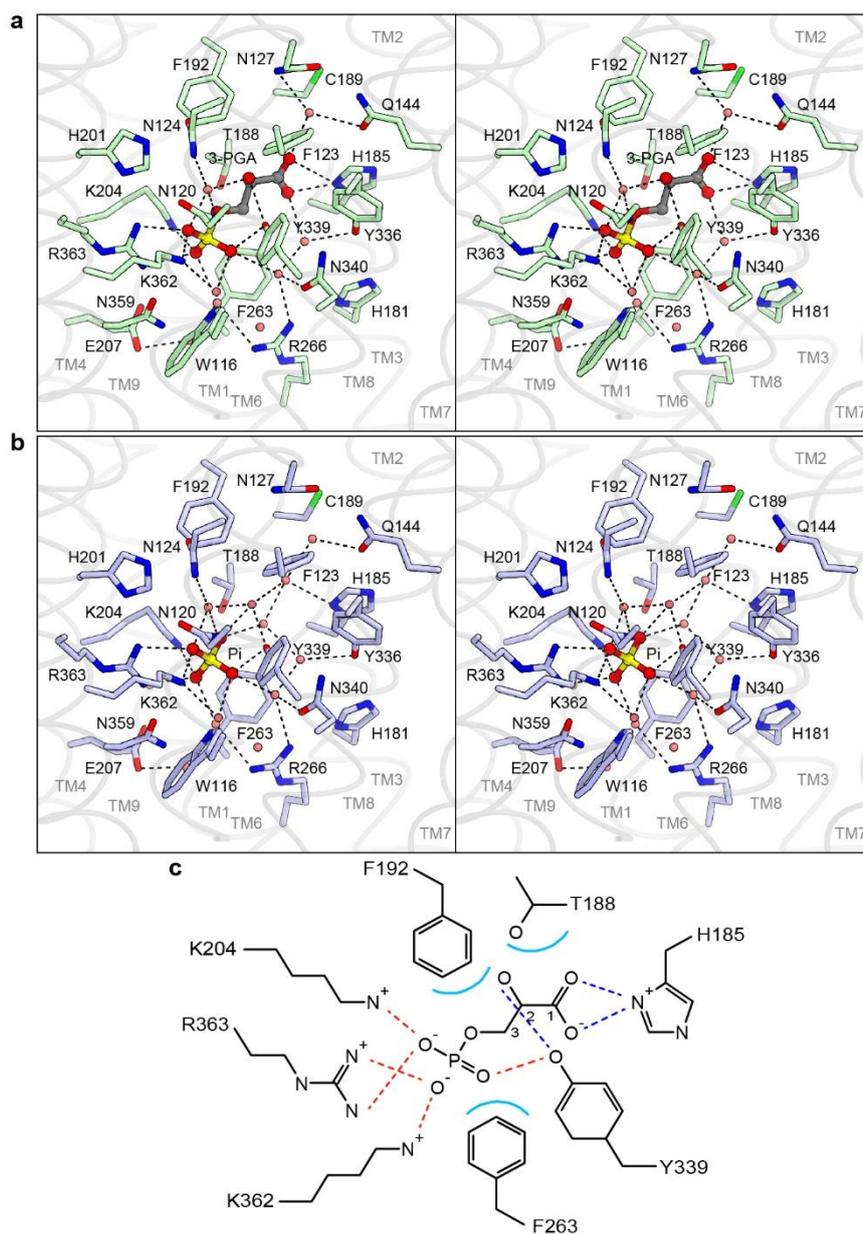
Figure 36 | 3-PGA and Pi recognition mechanisms

(a) Stereo view of the 3-PGA binding site. Dotted lines indicate polar interactions. Red balls represent water molecules.

(b) Stereo view of the Pi binding site.

(c) Schematic diagram of the 3-PGA coordination.

3.3.6 Similarity and diversity among pPT subtypes

A sequence comparison revealed that the four residues directly recognizing the phosphate (Lys204, Tyr339, Lys362 and Arg363) are strictly conserved in all higher plant pPTs (Figure 37 and Figure 38). Besides these residues, most of the residues near the phosphate moiety are also strictly conserved (Figure 37a). In contrast, the residues distant from the phosphate, or near the sugar moiety, are varied among the different pPT subtypes (Figure 37a). These findings suggest that the variant residues of the different pPT subtypes recognize the attached sugar moieties and thus determine their distinct substrate specificities, while the conserved residues similarly recognize the phosphate.

To better understand the substrate selectivities of the pPT family members (*74*), we generated homology models of five representative pPTs, namely *Arabidopsis thaliana* TPT (AtTPT), PPT1 (AtPPT1), GPT1 (AtGPT1), and XPT (AtXPT) and *Toxoplasma gondii* APT (TgAPT). (Figure 37b–m). The AtTPT model suggests that the plant TPTs similarly recognize the substrates as in the current GsGPT structure, since the residues recognizing 3-PGA are highly conserved (His184, Lys203, Tyr338, Lys359 and Arg360 in AtTPT) (Figure 37b,e,j). TPT prefers three-carbon compounds phosphorylated at the C3 position (triose-P and 3-PGA) to those phosphorylated at the C2 position (PEP and 2-PGA) by ~10-fold (*115*, *144*). The AtTPT model indicates that the 'branched' C3 methylene group of PEP would sterically clash with the bulky Phe262 sidechain, explaining the lower preference for PEP (*75*) (Figure 37c,f). In contrast, the C3 carbon of PEP can be accommodated in the widened pocket of the AtPPT1 model, where Phe is replaced by Asn262, consistent with the PPT's preference for PEP (*79*) (Figure 37d,g). The apicomplexan pPTs, including TgAPT, PfipPT and PfopPT, have 'dual specificity', as they transport both triose-P and PEP with similar affinities (*84*, *86*). The TgAPT model explains its dual specificity well, as it can accommodate both triose-P and PEP (Figure 37h,i).

GPT transports glucose-6-phosphate (Glc-6-P), the largest substrate of all pPTs, as well as smaller substrates such as triose-P and 3-PGA (*78*). The AtGPT1 model has the largest pocket space, which can accommodate the bulky C6 sugar (Figure 37k,l), consistent with its broad substrate specificity. As compared with GPT, the AtXPT model has a rather small pocket, which might be suitable for the C5 sugar moiety of its substrate, xylulose-5-phosphate (*69*) (Xul-5-P) (Figure 37m). Collectively, our crystal structures and the homology models address how different pPT members transport distinct sugar phosphates and thereby play diverse roles in plastid metabolism (*74*).
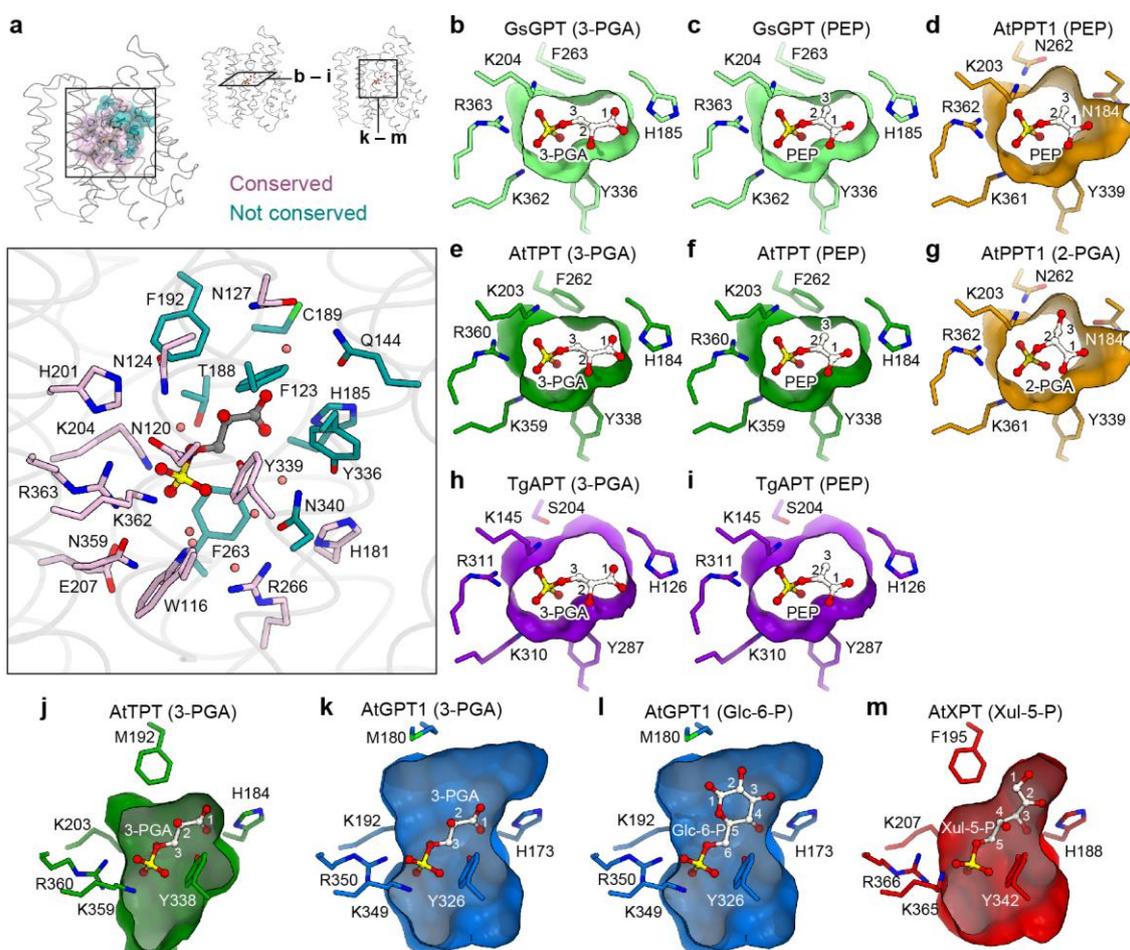
Figure 37 | Deducing the substrate specificities in different pPTs

(a) Amino-acid sequence conservation within the substrate-binding site of GsGPT. Conserved residues are colored violet and non-conserved residues are cyan.

(b–m) Homology-modelled structures of the substrate-binding sites of the pPTs. Key residues involved in substrate recognition are shown as stick models. Substrate molecules were modelled manually, based on the coordination of 3-PGA in GsGPT. Protein surfaces are shown for the regions around the substrate. In (c) and (f), the C3 carbon of PEP sterically clashed with the sidechain of Phe262, indicating non-preferable binding.
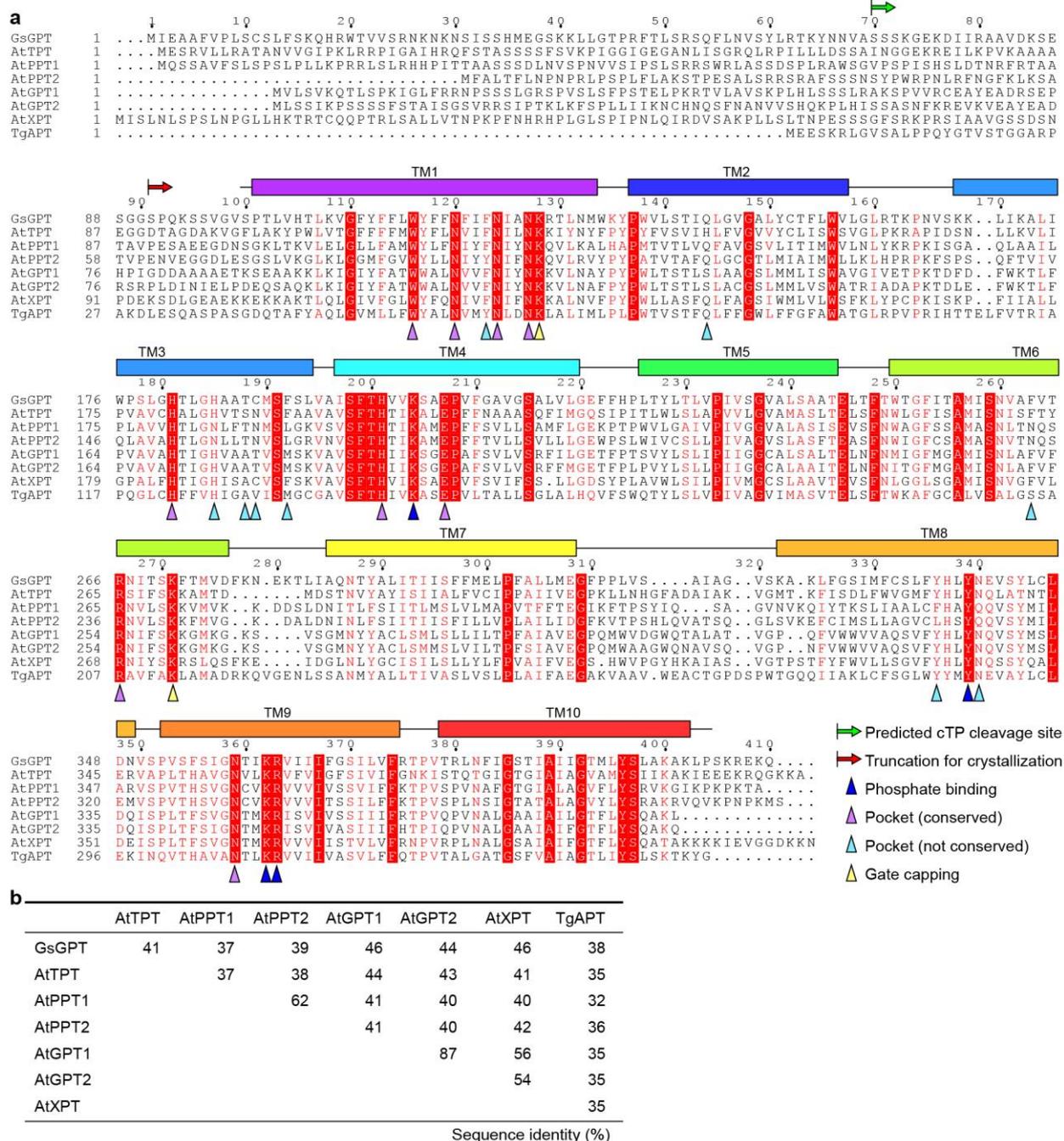
Figure 38 | Sequence alignment of the pPTs

(a) Amino acid sequence alignment of the pPTs.

(b) Sequence identity matrix of the pPTs. Identities were calculated for the mature translocator regions

(corresponding to residues 101–410 in GsGPT).

3.3.7 Basis of strict 1:1 exchange

Previous biochemical studies have shown that the transport by TPT is mediated by 'alternating-access' (*145*), in which the substrate-binding site is alternately exposed on either side of the membrane. In the current structure, the substrate is completely occluded from both sides of the membrane by the two gates (Figure 39a). The 'outside gate' is formed by Phe192 and Ile197 on the tips of TM3 and TM4, and seals the substrate from the outside solvent (Figure 39b). The 'inside gate' is formed by Leu347, Phe352 and Pro355 on the tips of TM8 and TM9, and similarly seals the substrate from the inside solvent (Figure 39c). The helix ends of both gates are further capped by the conserved Lys128 and Lys271 residues (Figure 39b,c).

To deduce the conformational change during the alternating-access, we compared this occluded structure of GsGPT with the available outward-open structure of the DMT transporter SnYddG. The structural superimposition revealed a prominent structural difference at TM3 and TM4 with a ~30º outward tilting in GsGPT (Figure 39a), suggesting that these helices undergo rocker-switch (*146*) movements to open and close the outside gate. The pseudo-symmetric structure of GsGPT suggests that similar motion would occur in the symmetrical counterpart, TM8 and TM9, to open and close the inside gate.

To further understand the conformational changes, we performed molecular dynamics simulations of GsGPT in the presence or absence of the bound Pi (Figure 40a–e). In the Pi-bound simulation, GsGPT did not undergo any significant structural change during 100 ns and remained in the occluded conformation (Figure 40f–i). In contrast, in the apo simulation, GsGPT underwent rapid conformational changes within about 10 ns to the inward-open or outward-open conformations, and stably adopted these open conformations until the end of the simulation (~100 ns) (Figure 40j–m). These conformational changes were consistent with our model proposed from the structural comparison with SnYddG, which involves the rocker-switch movements of the helix bundles TM3-TM4-TM6 and TM1-TM8-TM9 to open and close the two gates.

The different behaviors in the Pi-bound and apo simulations suggest that the conformational change of GsGPT is completely dependent on the substrate binding (Figure 41). Without a substrate, due to the electrostatic repulsion between the cationic residues (Lys204, Lys362 and Arg363) in the middle of the helix bundles, GsGPT prefers the outward- or inward-open states, as shown in the MD simulation. In contrast, phosphate or organic phosphate binding allows the close approximation of these cationic residues and thus leads to the occluded state, as in the current crystal structures. This ligand-dependent conformational change ensures the substrate-dependent transition between the inward- and outward-open states, and thus explains the strict

1:1 exchange kinetics of the pPTs (*76*).
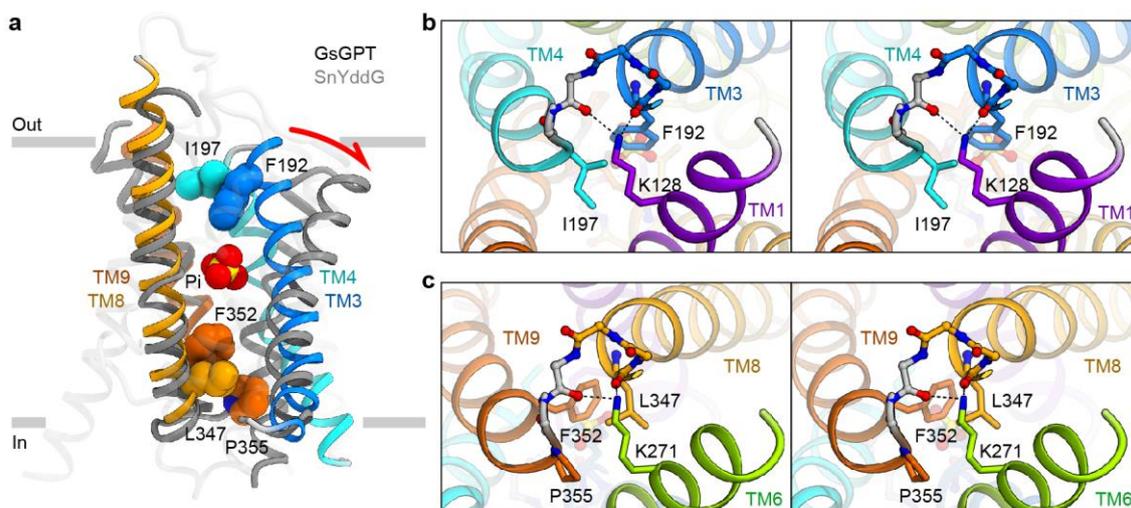
Figure 39 | Occluded structure and conformational change of GsGPT

(a) Superimposition of the occluded GsGPT structure and the outward-open SnYddG structure (PDB 5I20).

The substrate and gate-forming residues are shown as CPK models. The red arrow highlights the putative

rocker-switch movements in TM3 and TM4.

(b) Stereo view of the outside gate.

(c) Stereo view of the inside gate.
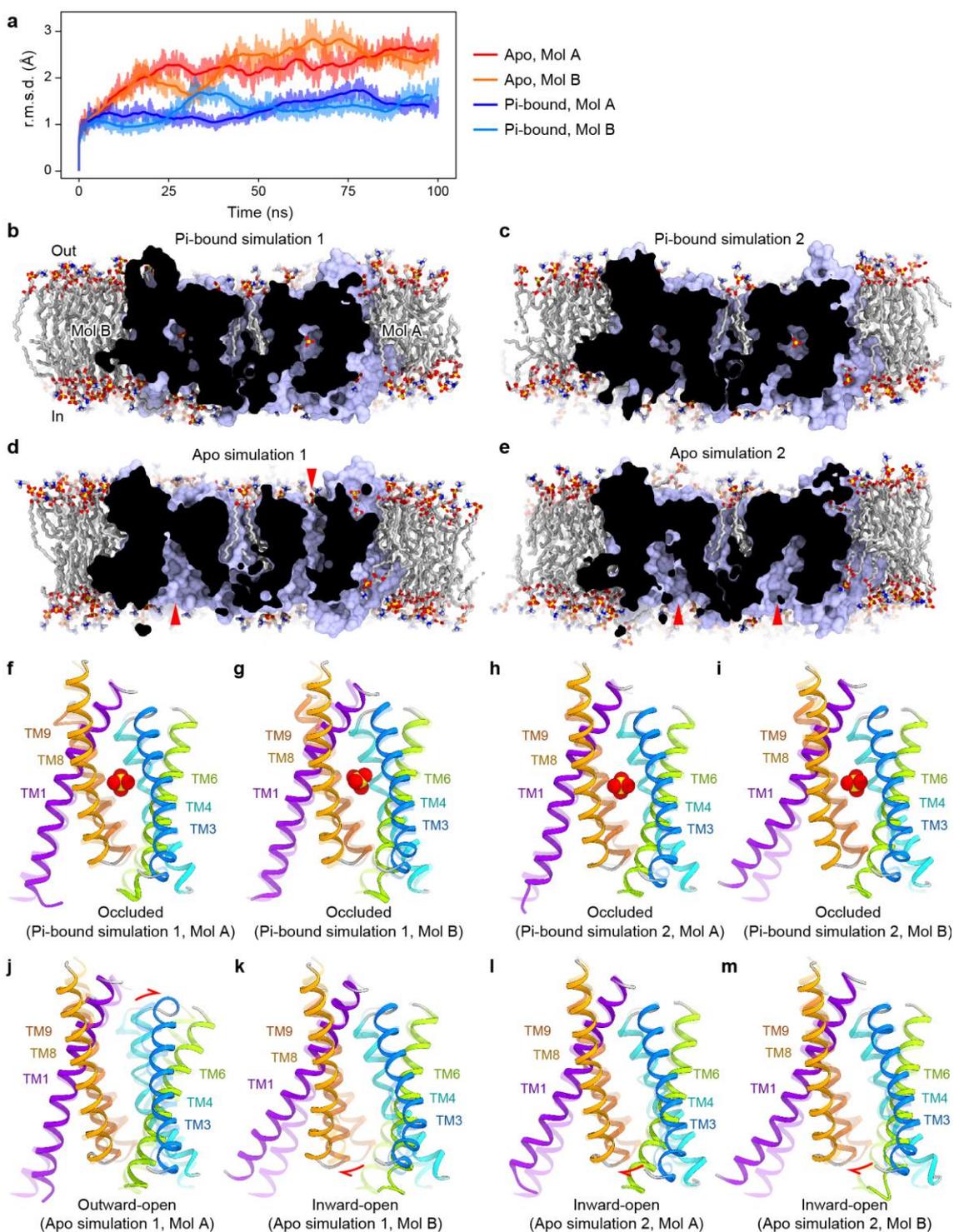
Figure 40 | Molecular dynamics simulation of GsGPT

(a) R.m.s.d. plot of each monomer (Mol A and B) in the Pi-bound and apo simulations.

(b–e) Final structures in the Pi-bound and apo simulations, performed in two independent runs. Red

arrowheads highlight the pocket opening. (f–m) Comparison of the initial (0 ns, transparent) and the final

structures (100 ns, opaque). Red arrows highlight helix movements.
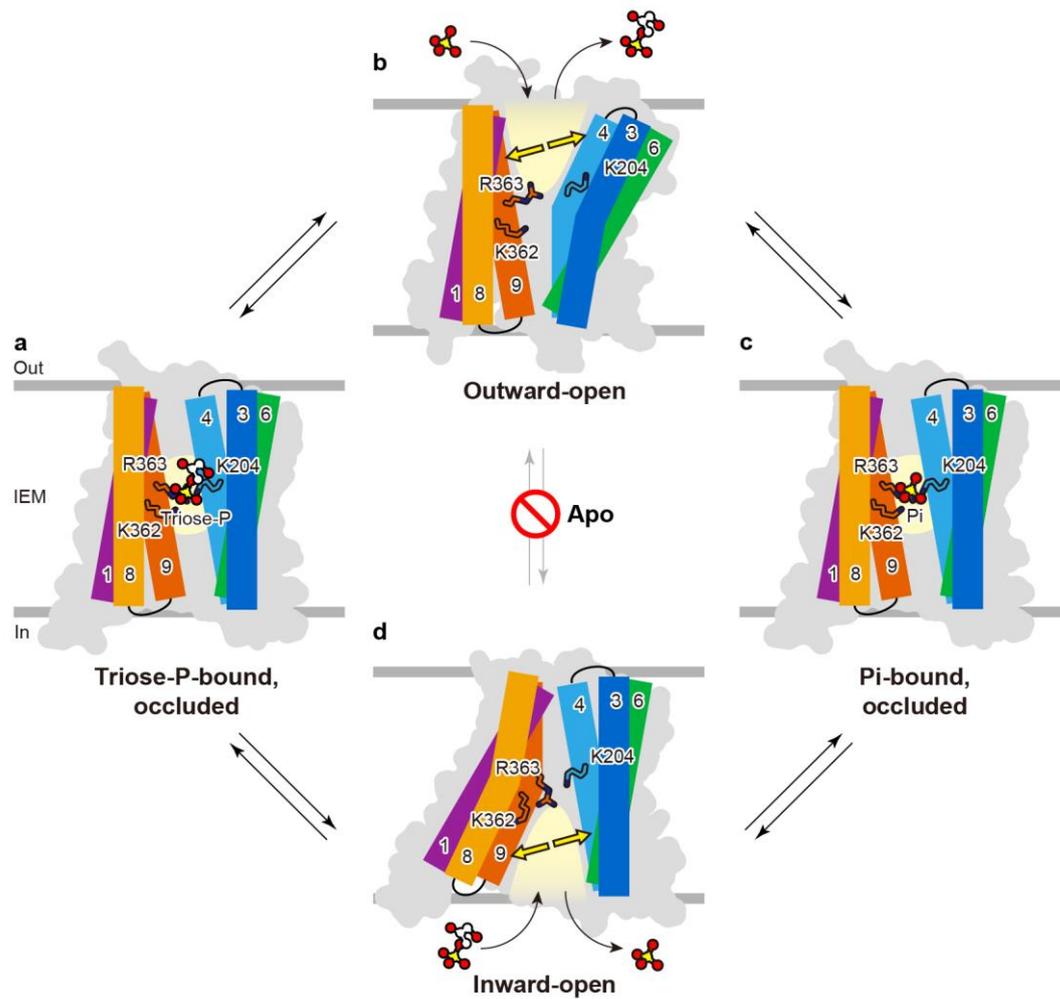
Figure 41 | Model of strict 1:1 exchange mechanism by GsGPT

(a–d) Illustration of a hypothetical conformational cycle of GsGPT. Substrate binding enables the association of the helix bundles (a, c), triggering the conformational transition. Without a substrate, the translocator cannot undergo the conformational transition due to the electrostatic repulsion (b, d).

3.4 Discussion

The proposed coupling mechanism between the substrate binding and the conformational change is quite different from the transport mechanism proposed for YddG, another 10-TM member of the DMT superfamily. YddG is a uniporter (*140*) that permeates substrates down a concentration gradient, indicating the lack of structural coupling. This difference could be explained by the composition of its substrate-binding site. The substrate-binding pocket of YddG mostly consists of hydrophobic residues (*140*), which would lack electrostatic repulsion. YddG can thus adopt the occluded state without any substrate, consistent with its uniporter function. Therefore, even though GsGPT and YddG share the similar 10-TM DMT fold, the different compositions of their substrate-binding sites result in distinct transport mechanisms.

Previous studies suggested that the members of the NST/TPT family share a common substrate-binding site (*147*). To explore this possibility, we created a sequence alignment of representative NST/TPT members (Figure 42). The alignment shows that the two phosphate-binding lysine residues (Lys204 and Lys362) are conserved at the corresponding positions in most plant NSTs (the KT, KV/A/G, KD and KR groups), supporting their proposed role in negative charge recognition (*147*). Meanwhile, these residues are not conserved in important animal NSTs (for example, the SLC35A subfamily), suggesting that these members might use distinct residues for recognizing the substrates. Mapping of known disease-causing mutations in NSTs (*148*) indicates their locations on the TM helices surrounding the central pocket, suggesting their involvement in substrate recognition. Taken together, these observations suggest that the pPTs and NSTs share a common substrate-binding site, but use distinct residues for recognizing their respective substrates.

In conclusion, we determined the high-resolution structures of TPT in complex with two counter-substrates. The structures resolve the long-standing controversy over its helix topology (*75*, *90*) and provide the framework to address its substrate recognition and strict 1:1 exchange mechanism. Further mechanistic understanding of the pPT family members could provide opportunities to engineer chloroplast transporters for improving crop productivity (*149*, *150*), or to develop new drugs targeting plastid organelles of apicomplexan parasites (*86*).

```
                    1        10        20        30        40        50        60        70        80
GsGPT               1 ...MIEAAFVPLSCSLFSKQHRWTVVSRNKNKNSISSHMEGSKKLLGTPRFTLSRSQFLNVSYLRTKYNNVASSSKGEKDIIRAAVD
AtTPT               1 ....MESRVLLRATANVVGIPKLRRPIGAIHRQFSTASSSSFSVKPIGGIGEGANLISGRQLRPILLLDSSAINGGEKREILKPVKA
AtGONST5|At1g21870    ......................................................................................
AtUAAT5|At1g12500     1 .....................MVEAQSWTTRRMSNPRFDAAATAAPTIVDIPGTP
AtUXT3|At1g06890      ......................................................................................
AtUAFT1|At5g25400     ......................................................................................
AtGONST1|At2g13650    1 ....................................................................MKLYEHDG
hSLC35C1|P78382       1 ...........................................................MNRAPLKRSRILHMAL
hSLC35D1|Q9NTN3       1 ...........................................................MAEVHRRQHARVKGEAP
hSLC35B2|Q8TB61       1 MDARWWAVVVLAAFPSLGAGGETPEAPPESWTQLWFFRFVVNAAGYASFMVPGYLLVQYFRRKNYLETGRGLCFPLVKACVFGNEPK
hSLC35A1|P78382       ......................................................................................
hSLC35A2|P78381       1 ..........................................................MAAVGAGGGSTA
hSLC35A3|Q9Y2D2       ......................................................................................

                                              TM1                                        TM2
                   90        100       110       120       130       140       150
GsGPT              85 KSESGGSPQKSSVGVSPTLVHTLKVGFYFFLWYFFNFIFNIANKRTLNM...........WKY.....PWVLSTIQLGVGALYCTF
AtTPT              84 AAAEGGDTAGDAKVGFLAKYPWLVTGFFFMWYFLNVIFNILNKKIYN........YFPY...PYFVSVIHLFVGVVYCLI
AtGONST5|At1g21870  1 .........MEEGSLWRQWTMFRSLLSILQWWGFNVTVIIMNKWIFQK.....LDFKF...PLSVSCVHFICSSIGAYI
AtUAAT5|At1g12500  35 PHSSASSPLKPFFLSSPTVSPTILTAAIIAAWFGSNIGVLLLNKYLLFY........YGFRY...PIFLTMTHMLSCAAYSSA
AtUXT3|At1g06890    1 ...........MSEGQKFQLGTIGALSLSVVSSVSIVICNKALIST.........LGFTF...ATTLTSWHLLVTFCSLHV
AtUAFT1|At5g25400   1 .............MGKGGSLSEGVIKNIIISYTYVAIWIFLSFTVIVYNKYILDKK.........MYDWPF...PISLTMIHMSFCSTLAFL
AtGONST1|At2g13650  9 VDLEDGKTVKSGGDKPIPRKIHNRALLSGLAYCISSCSMILVNKFVLSSY........NFNA...GIFLMLYQNFVSVIIVVG
hSLC35C1|P78382    17 TGASDPSAEAEANGEKPFLLRALQIALVVSLYWVTSISMVFLNKYLLDSP......SLRLDT...PIFVTFYQCLVTTLLCKG
hSLC35D1|Q9NTN3    18 AKSSTLRDEEELGMASAETLTVFLKLLAAGFYGVTLIFLLLYSVLTNY.......RFPS....SLCVGLGQMVATVAVLWV
hSLC35B2|Q8TB61    88 ASDEVPLAPRTEAAETTPMWQALKLLFCATGLQVSYLTWGVLQERVMTRSYGATATSPGERFTD.SQFLVLMNRVLALIVAGLSCV.
hSLC35A1|P78382     1 ..........MAAPRDNVTLLFKLYCLAVMTLMAAVYTIALRYTRTSD.....KELYFSTTAVCITEVIKLLLSVGIL.A
hSLC35A2|P78381    12 APGPGAVSAGALEPGTASAAHRRLKYISLAVLVVQNASLILSIRYARTLP.........GDRFFATTAVVMAEVKGLTCLLLLFA
hSLC35A3|Q9Y2D2      ...........MFANLKYVSLGILVFQTTSLVLTMRYSRTLK.........EEGPRYLSSTAVVAELLKIMACILLVYK

                                         TM3                              TM4
                  160       170       180       190       200       210       220
GsGPT             155 LW.VL.....GL.RTKPNV.........SKKL.IKALIWPSLGHTLGHAATCMSFSLVAISFTHVVKSAEPVFGAVGSALVLGEFFH
AtTPT             154 SW.SV.....GL.PKRAPI.........DSNL.LKVLIPVAVCHALGHVTSNVSFAAVAVSFTHTIKALEPFFNAAASQFIMGQSIP
AtGONST5|At1g21870 63 VIKVL.....KLKPLIVVD.......PEDR.WRRIFPMSFVFCINIVLGNISLRYIPVSFMQTIKSLTPATTVVLQWLVWRKYFD
AtUAAT5|At1g12500 107 VINIA.....GIVPRQHIL........SRRQ.FLKILSLSAIFCLSVVCGNTSLRYIPVSFNQAIGATTPFFTAVFSFLITCKTES
AtUXT3|At1g06890   59 AL.WM.....KMFE....H........KPFD.PRAVMGFGILNGISIGLLNLSLGFNSVGFYQMTKLAIIPCTVLLETLFFRKKFS
AtUAFT1|At5g25400  68 LIKVF.....KFVE.PVSM.......SRDTYLRSVVPIGALYSLSLWLSNSAYIYLSVSFIQMLKALMPVAVYSIGVLFKKEGFK
AtGONST1|At2g13650 81 LS.LM.....GLITTEPLTLRL.......MKVWFPVNVIFVGMLITSMFSLKYINVAMVTVLKNVTNVITAVGEMYLFNKQHD
hSLC35C1|P78382    91 LS.ALAACCPGAVDFPSLR.......LDLRV.ARSVLPSLVVFIGMITFNNLCLKYVGVAFYNVGRSLTTVFNVLLSYLLLKQTTS
hSLC35D1|Q9NTN3    90 GK.AL.....RVVKFPDLDRN......VP.....RKTFPLPLLYFGNQITGLFSTKKLNLPMFTVLRRFSILFTMFAEGVLLKKTFS
hSLC35B2|Q8TB61   173 LCK.Q.....PR..............HGAP.MYRYSFASLSNVLSSWCQYEALKFVSFPTQVLAKASKVIPVMLMGKLVSRRSYE
hSLC35A1|P78382    65 KETGS.....L...GRFKASLRENVLGSPKE.LLKLSVPSLVYAVQNNMAFLALSNLDAAVYQVTYQLKIPCTALCTVLMLNRTLS
hSLC35A2|P78381    89 QKRGN......V...KHLVLFLHEAVLVQYVD.TLKLAVPSLIYTLQNNLQYVAISNLPAATFQVTYQLKILTTALFSVLMLNRSLS
hSLC35A3|Q9Y2D2    61 DSKCS......L...RALNRVLHDEILNKPME.TLKLAIPSGIYTLQNNLLYVALSNLDAATYQVTYQLKILTTALFSVSMLSKKLG

              TM5                          TM6                              TM7
                  230       240       250       260       270       280       290
GsGPT             225 PLTYLTLVPIVSGVALSAATEL..........TFTWTGFITAMISNVAFVTRNITSKFTMVDF...KNEKTLIAQNTYALITIISFF
AtTPT             224 ITLWLSLAPVVLGVAMASLTEL..........SFNWLGFISAMISNISFTYRSIFSKKAMTD...MDSTNVYAYISIIALF
AtGONST5|At1g21870 135 WRIWASLVPIVGGILLTSITEL..........SFNVFGFCAALFGCLATSTKTILAESLL.......HGYKFDSINTVYYMAPFATM
AtUAAT5|At1g12500 179 TEVYLALLPVVSGIVLASNSEP..........SFHLFGFLICVASTAGRALKSVVQGIILTS.....ESEKLHSMNLLLYMAPMAAC
AtUXT3|At1g06890  126 RKIQFSLTILLLGVGIATVTDL..........QLNMLGSVLSLLAVVTTCVAQIMTNTIQKK......FKVSSTQLLYQSCPYQAI
AtUAFT1|At5g25400 140 SETMMNMLSISFGVAIAAYGEA..........RFDVWGVILQLGAVAFEATRLVMIQILLTS....KGITLNPITSLYYVAPCCLA
AtGONST1|At2g13650 151 NRVWAALFLMIISAVSGGITDL..........SFNAVGYAWQIANCFLTASYSLTLRKTMDTAKQVTQSGNLNEFSMVLLNNTLSLP
hSLC35C1|P78382   168 FYALLTCGIIIGGFWLGVDQEGAEG.......TLSWLGTVFGVLASLCVSLNAIYTTKVLP......AVDGSIWRLTFYNNVNACI
hSLC35D1|Q9NTN3   160 WGIKMTVFAMIIGAFVAASSDL..........AFDLEGYAFILINDVLTAANGAYVKQKLD......SKELGKYGLLYYNALFMIL
hSLC35B2|Q8TB61   237 HWEYLTATLISIGVSMFL..LSSGPEPR.SSPATTLSGLILLAGFAIGYVTGLQIKFTSPLTHNVSGTAKACAQTVLAVLYYEET
hSLC35A1|P78382   142 KLQWVSVFMLCAGVTLVQWKPAQA..TKVVVEQNPLLGFGAIAIAVLCSGFAGVYFEKVLKSS...DTSLWVRNIQMYLSGIIVTLA
hSLC35A2|P78381   166 RLQWASLLLLFTGVAIVQAQQAGGGGPRPLD.QNPGAGLAAVVASCLSSGFAGVYFEKILKGS...SGSVWLRNLQLGLFGTALGLV
hSLC35A3|Q9Y2D2   138 VYQWLSLVILMTGVAFVQWPSDSQLDSKELSAGSQFVGLMAVLTACFSSGFAGVYFEKILKET...KQSVWIRNIQLGFFGSIFGLM

          TM7                                TM8                           TM9
                  300       310       320       330       340       350       360       370
GsGPT             299 MELPFALLMEGFPPLVS......AIAGVSKA.KLFGSIMFCSLFYHLYNEVSYLCLDNVSPVSFSIGNTIKRVIIIFGSILVFR.T
AtTPT             292 VCIPPAIIVEGPKLLNHGFA...DAIAKVGMT.KFISDLFWVGMFYHLYNQLATNTLERVAPLTHAVGNVLKRVFVIGFSIVIFG.N
AtGONST5|At1g21870 205 ILGLPAFLLERNGILDWFEA.......HPSPWSALIILFNSGVLAFCLNFSIFYVIQSTTAVTFNVAGNLKVAVAVFVSWMIFR.N
AtUAAT5|At1g12500 251 ILLPFTLYIEGNVLRVLIEK.......ARTDPLIIFLLAGNATVAYLVNLTNFLVFKHTSALTLQVLGNGKAAVAGSVLIFR.N
AtUXT3|At1g06890  196 TLFVTGPFLDGLLTNQNVFA.......FKYTSQVVFFIVLSCLISVSVNFSTFLVIGKTSPVTYQVLGHLKTCLVLAFGYVLLR.D
AtUAFT1|At5g25400 212 FLFIPWVVEFPILRDTSSF.........HFDYLIFGTNSFCAFALNLAVFLLVGKTSALTMNVAGVVKDWLLIAFSWSVIK.D
AtGONST1|At2g13650 228 LGLLLSYFFNEMDYLYQT.........P.LLRLPSFWMVMTLSGLLGLAISFTSMWFLHQTGATTYSLVGSLNKIPLSIAGIVLF.NV
hSLC35C1|P78382   241 LFLPLLLLGELQALRDF..........A.QLGSAHFWGMMTLGGLFGFAIGYVTGLQIKFTS
hSLC35D1|Q9NTN3   230 PTLAIAYFTGDQAKAVEF.........E.GWADTLFLLQFTLSCVMGFILMYATVLCTQYNSALTTTIVGCIKNILITYIGMVFGGDY
hSLC35B2|Q8TB61   313 FTVG..SLLE......QGALLEGTRFM.GRHSEFAAHAL.LLSICSACGQLFIFYTIGQFGAAVFTIIMTLRQAFAILLSCL.LYGH
hSLC35A1|P78382   224 GVYL...SDGAEIKEKG.......FF.YGYTYYVWFVIFLASVGGLYTSV....VVKYTDNIMKGFSAAAAIVLSTIASVMLFG.L
hSLC35A2|P78381   249 GLWW......AEGTAVATRG.......FF.FGYTPAVWGVVLNQAFGGLLVAV....VKYADNILKGFATSLSIVSVASIRLFG.F
hSLC35A3|Q9Y2D2   222 GVYI....YDGELVSKNG.......FF.QGYNRLTWIVVVLQALGGLVIAA....VIKYADNILKGFATSLSIILSTLISYFWLQDF

          TM10
                  380       390       400       410
GsGPT             377 PVTRLNFIGSTIAIIGTMLYSLAKAKLPSKREKQ.......................
AtTPT             374 KISTQTGIGTGIAIAGVAMYSIIKAKIEEEKRQGKKA...................
AtGONST5|At1g21870 283 PISPMNAVGCGITLVGCTFYGYVRHMLSQQQPGTPRTPRTPRNKMELIPLVNDKLESKI
AtUAAT5|At1g12500 329 PVTVMGIAGFGVTIMGVVLYSSEARKRSKLLNQK.......................
AtUXT3|At1g06890  274 PFDWRNILGILVAVIGMVVYSYYCSIETQQKASETSTQLPQMKESEKDPLIAAENGSGVLSDGGGGVQQKTVAPVWNSNKDFQA
AtUAFT1|At5g25400 286 VTPINLFGYGIAFLGVAYYNHAKLQALKAKEAQKTAQQVDEETGRLLEEREGNEGGRKNEPED.
AtGONST1|At2g13650 305 PTSLQNSASILFGLVAVYFARAKMREKS............................
hSLC35C1|P78382   319 .KSFLWWTSNMMVLGGSSAYTWVRGWEMKKTPEEPSPKDSEKSAMGV.............
hSLC35D1|Q9NTN3   308 IFTWTNFIGLNISAGSLVYSYITTEEQLSEKQSEANNKLDIKGKGAV.............
hSLC35B2|Q8TB61   389 TVTVVGGLGVAVVFAALLRVYARGRLKQRGKKAVPVESPVQKV.................
hSLC35A1|P78382   294 QITLTFALGTLLVCVSIYLYGLPRQDTTSIQQGETASKERVIGV...............
hSLC35A2|P78381   319 HVDPLFALGAGLVIGAVYLYSLPRGAAKAIASASASASASGPCVHQQPPGQPPPPQLSSHRGDLITEPFLPKLLTKVKGS......
hSLC35A3|Q9Y2D2   293 VPTSVFFLGAILVITATFLYGYDPKPAGNPTKA.........................
```

Figure 42 | Alignment of GsGPT and NSTs.

Sequence alignment between selected NST/TPT family members. The multiple sequence alignment was calculated for 14 members from the human SLC35 subfamilies A–D (*148*), 50 members from the *Arabidopsis* NST/TPT family (*151*) and GsGPT, using Clustal Omega (*152*). Only selected sequences are shown, and gaps were manually refined based on the secondary structure. Blue boxes highlight the positions of the phosphate-binding residues in GsGPT.

Chapter 4 Concluding remarks

4.1 Summary of the thesis

In this thesis, I have determined the structures of SemiSWEET and TPT in different states that reflect their functional properties. In Chapter 2, I have described the structure of SemiSWEET, a bacterial homologue of SWEET, in the two opposite conformations. These structures revealed the molecular basis of 'alternating-access' in one of the smallest transporters identified to date. Although this study has only focused on a bacterial SWEET member, the conserved 3-TM repeat domain architecture suggests that the mechanistic insights gained here will also apply to eukaryotic SWEET proteins. In Chapter 3, I have described the structures of TPT bound to two different counter-substrates. These structures have defined the recognition mechanisms of phosphorylated metabolites and phosphate ion during photosynthetic product transport. The structures have also provided a model that accounts for the strict antiport activity of the pPT family proteins, which play important roles in phosphate homeostasis and regulation during photosynthesis.

4.2 Insights from other studies

After the publication of two original papers that constitute this thesis, several other groups have published the papers that are relevant to this study. In this section, I briefly discuss those recent findings and further implications on the function and mechanisms of these transporters.

4.2.1 Eukaryotic SWEETs

In 2015, the first structure of a eukaryotic SWEET transporter was reported (*153*) . The structure was of *Oryza sativa* SWEET2b (OsSWEET2b), the closest relative of *Arabidopsis* vacuolar sugar transporter AtSWEET2 (*63*). OsSWEET2b exhibits a 7-TM architecture, and the two 3-TM repeats (TM1–3 and TM5–7) form a substrate-translocation pathway, consistent with a SemiSWEET dimer (Figure 43a). The linker helix TM4 packs tightly against the first 3-TM repeat, but not the second one, playing structural roles rather than just being an inversion linker (*68*). This finding explains the previous observation that the split expression of TM1–4 and TM5–7 of AtSWEET1 remained functional, while TM1–3 and TM4–7 did not (*68*).

Notably, the structure of OsSWEET2b revealed a trimeric assembly, with the three protomers arranged in a non-crystallographic three-fold symmetry (*153*) (Figure 43b). The same trimeric assembly was consistently observed in two different crystal forms, supporting its functional relevance. A previous study has proposed the oligomerization of SWEET on the basis of dominant negative effects of several mutants to the wild-type proteins (*68*). The trimeric architecture of

OsSWEET2b thus explains such phenomenon based on the inter-protomer structural coupling, that is, the mutant protein inhibits the conformational change of the wild-type proteins within the trimer (Figure 43c).

In 2017, another group reported the structure of *Arabidipsis thaliana* SWEET13 (AtSWEET13) (*154*). AtSWEET13 has been identified as a sucrose transporter, and the structure contained a substrate analog, 2'-deoxycytidine-5'-monophosphate (dCMP), in its inward-open substrate-binding pocket. The asparagine (Asn76 and Asn176) and the tryptophan pairs (Trp58 and Trp180), which correspond to Trp50 and Asn66 in *E. coli* SemiSWEET, form hydrogen-bonding and stacking interactions with dCMP, consistent with the model that these residues recognize the substrate. The crystallization construct of AtSWEET13 contained four thermostabilizing mutations, V23L, S54N, V145M and S176N, located at the substrate-binding pocket. Interestingly, equivalent substitutions are found in the SWEET members that transport monosaccharides. Consistent with this, the crystallization construct showed the decrease in the sucrose transport activity, but retained a glucose transport activity. Therefore, these data indicate that the pocket size determines the mono- and disaccharide selectivity of different SWEET family members.
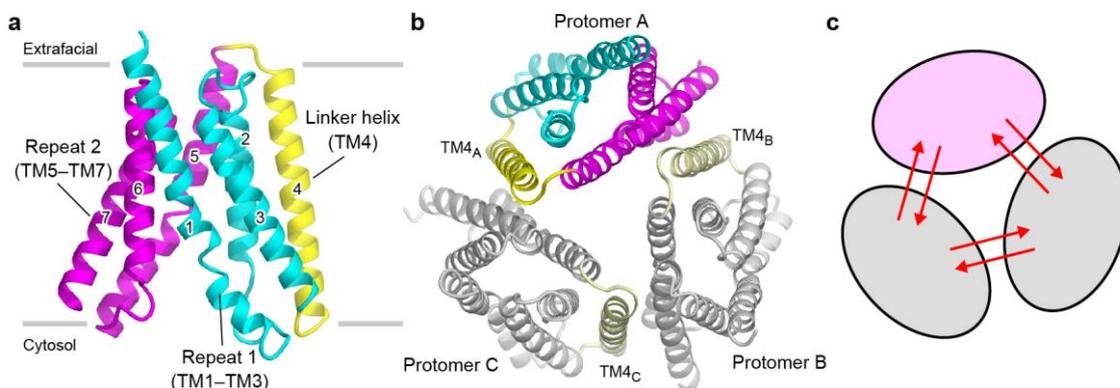
Figure 43 | Structure of OsSWEET2b

(a) Ribbon representation of the OsSWEET2b protomer. The first 3-TM repeat (TM1–TM3) is colored in cyan, and the second one (TM5–TM7) is in magenta. The linker helix (TM4) is colored in yellow.

(b) Trimeric assembly of OsSWEET2b. Protomer A is colored as in (a), and protomers B and C are in grey. TM4, colored in yellow, participates in the oligomerization.

(c) Schematics of the inter-protomer structural coupling. The ellipses represent protomers.
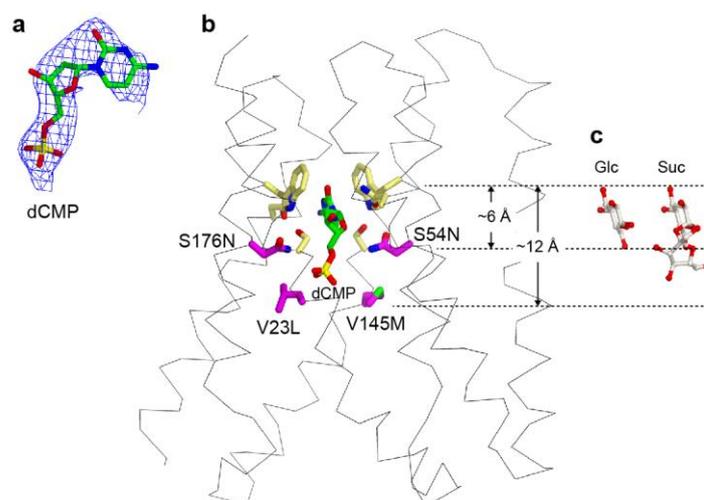
Figure 44 | Structure of AtSWEET13 and its substrate recognition mechanism

(a) The 2*F*o-*F*c electron density map of 2'-deoxycytidine-5'-monophosphate (dCMP), contoured at 1.0σ.

(b) The structure of AtSWEET13, highlighting the residues in the substrate-binding pocket. The mutated residues (V23L, S54N, V145M and S176N) are colored in magenta.

(c) Sizes of glucose (Glc) and sucrose (Suc). The mutations introduced in AtSWEET13 would reduce the pocket size from approximately 12 Å to 6 Å, making it too narrow for a sucrose to be accommodated. Note, the structure is in the inward-open conformation, and these mutated residues should come closer to each other upon conformational change, further constricting the substrate-binding site.

4.2.2 Molecular dynamics simulation

In 2017, a study reported the molecular dynamics simulation of SemiSWEET (*155*). The authors performed long-time, unbiased simulations (~2 μs) starting from the occluded conformation of *L. biflexa* SemiSWEET (*109*) and from the outward-open conformation that were newly determined. The authors conclude that 1) the gates play important roles not only in blocking substrate passage but also in driving the conformational changes by forming favorable interactions, 2) the formation and breakage of hydrogen-bonding interactions during the conformational change appear to offset each other to lower the activation energy during the transport cycle and 3) the presence or absence of the substrate does not largely affect the transporter conformation. These findings corroborate our facilitative diffusion model, since they explain the 'driving force' of conformational changes in the presence or absence of a substrate.

4.2.3 Nucleotide sugar transporters (NSTs)

The pPTs are closely related to the Golgi-localized nucleotide sugar transporters (NSTs), which are conserved in all eukaryotes. In 2017, the first structure of a yeast NST, known as Vrg4, was reported (*156*). The structure of Vrg4 revealed a DMT-superfamily fold similar to SnYddG and GsGPT, with 10 TM helices surrounding the central substrate-binding pocket. The structures were determined in the absence or presence of a substrate, GDP–mannose. Although the resolution was limited at 3.6 Å, the authors could model the substrate, with its mannose moiety pointing toward TM4 and TM9, and its GDP moiety toward TM2 and TM8, corroborating previous biochemical data (*157*, *158*).

The most intriguing finding was that Vrg4 requires a short-chain lipid (mirystoyl fatty acid; 14 carbons) for optimal activity in the liposome assays (*156*). The authors reasoned that the hydrophobic mismatch causes the inactivation of Vrg4 in the liposomes with long-chain lipids (16 and 18 carbons). The Golgi is known to have a thinner bilayer because of its different lipid composition to that of the plasma membrane (*159*). Consistent with this, Vrg4 adopts a compact structure with an apparent hydrophobic thickness of ~26 Å (Figure 45), significantly thinner than those of typical plasma membrane proteins, which are approximately 31–35 Å (*160*). The authors proposed that lipids may play regulatory roles during the dynamic membrane trafficking process in the Golgi complex (*161*, *162*), so that Vrg4 only functions in the correct cellular context. However, whether the bilayer thickness itself or the binding of specific lipids is important for activity is currently unclear. We found that, when compared with Vrg4, GsGPT also adopts a compact structure with an apparent hydrophobic thickness of ~27 Å. Thus, GsGPT might also favor thinner bilayer for proper function.
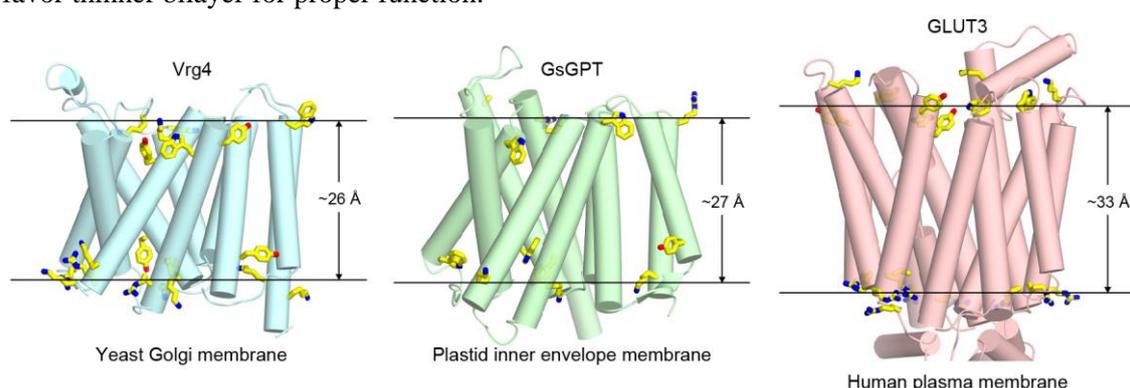
Figure 45 | Structure of Vrg4 and hydrophobic thickness

The structures of the Golgi-localized nucleotide sugar transporter Vrg4 (left), the plastid inner envelope membrane-localized GsGPT (center) and the plasma membrane-localized glucose transporter GLUT3 (right). The hydrophobic thicknesses were calculated on PPM web server (*163*). Membrane-anchoring residues (Trp, Tyr, Lys and Arg) are shown as stick models.

4.3 Perspectives

Although our and other's studies have provided important insights into the structure and functions of SWEET and pPT family proteins, several questions still remain. First, plant SWEET paralogues are known to form homo- and hetero-oligomers (*68*), but how one paralogue preferentially associates with a specific paralogue to the others is not understood. Therefore, future studies should address the structural basis of selectivity in homo- and hetero-oligomerization of SWEETs. Second, although the structure of AtSWEET13 provided some insights into mono- or disaccharide discrimination mechanism, detailed recognition mechanisms for different sugars, including glucose, fructose, galactose and sucrose, remain elusive (*4*, *52–57*). As did the high-resolution resolution structure of GLUT3 reveal its α- and β-anomer recognition mechanisms (*164*), detailed structural and functional analyses are required to fully understand the sugar selectivity mechanisms of different SWEET family proteins. Third, the structures of GsGPT have suggested possible involvement of lipid molecules in dimerization, but its physiological relevance remains to be characterized. The lipid compositions of the chloroplast envelope membranes are known to be different from that of the plasma membrane or the chloroplast thylakoid membrane (*165–167*). Future studies should address the role of lipids on the structure, dynamics and transport activity of the pPTs. Finally, the structure of GsGPT described here is hitherto the only known structure of the plastid envelope protein (*168*). Structural elucidation of other members will allow further understanding on the unique aspects of plastid envelope proteins, such as membrane insertion mechanism (*169–172*), evolutionary origin (*173–176*), post-translational modification (*177–179*), redox regulation (*180*, *181*) and protein-protein or lipid-protein interactions (*182–187*).

Acknowledgements

References

1. S. Lalonde, D. Wipf, W. B. Frommer, Transport mechanisms for organic forms of carbon and nitrogen between source and sink., *Annu Rev Plant Biol* **55**, 341–72 (2004).

2. J. Schroeder *et al.*, Using membrane transporters to improve crops for sustainable food production, *Nature* **497**, 60–66 (2013).

3. M. H. Saier *et al.*, The major facilitator superfamily., *J. Mol. Microbiol. Biotechnol.* **1**, 257–79 (1999).

4. L.-Q. Chen *et al.*, Sugar transporters for intercellular exchange and nutrition of pathogens, *Nature* **468**, 527–532 (2010).

5. M. Zimmermann, J. Milburn, Transport in plants I: Phloem transport, **1** (1975), doi:10.1007/978-3-642-66161-.

6. N. Sauer, J. Stolz, SUC1 and SUC2: two sucrose transporters from Arabidopsis thaliana; expression and characterization in baker's yeast and identification of the histidine-tagged protein, *Plant J* **6**, 67–77 (1994).

7. J. W. Riesmeier, B. Hirner, W. B. Frommer, Potato sucrose transporter expression in minor veins indicates a role in phloem loading., *Plant Cell* **5**, 1591–8 (1993).

8. J. W. Riesmeier, L. Willmitzer, W. B. Frommer, Isolation and characterization of a sucrose carrier cDNA from spinach by functional expression in yeast., *EMBO J.* **11**, 4705–13 (1992).

9. N. Sauer *et al.*, AtSUC8 and AtSUC9 encode functional sucrose transporters, but the closely related AtSUC6 and AtSUC7 genes encode aberrant proteins in different Arabidopsis ecotypes, *Plant J* **40**, 120–130 (2004).

10. N. Aoki, T. Hirose, G. N. Scofield, P. R. Whitfeld, R. T. Furbank, The sucrose transporter gene family in rice., *Plant Cell Physiol.* **44**, 223–32 (2003).

11. C. Kühn, C. Grof, Sucrose transporters of higher plants, *Curr Opin Plant Biol* **13**, 287–297 (2010).

12. D. R. Bush, Electrogenicity, pH-Dependence, and Stoichiometry of the Proton-Sucrose Symport., *Plant Physiol.* **93**, 1590–6 (1990).

13. R. GIAQUINTA, Possible role of pH gradient and membrane ATPase in the loading of sucrose into the sieve tubes, *Nature* **267**, 369–370 (1977).

14. C. Kühn *et al.*, The sucrose transporter StSUT1 localizes to sieve elements in potato tuber phloem and influences tuber physiology and development., *Plant Physiol.* **131**, 102–13 (2003).

15. Kühn, A Comparison of the Sucrose Transporter Systems of Different Plant Species, *Plant Biology* **5**, 215–232 (2003).

16. Y. Zhou, H. Qu, K. Dibley, C. Offler, J. Patrick, A suite of sucrose transporters expressed in

coats of developing legume seeds includes novel pH-independent facilitators, *The Plant Journal* **49**, 750–764 (2007).

17. A. Sivitz *et al.*, Arabidopsis Sucrose Transporter AtSUC9. High-Affinity Transport Activity, Intragenic Control of Expression, and Early Flowering Mutant Phenotype, *Plant Physiology* **143**, 188–198 (2006).

18. W. D. Hitz, P. J. Card, K. G. Ripp, Substrate recognition by a sucrose transporting protein., *J. Biol. Chem.* **261**, 11986–91 (1986).

19. Y.-L. Ruan, Sucrose Metabolism: Gateway to Diverse Carbon Use and Sugar Signaling, *Annual Review of Plant Biology* **65**, 33–67 (2014).

20. A. Sturm, Invertases. Primary structures, functions, and roles in plant development and sucrose partitioning., *Plant Physiol.* **121**, 1–8 (1999).

21. Y.-L. Ruan, Y. Jin, Y.-J. Yang, G.-J. Li, J. Boyer, Sugar Input, Metabolism, and Signaling Mediated by Invertase: Roles in Development, Yield Potential, and Response to Drought and Heat, *Mol Plant* **3**, 942–955 (2010).

22. N. Sauer, W. Tanner, The hexose carrier from Chlorella, *FEBS Letters* **259**, 43–46 (1989).

23. N. Sauer, K. Friedländer, U. Gräml-Wicke, Primary structure, genomic organization and heterologous expression of a glucose transporter from Arabidopsis thaliana., *EMBO J.* **9**, 3045–50 (1990).

24. M. Büttner, N. Sauer, Monosaccharide transporters in plants: structure, function and physiology, *Biochimica Et Biophysica Acta Bba - Biomembr* **1465**, 263–274 (2000).

25. M. Büttner, The monosaccharide transporter(-like) gene family in Arabidopsis, *Febs Lett* **581**, 2318–2324 (2007).

26. M. Büttner, The Arabidopsis sugar transporter (AtSTP) family: an update, *Plant Biol* **12**, 35–41 (2010).

27. E. Truernit, J. Schmid, P. Epple, J. Illig, N. Sauer, The sink-specific and stress-regulated Arabidopsis STP4 gene: enhanced expression of a gene encoding a monosaccharide transporter by wounding, elicitors, and pathogen challenge., *Plant Cell* **8**, 2169–82 (1996).

28. Noiraud, Maurousset, Lemoine, Identification of a mannitol transporter, AgMaT1, in celery phloem., *Plant Cell* **13**, 695–705 (2001).

29. S. Schneider *et al.*, Arabidopsis INOSITOL TRANSPORTER2 Mediates H+ Symport of Different Inositol Epimers and Derivatives across the Plasma Membrane, *Plant Physiol* **145**, 1395–1407 (2007).

30. S. Schneider *et al.*, Arabidopsis INOSITOL TRANSPORTER4 Mediates High-Affinity H+ Symport of Myoinositol across the Plasma Membrane, *Plant Physiol* **141**, 565–577 (2006).

31. Y.-S. Klepek *et al.*, Arabidopsis POLYOL TRANSPORTER5, a New Member of the Monosaccharide Transporter-Like Superfamily, Mediates H+-Symport of Numerous

Substrates, Including myo-Inositol, Glycerol, and Ribose, *Plant Cell* **17**, 204–218 (2005).

32. N. Noiraud, L. Maurousset, R. Lemoine, Transport of polyols in higher plants, *Plant Physiol Bioch* **39**, 717–728 (2001).

33. A. Wormit *et al.*, Molecular Identification and Physiological Characterization of a Novel Monosaccharide Transporter from Arabidopsis Involved in Vacuolar Sugar Transport, *Plant Cell Online* **18**, 3476–3490 (2006).

34. S. Aluri, M. Büttner, Identification and functional expression of the Arabidopsis thaliana vacuolar glucose transporter 1 and its role in seed germination and flowering, *Proc National Acad Sci* **104**, 2537–2542 (2007).

35. Y. Mathieu *et al.*, Regulation of Vacuolar pH of Plant Cells: I. Isolation and Properties of Vacuoles Suitable for P NMR Studies., *Plant Physiol.* **89**, 19–26 (1989).

36. M. THOM, E. KOMOR, Role of the ATPase of sugar-cane vacuoles in energization of the tonoplast, *Eur J Biochem* **138**, 93–99 (1984).

37. K. Yamada *et al.*, Functional Analysis of an Arabidopsis thaliana Abiotic Stress-inducible Facilitated Diffusion Transporter for Monosaccharides, *J Biol Chem* **285**, 1138–1146 (2010).

38. T. Kiyosue, H. Abe, K. Yamaguchi-Shinozaki, K. Shinozaki, ERD6, a cDNA clone for an early dehydration-induced gene of Arabidopsis, encodes a putative sugar transporter1The nucleotide sequence reported in this paper has been submitted to DDBJ with the accession number of D89051.1, *Biochimica Et Biophysica Acta Bba - Biomembr* **1370**, 187–191 (1998).

39. A. Weber *et al.*, Identification, purification, and molecular cloning of a putative plastidic glucose translocator., *Plant Cell* **12**, 787–802 (2000).

40. T. L. Slewinski, Diverse functional roles of monosaccharide transporters and their homologs in vascular plants: a physiological perspective., *Mol Plant* **4**, 641–62 (2011).

41. M. D. Marger, M. H. Saier, A major superfamily of transmembrane facilitators that catalyse uniport, symport and antiport., *Trends in biochemical sciences* **18**, 13–20 (1993).

42. P. J. Henderson, M. C. Maiden, Homologous sugar transport proteins in Escherichia coli and their relatives in both prokaryotes and eukaryotes., *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **326**, 391–410 (1990).

43. R. Rubin, S. Levy, R. Heinrikson, F. Kézdy, Gene duplication in the evolution of the two complementing domains of Gram-negative bacterial tetracycline efflux proteins, *Gene* **87**, 7–13 (1990).

44. J. Abramson *et al.*, Structure and Mechanism of the Lactose Permease of *Escherichia coli*, *Science* **301**, 610–615 (2003).

45. H. Kumar *et al.*, Structure of sugar-bound LacY, *Proceedings of the National Academy of Sciences* **111**, 1784–1788 (2014).

46. C. Law, P. Maloney, D.-N. Wang, Ins and Outs of Major Facilitator Superfamily Antiporters, *Annu Rev Microbiol* **62**, 289–305 (2008).

47. N. Yan, Structural advances for the major facilitator superfamily (MFS) transporters, *Trends Biochem Sci* **38**, 151–159 (2013).

48. N. Yan, Structural Biology of the Major Facilitator Superfamily Transporters, *Annu Rev Biophys* **44**, 257–283 (2015).

49. E. Quistgaard, C. Löw, F. Guettou, P. Nordlund, Understanding transport by the major facilitator superfamily (MFS): structures pave the way, *Nat Rev Mol Cell Bio* **17**, 123–132 (2016).

50. N. Sauer, Molecular physiology of higher plant sucrose transporters, *Febs Lett* **581**, 2309–2317 (2007).

51. M. Yuan, S. Wang, Rice MtN3/Saliva/SWEET Family Genes and Their Homologs in Cellular Organisms, *Mol Plant* **6**, 665–674 (2013).

52. F. Chardon *et al.*, Leaf Fructose Content Is Controlled by the Vacuolar Transporter SWEET17 in Arabidopsis, *Curr Biol* **23**, 697–702 (2013).

53. W.-J. Guo *et al.*, SWEET17, a Facilitative Transporter, Mediates Fructose Transport across the Tonoplast of Arabidopsis Roots and Leaves, *Plant Physiol* **164**, 777–789 (2014).

54. Y. Zhou *et al.*, Overexpression of OsSWEET5 in Rice Causes Growth Retardation and Precocious Senescence, *PLoS ONE* **9** (2014), doi:10.1371/journal.pone.0094210.

55. X. Wei, F. Liu, C. Chen, F. Ma, M. Li, The Malus domestica sugar transporter gene family: identifications based on genome and expression profiling related to the accumulation of fruit sugars., *Front Plant Sci* **5**, 569 (2014).

56. W. Lin *et al.*, Nectar secretion requires sucrose phosphate synthases and the sugar transporter SWEET9, *Nature* **508**, 546–549 (2014).

57. L.-Q. Chen *et al.*, Sucrose Efflux Mediated by SWEET Proteins as a Key Step for Phloem Transport, *Science* **335**, 207–211 (2012).

58. L. Chen, SWEET sugar transporters for phloem transport and pathogen nutrition, *New Phytol* **201**, 1150–1155 (2014).

59. F. Baker, K. Leach, D. Braun, SWEET as Sugar: New Sucrose Effluxers in Plants, *Mol Plant* **5**, 766–768 (2012).

60. D. Braun, SWEET! The Pathway Is Complete, *Science* **335**, 173–174 (2012).

61. A. Vassilyev, On the mechanisms of nectar secretion: revisited, *Ann Bot-london* **105**, 349–354 (2010).

62. D. Sosso *et al.*, Seed filling in domesticated maize and rice depends on SWEET-mediated hexose transport, *Nat Genet* **47**, 1489–1493 (2015).

63. H. Chen *et al.*, The Arabidopsis vacuolar sugar transporter SWEET2 limits carbon

sequestration from roots and restricts Pythium infection, *Plant J* **83**, 1046–1058 (2015).

64. P. Klemens *et al.*, Overexpression of the Vacuolar Sugar Carrier AtSWEET16 Modifies Germination, Growth, and Stress Tolerance in Arabidopsis, *Plant Physiol* **163**, 1338–1352 (2013).

65. L.-Q. Chen *et al.*, A Cascade of Sequentially Expressed Sucrose Transporters in the Seed Coat and Endosperm Provides Nutrition for the Arabidopsis Embryo, *Plant Cell* **27**, 607–619 (2015).

66. I. Kryvoruchko *et al.*, MtSWEET11, a Nodule-Specific Sucrose Transporter of Medicago truncatula Root Nodules, *Plant Physiology* , pp.01910.2015 (2016).

67. Y. Kanno *et al.*, AtSWEET13 and AtSWEET14 regulate gibberellin-mediated physiological processes., *Nat Commun* **7**, 13245 (2016).

68. Y. H. Xuan *et al.*, Functional role of oligomerization for bacterial and plant SWEET sugar transporter family., *Proc. Natl. Acad. Sci. U.S.A.* **110**, E3685–94 (2013).

69. R. Lilley, C. Chon, A. Mosbach, H. Heldt, The distribution of metabolites between spinach chloroplasts and medium during photosynthesis in vitro, *Biochimica Et Biophysica Acta Bba - Bioenergetics* **460**, 259–272 (1977).

70. H. W. Heldt, L. Rapley, Specific transport of inorganic phosphate, 3-phosphoglycerate and dihydroxyacetonephosphate, and of dicarboxylates across the inner membrane of spinach chloroplasts., *FEBS Lett.* **10**, 143–148 (1970).

71. K. Fischer, A. Weber, Transport of carbon in non-green plastids, *Trends Plant Sci* **7**, 345–351 (2002).

72. R. Fliege, U. Flügge, K. Werdan, H. Heldt, Specific transport of inorganic phosphate, 3-phosphoglycerate and triosephosphates across the inner membrane of the envelope in spinach chloroplasts., *Biochim Biophys Acta* **502**, 232–47 (1978).

73. H. Heldt, L. Rapley, Unspecific permeation and specific uptake of substances in spinach chloroplasts, *FEBS Letters* **7**, 139–142 (1970).

74. A. Weber, N. Linka, Connecting the Plastid: Transporters of the Plastid Envelope and Their Role in Linking Plastidial with Cytosolic Metabolism, *Annu Rev Plant Biol* **62**, 53–77 (2011).

75. A. P. Weber, R. Schwacke, U.-I. I. Flügge, Solute transporters of the plastid envelope membrane., *Annu Rev Plant Biol* **56**, 133–64 (2005).

76. F. Facchinelli, A. Weber, The Metabolite Transporters of the Plastid Envelope: An Update, *Front Plant Sci* **2** (2011), doi:10.3389/fpls.2011.00050.

77. U. Flügge, H. Heldt, The phosphate-triose phosphate-phosphoglycerate translocator of the chloroplast, *Trends Biochem Sci* **9**, 530–533 (1984).

78. C. R. Stocking, S. Larson, A chloroplast cytoplasmic shuttle and the reduction of extraplastid NAD., *Biochem. Biophys. Res. Commun.* **37**, 278–82 (1969).

79. Fischer *et al.*, A new class of plastidic phosphate translocators: a putative link between primary and secondary metabolism by the phosphoenolpyruvate/phosphate antiporter., *Plant Cell* **9**, 453–62 (1997).

80. U.-I. I. Flügge, R. E. Häusler, F. Ludewig, M. Gierth, The role of transporters in supplying energy to plant plastids., *J. Exp. Bot.* **62**, 2381–92 (2011).

81. M. Eicks, V. Maurino, S. Knappe, U.-I. Flügge, K. Fischer, The Plastidic Pentose Phosphate Translocator Represents a Link between the Cytosolic and the Plastidic Pentose Phosphate Pathways in Plants, *Plant Physiol* **128**, 512–522 (2002).

82. L. Lim, G. McFadden, The evolution, metabolism and functions of the apicoplast, *Philosophical Transactions Royal Soc B Biological Sci* **365**, 749–763 (2010).

83. K. Mullin *et al.*, Membrane transporters in the relict plastid of malaria parasites., *P Natl Acad Sci Usa* **103**, 9572–7 (2006).

84. C. Brooks *et al.*, The toxoplasma apicoplast phosphate translocator links cytosolic and apicoplast metabolism and is essential for parasite survival., *Cell Host Microbe* **7**, 62–73 (2009).

85. T. Banerjee, D. Jaijyan, N. Surolia, A. Singh, A. Surolia, Apicoplast triose phosphate transporter (TPT) gene knockout is lethal for Plasmodium, *Mol Biochem Parasit* **186**, 44–50 (2012).

86. L. Lim, M. Linka, K. Mullin, A. Weber, G. McFadden, The carbon and energy sources of the non-photosynthetic plastid in the malaria parasite, *Febs Lett* **584** (2009), doi:10.1016/j.febslet.2009.11.097.

87. U.-I. Flügge, Phosphate translocation in the regulation of photosynthesis, *Journal of Experimental Botany* , 1317–1323 (1995).

88. R. Gerhardt, M. Stitt, H. W. Heldt, Subcellular Metabolite Levels in Spinach Leaves : Regulation of Sucrose Synthesis during Diurnal Alterations in Photosynthetic Partitioning., *Plant Physiol.* **83**, 399–407 (1987).

89. U.-I. I. Flugge, PHOSPHATE TRANSLOCATORS IN PLASTIDS., *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**, 27–45 (1999).

90. Flügge *et al.*, The triose phosphate-3-phosphoglycerate-phosphate translocator from spinach chloroplasts: nucleotide sequence of a full-length cDNA clone and import of the in vitro synthesized precursor protein into chloroplasts., *Embo J* **8**, 39–46 (1989).

91. B. Bruce, Chloroplast transit peptides: structure, function and evolution, *Trends Cell Biol* **10**, 440–447 (2000).

92. V. Saudek, Cystinosin, MPDU1, SWEETs and KDELR Belong to a Well-Defined Protein Family with Putative Function of Cargo Receptors Involved in Vesicle Trafficking, *PLoS ONE* **7** (2012), doi:10.1371/journal.pone.0030876.

93. Y. Zhai, W. Heijne, D. Smith, M. Saier, Homologues of archaeal rhodopsins in plants, animals and fungi: structural and functional predications for a putative fungal chaperone protein, *Biochimica Et Biophysica Acta Bba - Biomembr* **1511**, 206–223 (2001).

94. M. Town *et al.*, A novel gene encoding an integral membrane protein is mutated in nephropathic cystinosis, *Nature Genetics* **18**, 319–324 (1998).

95. T. Kawate, E. Gouaux, Fluorescence-Detection Size-Exclusion Chromatography for Precrystallization Screening of Integral Membrane Proteins, *Structure* **14**, 673–681 (2006).

96. M. Hattori, R. Hibbs, E. Gouaux, A Fluorescence-Detection Size-Exclusion Chromatography-Based Thermostability Assay for Membrane Protein Precrystallization Screening, *Structure* **20** (2012), doi:10.1016/j.str.2012.06.009.

97. Backmark *et al.*, Fluorescent probe for high-throughput screening of membrane protein expression, *Protein Sci* **22**, 1124–1132 (2013).

98. M. Caffrey, V. Cherezov, Crystallizing membrane proteins using lipidic mesophases, *Nature Protocols* **4**, 706–731 (2009).

99. W. Kabsch, XDS, *Acta Crystallogr D Biol Crystallogr* **66** (2010), doi:10.1107/S0907444909047337.

100. T. R. Schneider, G. M. Sheldrick, Substructure solution with SHELXD, *Acta Crystallogr Sect D Biological Crystallogr* **58**, 1772–9 (2002).

101. D. Sylvie, C. Vonrhein, E. Blanc, P. Roversi, G. Bricogne, *Macromolecular Crystallography Protocols, Volume 2* (springer, 2006), pp. 215–230.

102. P. D. Adams *et al.*, PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr D Biol Crystallogr* **66** (2010), doi:10.1107/S0907444909052925.

103. T. C. Terwilliger *et al.*, Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard, *Acta Crystallogr D Biol Crystallogr* **64**, 61–69 (2007).

104. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot, *Acta Crystallogr D Biol Crystallogr* **66** (2010), doi:10.1107/S0907444910007493.

105. A. J. McCoy *et al.*, Phaser crystallographic software, *J Appl Crystallogr* **40** (2007), doi:10.1107/S0021889807021206.

106. D. Deng *et al.*, Crystal structure of the human glucose transporter GLUT1, *Nature* **510**, 121–125 (2014).

107. B. Schenk *et al.*, MPDU1 mutations underlie a novel human congenital disorder of glycosylation, designated type If., *J. Clin. Invest.* **108**, 1687–95 (2001).

108. B. Liu, H. Du, R. Rutkowski, A. Gartner, X. Wang, LAAT-1 Is the Lysosomal Lysine/Arginine Transporter That Maintains Amino Acid Homeostasis, *Science* **337**, 351–

354 (2012).

109. Y. Xu *et al.*, Structures of bacterial homologues of SWEET transporters in two distinct conformations, *Nature* **515**, 448–452 (2014).

110. J. Wang *et al.*, Crystal structure of a bacterial homologue of SWEET transporters., *Cell Res.* **24**, 1486–9 (2014).

111. H.-W. Heldt, Three decades in transport business: studies of metabolite transport in chloroplasts – a personal perspective, *Photosynth Res* **73**, 265–272 (2002).

112. K. Fischer, The Import and Export Business in Plastids: Transport Processes across the Inner Envelope Membrane, *Plant Physiol* **155**, 1511–1519 (2011).

113. U.-I. Flügge *et al.*, The major chloroplast envelope polypeptide is the phosphate translocator and not the protein import receptor, *Nature* **353**, 364–367 (1991).

114. U. I. Flügge, Hydrodynamic properties of the Triton X-100-solubilized chloroplast phosphate translocator, *Biochimica et Biophysica Acta (BBA) - Biomembranes* **815**, 299–305 (1985).

115. A. Gross, G. Brückner, H. Heldt, U. Flügge, Comparison of the kinetic properties, inhibition and labelling of the phosphate translocators from maize and spinach mesophyll chloroplasts, *Planta* **180** (1989), doi:10.1007/BF00194006.

116. H. Wallmeier, A. Weber, A. Gross, U.-I. Flügge, *Transport and Receptor Proteins of Plant Membranes* (springer, 1992), pp. 77–89.

117. K. Fischer *et al.*, Cloning and in vivo expression of functional triose phosphate/phosphate translocators from C3- and C4-plants: evidence for the putative participation of specific amino acid residues in the recognition of phosphoenolpyruvate., *Plant J.* **5**, 215–26 (1994).

118. D. Drew *et al.*, GFP-based optimization scheme for the overexpression and purification of eukaryotic membrane proteins in Saccharomyces cerevisiae, *Nat Protoc* **3**, nprot.2008.44 (2008).

119. P. Chae *et al.*, Maltose–neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins, *Nat Methods* **7**, nmeth.1526 (2010).

120. Y. Hayashi, H. Matsui, T. Takagi, *Membrane protein molecular weight determined by low-angle laser light-scattering photometry coupled with high-performance gel chromatography.* (1989), pp. 514–28.

121. J. Foadi *et al.*, Clustering procedures for the optimal selection of data sets from multiple crystals in macromolecular crystallography, *Acta Crystallogr Sect D Biological Crystallogr* **69**, 1617–1632 (2013).

122. M. Linka, A. Jamai, A. Weber, Functional Characterization of the Plastidic Phosphate Translocator Gene Family from the Thermo-Acidophilic Red Alga Galdieria sulphuraria Reveals Specific Adaptations of Primary Carbon Partitioning in Green Plants and Red Algae,

Plant Physiol **148**, 1487–1496 (2008).

123. M. Javanainen, Universal Method for Embedding Proteins into Complex Lipid Bilayers for Molecular Dynamics Simulations, *J Chem Theory Comput* **10**, 2577–2582 (2014).

124. J. Klauda *et al.*, Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types, *J Phys Chem B* **114**, 7830–7843 (2010).

125. B. Hess, C. Kutzner, D. Spoel, E. Lindahl, GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation, *J Chem Theory Comput* **4**, 435–447 (2008).

126. G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, *J Chem Phys* **126**, 014101 (2007).

127. S. Nosé, A unified formulation of the constant temperature molecular dynamics methods, *J Chem Phys* **81**, 511–519 (1984).

128. W. Hoover, Canonical dynamics: Equilibrium phase-space distributions, *Phys Rev A* **31**, 1695–1697 (1985).

129. Parrinello, Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, *J Appl Phys* **52**, 7182–7190 (1981).

130. B. Hess, P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation., *J Chem Theory Comput* **4**, 116–22 (2008).

131. T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems, *J Chem Phys* **98**, 10089–10092 (1993).

132. B. Kammerer *et al.*, Molecular characterization of a carbon transporter in plastids from heterotrophic tissues: the glucose 6-phosphate/phosphate antiporter., *Plant Cell* **10**, 105–17 (1998).

133. Loddenkötter, Kammerer, Fischer, Flügge, Expression of the functional mature chloroplast triose phosphate translocator in yeast internal membranes and purification of the histidine-tagged protein by a single metal-affinity chromatography step., *Proc National Acad Sci* **90**, 2155–2159 (1993).

134. S. KORE-EDA *et al.*, Characterization of the Plastidic Phosphate Translocators in the Inducible Crassulacean Acid Metabolism Plant Mesembryanthemum crystallinum, *Biosci Biotechnology Biochem* **77**, 1511–1516 (2014).

135. R. E. Häusler, N. H. Schlieben, U. I. Flügge, Control of carbon partitioning and photosynthesis by the triose phosphate/phosphate translocator in transgenic tobacco plants (Nicotiana tabacum). II. Assessment of control coefficients of the triose phosphate/phosphate translocator., *Planta* **210**, 383–90 (2000).

136. M. Matsuzaki *et al.*, Genome sequence of the ultrasmall unicellular red alga Cyanidioschyzon merolae 10D, *Nature* **428**, nature02398 (2004).

137. G. Schönknecht *et al.*, Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote., *Science* **339**, 1207–10 (2013).

138. M. Caffrey, Crystallizing membrane proteins for structure determination: use of lipidic mesophases., *Annu Rev Biophys* **38**, 29–51 (2009).

139. R. E. Martin, K. Kirk, The malaria parasite's chloroquine resistance transporter is a member of the drug/metabolite transporter superfamily., *Molecular biology and evolution* **21**, 1938–49 (2004).

140. H. Tsuchiya *et al.*, Structural basis for amino acid export by DMT superfamily transporter YddG., *Nature* **534**, 417–20 (2016).

141. D. L. Jack, N. M. Yang, M. H. Saier, The drug/metabolite transporter superfamily., *Eur. J. Biochem.* **268**, 3620–39 (2001).

142. E. Krissinel, K. Henrick, Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions., *Acta Crystallogr Sect D Biological Crystallogr* **60**, 2256–68 (2004).

143. S. E. Rollauer *et al.*, Structure of the TatC core of the twin-arginine protein transport system., *Nature* **492**, 210–4 (2012).

144. A. Nozawa *et al.*, A Cell-Free Translation and Proteoliposome Reconstitution System for Functional Analysis of Plant Solute Transporters, *Plant Cell Physiol* **48**, 1815–1820 (2007).

145. U. Flügge, Reaction mechanism and asymmetric orientation of the reconstituted chloroplast phosphate translocator., *Biochim. Biophys. Acta* **1110**, 112–8 (1992).

146. D. Drew, O. Boudker, Shared Molecular Mechanisms of Membrane Transporters., *Annu. Rev. Biochem.* **85**, 543–72 (2016).

147. S. Knappe, U.-I. Flügge, K. Fischer, Analysis of the Plastidic phosphate translocator Gene Family in Arabidopsis and Identification of New phosphate translocator-Homologous Transporters, Classified by Their Putative Substrate-Binding Site, *Plant Physiol* **131**, 1178–1190 (2003).

148. B. Hadley *et al.*, Structure and function of nucleotide sugar transporters: Current progress., *Comput Struct Biotechnol J* **10**, 23–32 (2014).

149. L. Zhang *et al.*, Overriding the co-limiting import of carbon and energy into tuber amyloplasts increases the starch content and yield of transgenic potato plants, *Plant Biotechnol J* **6**, 453–464 (2008).

150. M.-H. Cho, A. Jang, S. Bhoo, J.-S. Jeon, T.-R. Hahn, Manipulation of triose phosphate/phosphate translocator and cytosolic fructose-1,6-bisphosphatase, the key components in photosynthetic sucrose synthesis, enhances the source capacity of transgenic Arabidopsis plants, *Photosynth Res* **111**, 261–268 (2012).

151. C. Rautengarten *et al.*, The Golgi localized bifunctional UDP-rhamnose/UDP-galactose

transporter family of Arabidopsis., *Proc. Natl. Acad. Sci. U.S.A.* **111**, 11563–8 (2014).

152. F. Sievers *et al.*, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, *Mol Syst Biol* **7**, 539–539 (2011).

153. Y. Tao *et al.*, Structure of a eukaryotic SWEET transporter in a homotrimeric complex, *Nature* **527** (2015), doi:10.1038/nature15391.

154. L. Han *et al.*, Molecular mechanism of substrate recognition and transport by the AtSWEET13 sugar transporter., *Proc. Natl. Acad. Sci. U.S.A.* **114**, 10089–10094 (2017).

155. N. Latorraca *et al.*, Mechanism of Substrate Translocation in an Alternating Access Transporter, *Cell* **169**, 96–107.e12 (2017).

156. J. Parker, S. Newstead, Structural basis of nucleotide sugar transport across the Golgi membrane, *Nature* , nature24464 (2017).

157. X.-D. Gao, N. Dean, Distinct Protein Domains of the Yeast Golgi GDP-mannose Transporter Mediate Oligomer Assembly and Export from the Endoplasmic Reticulum, *J Biol Chem* **275**, 17718–17727 (2000).

158. Handford, Rodriguez-Furlán, Orellana, Nucleotide-sugar transporters: structure, function and roles in vivo, *Braz J Med Biol Res* **39**, 1149–1158 (2006).

159. G. Meer, Lipids of the Golgi membrane, *Trends Cell Biol* **8**, 29–33 (1998).

160. H. Sharpe, T. Stevens, S. Munro, A Comprehensive Comparison of Transmembrane Domains Reveals Organelle-Specific Properties, *Cell* **142**, 158–169 (2010).

161. D. Banfield, Mechanisms of Protein Retention in the Golgi, *Csh Perspect Biol* **3**, a005264 (2011).

162. M. Wong, S. Munro, The specificity of vesicle traffic to the Golgi is encoded in the golgin coiled-coil proteins, *Science* **346**, 1256898 (2014).

163. M. Lomize, I. Pogozheva, H. Joo, H. Mosberg, A. Lomize, OPM database and PPM web server: resources for positioning of proteins in membranes, *Nucleic Acids Research* **40**, D370–D376 (2012).

164. D. Deng *et al.*, Molecular basis of ligand recognition and transport by glucose transporters, *Nature* (2015), doi:10.1038/nature14655.

165. R. O. Mackender, R. M. Leech, The Galactolipid, Phospholipid, and Fatty Acid Composition of the Chloroplast Envelope Membranes of Vicia faba. L., *Plant Physiol.* **53**, 496–502 (1974).

166. M. A. Block, R. Douce, J. Joyard, N. Rolland, Chloroplast envelope membranes: a dynamic interface between plastids and the cytosol., *Photosyn. Res.* **92**, 225–44 (2007).

167. R. O. Mackender, R. M. Leech, Isolation of chloroplast envelope membranes., *Nature* **228**, 1347–9 (1970).

168. Flugge, Heldt, Metabolite Translocators of the Chloroplast Envelope, *Annu Rev Plant Phys*

42, 129–144 (1991).

169. J. Soll, E. Schleiff, Protein import into chloroplasts., *Nature reviews. Molecular cell biology* **5**, 198–208 (2004).

170. P. Jarvis, C. Robinson, Mechanisms of protein import and routing in chloroplasts., *Curr. Biol.* **14**, R1064–77 (2004).

171. J. Lee, D. Kim, I. Hwang, Specific targeting of proteins to outer envelope membranes of endosymbiotic organelles, chloroplasts, and mitochondria, *Frontiers Plant Sci* **5**, 173 (2014).

172. B. Bölter, J. Soll, Once upon a Time - Chloroplast Protein Import Research from Infancy to Future Challenges., *Mol Plant* **9**, 798–812 (2016).

173. J. Janouškovec, A. Horák, M. Oborník, J. Lukeš, P. Keeling, A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids, *Proc National Acad Sci* **107**, 10949–10954 (2010).

174. A. Weber, M. Linka, D. Bhattacharya, Single, Ancient Origin of a Plastid Metabolite Translocator Family in Plantae from an Endomembrane-Derived Ancestor, *Eukaryot Cell* **5**, 609–612 (2006).

175. C. X. Chan, A. Reyes-Prieto, D. Bhattacharya, Red and green algal origin of diatom membrane transporters: insights into environmental adaptation and cell evolution., *PLoS ONE* **6**, e29138 (2011).

176. H. M. Tyra, M. Linka, A. P. Weber, D. Bhattacharya, Host origin of plastid solute transporters in the first photosynthetic eukaryotes., *Genome Biol.* **8**, R212 (2007).

177. B.-J. Nature, Phosphorylation of chloroplast membrane polypeptides, *Nature* (1977) (available at https://www.nature.com/nature/journal/v269/n5626/abs/269344a0.html).

178. P. Dörmann, C. Benning, Galactolipids rule in seed plants., *Trends Plant Sci* **7**, 112–8 (2002).

179. Y.-S. Law *et al.*, Phosphorylation and Dephosphorylation of the Presequence of Precursor MULTIPLE ORGANELLAR RNA EDITING FACTOR3 during Import into Mitochondria from Arabidopsis, *Plant Physiol.* **169**, 1344–1355 (2015).

180. M. Balsera, J. Soll, B. Buchanan, Redox extends its regulatory reach to chloroplast protein import, *Trends Plant Sci* **15**, 515–521 (2010).

181. A. Stengel, J. P. Benz, J. Soll, B. Bölter, Redox-regulation of protein import into chloroplasts and mitochondria: similarities and differences., *Plant Signal Behav* **5**, 105–9 (2010).

182. Y. Yang, A. Zienkiewicz, A. Lavell, C. Benning, Co-evolution of Domain Interactions in the Chloroplast TGD1, 2, 3 Lipid Transfer Complex Specific to Brassicaceae and Poaceae Plants, *The Plant Cell* , tpc.00182.2017 (2017).

183. Z. Wang, C. Benning, Chloroplast lipid synthesis and lipid trafficking through ER–plastid

membrane contact sites, *Biochem Soc T* **40**, 457–463 (2012).

184. Z. Wang, N. S. Anderson, C. Benning, The phosphatidic acid binding site of the Arabidopsis trigalactosyldiacylglycerol 4 (TGD4) protein required for lipid import into chloroplasts., *J. Biol. Chem.* **288**, 4763–71 (2013).

185. J. Fan, Z. Zhai, C. Yan, C. Xu, Arabidopsis TRIGALACTOSYLDIACYLGLYCEROL5 Interacts with TGD1, TGD2, and TGD4 to Facilitate Lipid Transfer from the Endoplasmic Reticulum to Plastids, *Plant Cell* **27**, tpc.15.00394 (2015).

186. N. Li, C. Xu, Y. Li-Beisson, K. Philippar, Fatty Acid and Lipid Transport in Plant Cells, *Trends Plant Sci* **21**, 145–158 (2016).

187. P. Lundquist, O. Mantegazza, A. Stefanski, K. Stühler, A. Weber, Surveying the Oligomeric State of Arabidopsis thaliana Chloroplasts, *Molecular Plant* **10**, 197–211 (2017).

Original papers

Lee, Y., Nishizawa, T., Yamashita, K., Ishitani, R. & Nureki, O. (2015). Structural basis for the facilitative diffusion mechanism by SemiSWEET transporter. *Nature communications* **6**, 6112.

Lee, Y., Nishizawa, T., Takemoto, M., Kumazaki, K., Yamashita, K., Hirata, K., Minoda, A., Nagatoishi, S., Tsumoto, K., Ishitani, R. & Nureki. O. (2017). Structure of the triose-phosphate/ phosphate translocator reveals the basis of substrate specificity. *Nature Plants* **3**, 825–832.