Biochemistry and Chemical Biology
Structural Biology and Molecular Biophysics

# Tracing the substrate translocation mechanism in P-glycoprotein

**Theresa Gewering, Deepali Waghray, Kristian Parey, Hendrik Jung, Nghi N.B. Tran, Joel Zapata, Pengyi Zhao, Hao Chen, Dovile Januliene, Gerhard Hummer, Ina L. Urbatsch, Arne Moeller, Qinghai Zhang**

Osnabrück University, Department of Biology/Chemistry, Structural Biology section, 49076 Osnabrück, Germany • Department of Structural Biology, Max Planck Institute of Biophysics, 60438 Frankfurt am Main, Germany • Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, USA • Osnabrück University, Center of Cellular Nanoanalytic Osnabrück (CellNanOs), 49076 Osnabrück, Germany • Department of Theoretical Biophysics, Max Planck Institute of Biophysics, 60438 Frankfurt am Main, Germany • Department of Cell Biology and Biochemistry, Texas Tech University Health Sciences Center, Lubbock, TX 79430, USA • Department of Chemistry & Environmental Science, New Jersey Institute of Technology, Newark, New Jersey 07102, USA • Institute for Biophysics, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany



**Reviewed Preprint**

Published from the original preprint after peer review and assessment by eLife.

About eLife's process

**Reviewed preprint version 2**
January 8, 2024

**Reviewed preprint version 1**
October 5, 2023 (this version)

**Sent for peer review**
July 2, 2023

**Posted to preprint server**
June 18, 2023

## Abstract

P-glycoprotein (Pgp) is a prototypical ABC transporter of great biological and clinical significance that confers cancer multidrug resistance and mediates the bioavailability and pharmacokinetics of many drugs[1–3]. Decades of structural and biochemical studies have provided insights into how Pgp binds diverse compounds[4–9], but how they are translocated through the membrane has remained elusive. Here, we covalently attached a cyclic substrate to discrete sites of Pgp and determined multiple complex structures in inward- and outward-facing states by cryoEM. In conjunction with molecular dynamics simulations, our structures trace the substrate passage across the membrane and identify conformational changes in transmembrane helix 1 (TM1) as regulators of substrate transport. In mid-transport conformations, TM1 breaks at glycine 72. Mutation of this residue significantly impairs drug transport of Pgp in vivo, corroborating the importance of its regulatory role. Importantly, our data suggest that the cyclic substrate can exit Pgp without the requirement of a wide-open outward-facing conformation, diverting from the common efflux model for Pgp and other ABC exporters. The substrate transport mechanism of Pgp revealed here pinpoints critical targets for future drug discovery studies of this medically relevant system.

### eLife assessment

P-glycoprotein (Pgp) is a major ABC-transporter that couples ATP hydrolysis with the export of xenobiotics and drugs, influencing their pharmacokinetics. The authors provide **convincing** cryo-EM Pgp structures of drug complexes in previously unforeseen outward-facing conformations. These structures reveal **important** mechanistic insights that could be of broader impact, but because these transient-like states were captured by cysteine cross-linking the substrate to Pgp, support for the conclusions is **incomplete**, and further validation is required.

# Main Text

P-glycoprotein (Pgp, also known as MDR-1 and ABCB1) is a prominent member of the human ATP-binding cassette (ABC) efflux transporters that removes a large variety of chemically unrelated hydrophobic compounds from cells[3, 9]. Pgp is highly expressed in the liver, kidney, and intestines, where it limits absorption and enhances the excretion of cytotoxic compounds[10–12]. In the blood-brain, blood-testis, and blood-placenta barriers, Pgp protects and detoxifies those sanctuaries from xenobiotics[13–15]. Of central importance for many chemotherapeutic treatments, Pgp is a major determinant of drug bioavailability and pharmacokinetics and confers multidrug resistance in several diseases, most notably cancer[16–18]. Consequently, evasion, selective inhibition, and modulation of Pgp transport are important goals in drug development, but are hindered by a lack of detailed understanding of the drug transport mechanisms[19].

For Pgp and ABC transporters in general, ATP binding, hydrolysis, and subsequent release of the cleaved products (inorganic phosphate and ADP) fuel large-scale conformational changes that ultimately result in translocation of substrates across the membrane bilayer[20, 21]. In inward-facing (IF) conformations, the nucleotide-binding domains (NBDs) are separated, and substrates can access a large binding cavity that is open to the lower membrane leaflet and cytoplasm. ATP-induced dimerization of the NBDs promotes large rearrangements of the transmembrane domains (TMDs) from an IF to an outward-facing (OF) conformation, from which the substrate is released[22, 23]. Pgp structures with bound substrates and inhibitors in IF conformations have revealed one or more overlapping binding sites[6–8, 24, 25]. Multimodal binding mechanisms for chemically related compounds have been proposed[6, 9], including access of hydrophobic ligands from the lipid-protein interface to the binding sites of Pgp[8, 24]. However, structural data that elucidate how substrates are shuttled across the lipid bilayer to the extracellular space are unavailable at present. So far, only a single OF structure of mammalian Pgp has been described, but without detectable substrate density[26]. Most insights into the transport pathway arise from molecular dynamics simulations, which are also limited by the scarcity of available structural details[27, 28]. Consequently, the molecular mechanics that move the substrate through the transmembrane passage have remained elusive. To address this fundamental question and overcome the dynamic and transient nature of substrate translocation, we tethered a substrate molecule covalently to specific residues along the putative translocation pathway of Pgp and utilized cryoEM to capture Pgp structures with bound substrate during various stages of transport.

## Covalent ligand design for Pgp labeling

For covalent labeling of Pgp, we synthesized a derivative of the previously published QZ-Ala tripeptide substrate[24], by substituting one of the three Ala with Cys that is disulfide-linked to a 2,4-dinitrophenyl thiolate group (designated AAC-DNPT, See **Methods**, **Extended Data Fig. 1**). QZ-Ala is a strong ATPase stimulator and subject to Pgp transport as evidenced by its reduced cytotoxicity in cells overexpressing Pgp[24]. Additionally, we have measured Pgp-mediated transport of QZ-Ala in the MDCK-MDR1 monolayer permeability assay (see **Methods**). The efflux ratio ($R_E$) of QZ-Ala was determined as 2.5. In the presence of the Pgp inhibitor cyclosporin A, $R_E$ was 0.9, indicating that the Pgp-specific transport of QZ-Ala was inhibited. Derivatizing QZ-Ala is simplified by its structural symmetry, which reduces the complexity of modifying this transport substrate for attaching to Pgp in different states. Furthermore, attachment of DNPT strongly activates disulfide exchange, allowing rapid and efficient crosslinking between the cyclic peptide and accessible cysteines. Free DNPT, released during crosslinking, is bright yellow. This provides a simple, visible readout that can be monitored with a UV-Vis spectrometer (**Fig. 1a**). Four single-Cys mutants of Pgp (Pgp335, Pgp978, Pgp971, and Pgp302) were generated for crosslinking, and mutations were chosen near the two previously reported QZ-Ala binding sites[24]. L335C in

transmembrane helix 6 (TM6) and V978C in TM12 are situated at symmetric positions on opposing TMD halves, while L971C is located two helical turns above V978 (**Fig. 1b**). Residues L335C, V978C as well as I302C located in TM5 were previously shown to be labeled with verapamil[29]. Based on the appearance of the yellow DNPT byproduct, Pgp335 and Pgp978 reacted with AAC-DNPT within minutes, while Pgp971 and Pgp302 did not. Even after extended incubations only background absorbances were detected comparably to Cys-less Pgp (CL-Pgp) (**Fig. 1c**). Only after adding $Mg^{2+}$ATP to fuel substrate translocation, Pgp971 could react with AAC-DNPT, but Pgp302 could not. Covalent attachment of the cyclic peptide AAC was further validated by high-resolution mass spectrometry of trypsin-digested peptide fragments (**Extended Data Figs. 2** to **4**). The presence of noncovalent transport substrates such as QZ-Ala or verapamil increase ATP hydrolysis in CL-Pgp and the single Cys mutants (**Extended Data Fig. 5a**). Similar to QZ-Ala, AAC-DNTP stimulated ATPase activity of CL-Pgp at low concentrations; with half-maximal enhancement of activity (EC50) seen at submicromolar concentrations, suggesting that it also acts as substrate (**Extended Data Fig. 5b**). At higher concentrations, the rate of ATP hydrolysis is significantly reduced in the mutants compared to CL-Pgp indicative of cross-linking AAC-DNPT with the single Cys mutants (**Fig. 1d** and **Extended Data Fig. 5b**). Importantly, Pgp335 and Pgp978 retain notably higher ATPase activity after cross-linking than the respective apo-Pgp mutants in the absence of substrate (**Extended Data Fig. 5b** and **Fig. 1d**), suggesting that tethered AAC still acts as a substrate and that Pgp335 and Pgp978 cycle between IF and OF conformations.

## Pgp structures with bound substrates

To capture different conformations during the transport cycle of Pgp, we mutated the catalytic glutamate residues in the Walker-B motifs of each NBD (E552Q/E1197Q) in Pgp335 and Pgp978. When $Mg^{2+}$ATP is present, these mutations arrest the enzyme in an ATP-occluded, NBD-dimerized conformation that resembles the pre-hydrolysis state[26]. For Pgp971, we prepared the covalent complex in the presence of $Mg^{2+}$ ATP and stabilized the resulting OF complex with vanadate in a post-hydrolysis intermediate state[30, 31] (**Fig. 1c**). CryoEM analysis of the respective covalent complexes in both OF and IF conformations provided eight high-resolution structures with zero, one, or two ligands bound. In addition, a control dataset with ATP, but without the substrate was collected for Pgp335 (**Fig. 2**, **Extended Data Fig. 6** and **7**, and **Extended Data Table 1**).

The IF-structure of Pgp335, solved at 3.8 Å resolution, shows clear densities for two substrate molecules: one crosslinked and one not (IF335-2lig, **Fig. 2** and **Extended Data Fig. 7** and **8**). The non-crosslinked molecule is bound in a position that resembles the one previously reported in a crystal structure[24]. It interacts closely with the covalently-linked substrate, which is situated further up the membrane. Both ligand molecules are oriented along the vertical axis of the protein and positioned side-by-side through hydrophobic interactions. The uncleaved DNPT group of the non-crosslinked substrate is evident in the cryoEM density map. Two substrate molecules are also bound in the IF structure of Pgp978 (IF978-2lig); the non-covalently bound molecule is clearly visible at a similar position as in IF335-2lig, while the crosslinked substrate is located on the opposite side of the binding cavity near TM12 (**Extended Data Fig. 8**). Corroborating our labeling data (**Fig. 1c**), in the IF structure of Pgp971 (IF971-1lig) at 4.3 Å resolution, only the non-crosslinked substrate is bound at the approximate location seen for non-crosslinked substrate in IF335-2lig and IF978-2lig (**Fig. 2**, **Extended Data Figs. 8** and **9e**).

Multi-model cryoEM analyses of ATP-bound Pgp335 revealed three different OF conformations within a single dataset: ligand-free (42%), single-ligand-bound (34%), and double-ligand-bound (26%) (OF335-nolig, OF335-1lig and OF335-2lig), at 2.6 Å, 2.6 Å, and 3.1 Å resolution, respectively (**Extended Data Fig. 10**, **Fig. 2**). For Pgp978 and Pgp971, only OF conformations with a single bound ligand (OF978-1lig and OF971-1lig, **Fig. 2**) were detected at 2.9 Å and 3 Å resolution, respectively.

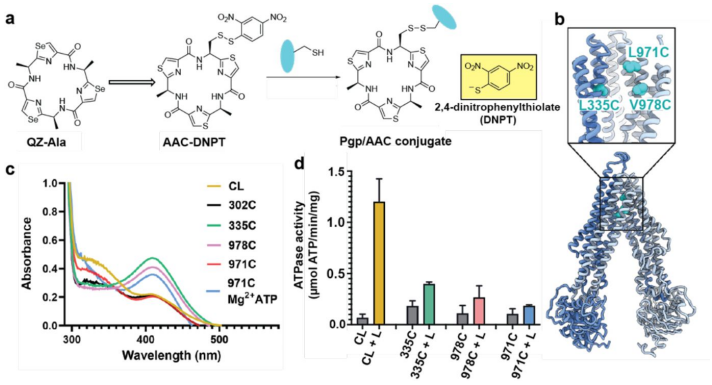



**Fig. 1.**

### Covalent ligand design for Pgp transport studies.

**(a)** Design of AAC-DNPT, a derivative of the cyclic peptide QZ-Ala, for disulfide crosslinking with single-cysteine mutants of Pgp; yellow product DNPT used as readout. **(b)** Positions of L335 (TM6), V978 and L971 (TM12) in Pgp, residues targeted for cyclic peptide labeling. **(c)** Monitoring covalent reaction between AAC-DNPT and cysteine mutants (+/-$Mg^{2+}$ATP) by UV-visible spectrometry; CL-Pgp and I302C were included as negative controls. **(d)** ATPase activities of L335C, V978C, and L971C before and after crosslinking with AAC-DNPT; CL-Pgp was used as a control. Data are averaged from triplicate measurements with standard error bars.

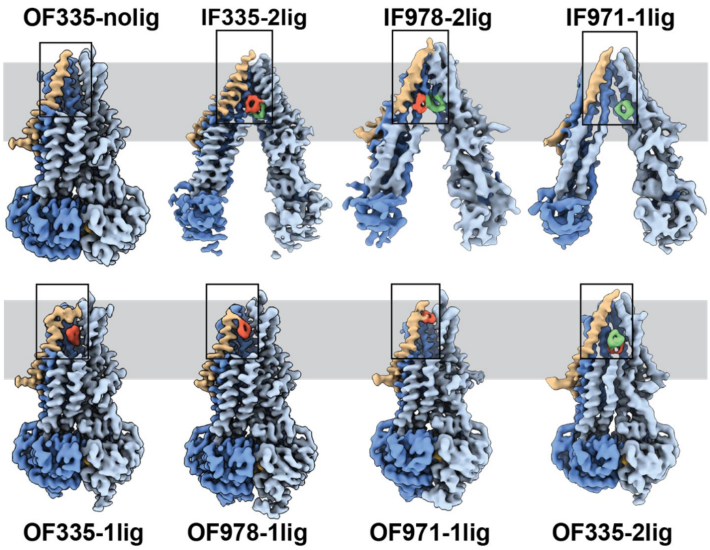



**Fig. 2.**

### High-resolution structures of Pgp with bound ligands.

For all mutants, IF and OF structures were obtained, with ligands revealed in the transmembrane region (grey shadow). For Pgp335 three different OF conformations were detected in one data set (OF335-nolig, OF335-1lig, OF335-2lig). To display the binding sites, the upper halves of TM 2,10,11 and 12 were removed from the densities. For the IF structures, TM9 was also removed for clarity. TM1 is shown in brown, with the covalently bound ligand in red and the non-crosslinked in green. Significant conformational changes were observed among these structures from which a further zoom-in of TM1 (black box region) was shown in **Fig. 4** and **Extended Data Fig. 8**.

## Substrate translocation pathway

In the OF structures with a single bound substrate, the central transmembrane helices TM1, TM6, TM7 and TM12 bulge outwards to accommodate the substrate in a small tunnel (**Extended Data Fig. 11**). Comparing the ligand positions between the IF and OF structures, in IF335-2lig to OF335-1lig, and IF978-2lig to OF978-1lig, the covalently attached ligand pivots around the respective Cys residue by almost 180°, moving it further up the translocation tunnel (**Extended Data Figs. 8** and **9a,b**, and Supplementary Videos 1 and 2). Our data do not reveal whether rotation of the ligand during the IF-OF conformational change is required for transport or if this reflects the fact that the ligand is covalently pinned to the crosslinking residue. However, the observation of ligand rotation illustrates how much space is available in the binding cavity during this conformational transition. The ligand position and orientation in OF335-1lig and OF978-1lig overall coincide well, with the latter shifting up slightly (**Extended Data Fig. 9d**). This finding indicates that the export tunnel is inherent to substrate translocation regardless of the initial binding position in the IF conformation. As such, our data suggest a single common pathway for substrate export in Pgp, contrary to the dual pseudosymmetric pathways suggested earlier[28, 32].

OF971-1lig captures a later-stage transport intermediate, in which the compound resides approximately 7 Å further up the translocation tunnel and is coordinated by highly conserved residues, including M74 and F78 (**Figs. 2** and **3b**, **Extended Data Figs. 8** and **12**). As no crosslinked substrate was detected in the corresponding IF structure, the substrate was likely transported to residue L971C before the crosslinking reaction could take place. Substrate entry from the extracellular side is less likely given the small opening of the substrate translocation pathway in all OF Pgp structures with or without bound substrate obtained here and previously[26]. In all substrate bound OF conformations methionines and phenylalanines at different positions seem to be important for ligand coordination in all of the Pgp structures (**Extended Data Fig. 12**). As expected, the substrate is held in the binding pocket mainly by the hydrophobic interactions throughout the different conformations (**Extended Data Figs. 8** and **12**). In the OF335-nolig structure, as well as in the virtually indistinguishable control, the translocation tunnel is collapsed, which shields the substrate binding pocket from the extracellular milieu and prevents re-entry from this side[26].

## Regulatory role of TM1

The series of IF and OF Pgp structures reveal a cascade of conformational changes in TM1, which displays significant plasticity throughout the different transport stages (**Fig. 2**, **Extended Data Fig. 8**, and Supplementary Videos 1, 2 and 3). In IF conformations, TM1 is a long, straight helix (**Fig. 2** and **Extended Data Fig. 8**). However, in OF335-1lig, OF978-1lig, and OF971-1lig, TM1 swings out at A63, dilating the transmembrane passage and providing sufficient space for accommodating the bound ligand. At the same time, a pronounced kink at G72 on TM1 destabilizes the helix and the extracellular loop of Pgp between TM1 and TM2 and leads to a partial unwinding. This conformational deformation shields the intramembranous tunnel from the extracellular environment, with TM1 acting as a lid, almost parallel to the membrane (**Fig. 2** and **4a**, **Extended Data Fig. 8**). Strikingly, along with the upward movement of the ligand (from L335C to V978C to L971C), the TM1 loop is progressively lifted and restabilized. Ultimately, as shown previously[26] and by our OF335-nolig structure, the helix is straightened upon the release of the ligand (**Fig. 4a**, **Extended Data Fig. 8**). Interestingly, the extracellular gate is sealed in all OF structures, which is especially surprising for OF971-1lig where the ligand is near the tunnel exit (**Fig. Extended Data Fig. 11**).

Using OF978-1lig as initial configuration, we performed molecular dynamics (MD) simulations of Pgp embedded in a lipid bilayer to validate the proposed substrate escape pathway. Accelerated by temperature (T=400 K), we observed movement and release of the non-covalently bound substrate

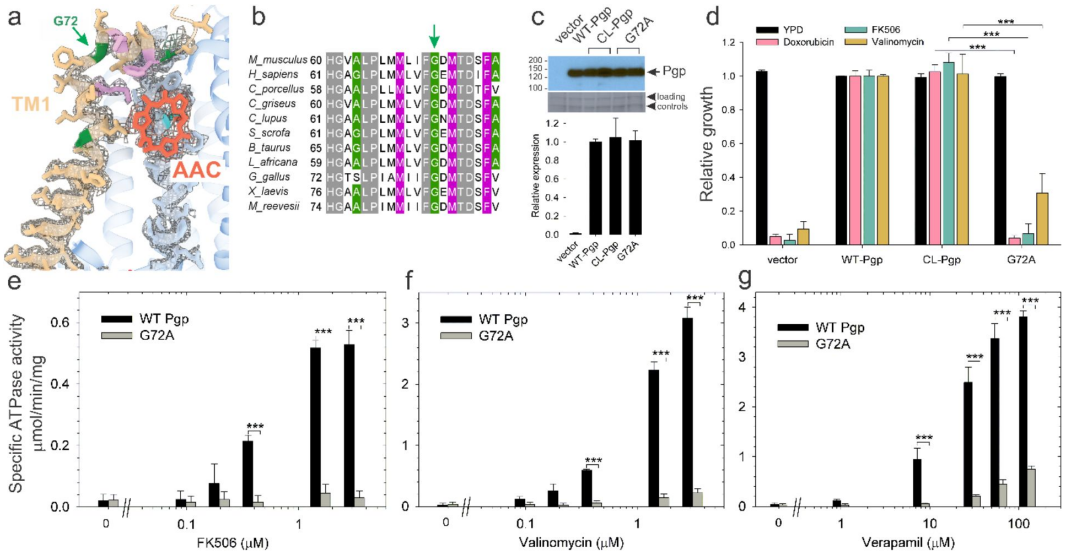



**Fig. 3.**

### Functional characterization of G72A mutant.

**(a)** TM1 breaks at the G72 position in the mid-transport states of Pgp. Shown is a close-up view of TM1 (brown) with side chain densities, the G72 position (green arrow), and the adjacently bound cyclic peptide AAC (red sticks) in the OF335-1lig structure. **(b)** Sequence alignment of TM1 in vertebrate Pgps including the highly conserved G72 residue (green arrow). **(c)** Level of expression of the G72A mutant relative to the WT- and CL-Pgp in *S. cerevisiae* cultures derived from the western blot analysis. **(d)** In vivo activity of the G72A mutant. The growth resistance of *S. cerevisiae* cultures transformed with either vector control plasmid, WT-Pgp, CL-Pgp, or the G72A mutant against several fungicides (doxorubicin, FK506, and valinomycin) relative to the YPD medium only (black bar) were analyzed. Compared to WT-Pgp and CL-Pgp, the G72A mutant severely compromised growth resistance against the tested drugs (p-values of <0.001 are indicated by ***). **(e-g)** ATPase activity of the purified G72A mutant. While basal ATP hydrolysis rates (in the absence or at low concentrations of drug) were indistinguishable between WT and G72A mutant proteins, stimulation of the ATPase activities by 1-3 µM FK506, 1-3 µM valinomycin, or 30-130 µM verapamil were severely impaired (p-values of <0.001 are indicated by ***).

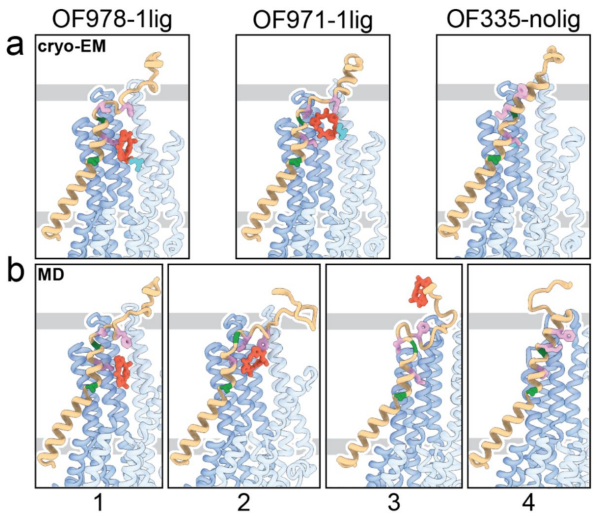



**Fig. 4.**

**Tracing the substrate movement in Pgp.**

Structures of OF Pgp from cryoEM **(a)** aligned with four snapshots from MD simulations **(b)** of substrate transport and release. TM1 is rendered in brown, with A63 and G72 in green, and M68, M74, and F78 in plum. The cyclic peptide is shown as red sticks. Significant conformational changes were observed for TM1 along with the upward movement and release of the substrate from OF Pgp.

to the extracellular side moderated by interactions between the ligand and hydrophobic residues in the upper part of TM1 (Supplementary Video 3). The intermediate steps coincided well with the conformations observed in cryoEM (**Fig. 4**, **Extended Data Figs. 13** and **14**). Without any bias imposed, the substrate progressed to a position closely matching OF971-1lig (**Fig. 2**, **4b**). Then, while TM1 gradually refolded from the intracellular side, the substrate wiggled out to form transient interactions with the still unstructured extracellular part of TM1 (**Fig. 4**). Importantly, the substrate escaped without Pgp experiencing a wide opening of the extracellular region. After the substrate left the exit tunnel, TM1 continued to refold to attain a straight helix configuration like OF335-nolig at the end of the simulation (**Fig. 4**). Overall, with exception of TM1, motions in Pgp to accommodate the moving substrate were mostly limited to small dilations of the transmembrane helices, supporting our observations from cryoEM.

Our cryoEM structures and MD simulations revealed an important regulatory role of TM1 during substrate transport, emphasizing the helix-breaking G72. Therefore, we mutated this residue, which is highly conserved in mammalian Pgp (**Fig. 3**), to a helix-stabilizing alanine[33] and tested the mutant's ability to export structurally diverse fungicides and to convey drug resistance *in vivo*. Compared to the wild-type and CL-Pgp, the G72A mutant expressed normally in *Saccharomyces cerevisiae* (**Fig. 3c**), but showed significantly reduced growth resistance to the test drugs including doxorubicin, FK506, and valinomycin (**Fig. 3d**). As drug transport was severely impaired, we further investigated the mutant's ATPase activity. In presence of transport substrates such as FK506, valinomycin, or verapamil the ATPase activity of WT Pgp was significantly stimulated; however, the G72A mutant displayed only low levels of basal ATPase activity (in the absence of drug) and the response to drug stimulation was greatly diminished over a range of concentrations tested (**Fig. 3e-g**). To reveal potential structural consequences responsible for the loss of function, we performed cryoEM on the Pgp335 G72A mutant after crosslinking the single cysteine L335C with AAC-DNTP and stabilizing the OF complex with $Mg^{2+}$ ATP and vanadate in a post-hydrolysis intermediate state. After multi-model classification, two conformations emerged. The first conformation, resolved at 3.0 Å, shows no substrate bound in analogy to OF335-nolig. The second dominant conformation is, however, much more heterogenous, resulting in a reduced and anisotropic resolution of approximately 4.6 Å (**Extended Data Fig. 15**). Here, the NBDs are well resolved but several transmembrane helices are missing, including TM1 in which residue G72A is located. The structural disorder, observed in the transmembrane region, is likely a result of the tethered ligand, causing local distortion of the surrounding α-helices as TM1 may not be able to kink at G72 in the mutant, which could explain its impaired transport function. Collectively, the significant differences in drug resistance profiles and drug-stimulated ATP hydrolysis rates between the G72A mutant and WT-Pgp along with our structural findings substantiate and highlight the key role of TM1 for the regulation of substrate transport in Pgp.

## Discussion

Our data suggest a mechanism for how Pgp mediates substrate transport. Transport commences in the IF conformation with binding of the substrate. Upon ATP-induced NBD-dimerization, the transporter converges into an OF conformation in which the intracellular gate is closed. The presence of the substrate in the OF conformation of Pgp dislocates TM1 at A63 and introduces a sharp kink at G72, forcing the helix to unwind partially between TM1 and TM2 (**Fig. 4** and **Extended Data Fig. 8**). This enlarges the extracellular loop which shields the transmembrane tunnel from the extracellular environment and guides the substrate to move up the translocation tunnel, allowing TM1 to gradually restabilize. During the last stages of transport (OF971-1lig to OF335-nolig), the TM1-TM2 regulatory loop straightens, and the substrate-interacting hydrophobic residues (M74 and F78) rotate and pivot away from the tunnel, which facilitates the escape of the substrate from the transporter. Although OF971-1lig captured a post-hydrolysis state, we cannot rule out the possibility of ligand release prior to ATP hydrolysis, when it is not covalently bound.

Notwithstanding, the proposed TM1 kinking and straightening mechanism of substrate expulsion would obviate the requirement for the Y-shaped, wide-open OF conformation reported for multiple bacterial ABC transporters[34, 35], suggesting a substrate-dependent transport mechanism in Pgp, or else a divergence of mechanisms among ABC transporters.

For the OF335-2lig structure, observed in the same data set as OF335-1lig and OF335-nolig, we hypothesize that the hydrophobic nature of the substrate promotes the binding of two substrate molecules to each other. MD simulations with this structure as a starting conformation reveal no separation of the two molecules, suggesting that they might be treated as one bigger substrate in the transport event. With the binding positions differing between OF335-1lig and OF335-2lig (**Extended Data Fig. 9c**, **12** and **16**), neither the cryoEM structure nor the simulations allow conclusions about a possible transport mechanism of two ligands at this point.

How ABC transporters move substrates through the lipid bilayer has been a long-standing question[36]. The scarcity of high-resolution structures of substrate-bound OF conformations[37], especially in the highly transient intermediate states, has severely limited our understanding of this vital process. Here, we were able to determine eight high-resolution structures of the archetypical ABC transporter Pgp, revealing the underlying molecular mechanics of substrate translocation. In general, the large number and diversity of ABC transporters, and the even wider variety of their substrates, make it unlikely that a universal substrate translocation mechanism will emerge. Our strategy of stalling the otherwise highly transient intermediates in this process through crosslinking and subsequent high-resolution cryoEM analysis could serve as a blueprint to understand the transport mechanisms of lipid transporters as well as other members of the transporter family.

# Methods

## Synthesis of compounds and analysis

The synthesis route and compound numbers are shown in **Extended Data Fig. 1**. All organic reactions were carried out under anhydrous conditions and argon atmosphere, unless otherwise noted. Reagents were purchased at the highest commercial quality and used without further purification. NMR spectra were acquired using Bruker DRX-300 or Bruker AV-600 instruments, and chemical shifts (δ) are reported in parts per million (ppm), which were calibrated using residual undeuterated solvent as an internal reference. High-resolution mass spectra (HRMS) were recorded on an Agilent Technologies 6230 TOF LC/MS with a Dual AJS ESI ion source. For flash column chromatography, the stationary phase was 60Å silica gel.

## General procedure for the synthesis of *N*-Boc-(*S*)-amino thioamides 1 and 2

To a solution of *N*-Boc-(*S*)-amino acid (1.0 g, 5.20 mmol) in THF (20 mL) at 0 °C was added triethylamine (0.8 mL, 5.80 mmol) and ethyl chloroformate (1.1 mL,10.92 mmol). The reaction mixture was stirred for 30 min before the addition of aqueous ammonium hydroxide (20 mL). The reaction mixture was stirred for a further 10 min. The mixture was extracted with ethyl acetate (EtOAc), and the organic layer was washed with brine and dried over anhydrous $Na_2SO_4$, then concentrated under reduced pressure to give *N*-Boc-(*S*)-amino amide as a white solid. The residue was redissolved in THF, Lawesson's reagent (3.0 g, 5.20 mmol) was added, and the reaction mixture was stirred at 50 °C for 12 h. After solvent removal, the residue was redissolved in EtOAc. The organic layer was washed with 1% NaOH, $H_2O$, and brine, then dried over anhydrous $Na_2SO_4$ and concentrated under reduced pressure. The resulting residue was purified by silica gel column chromatography using EtOAc/hexane (1:1) to give *N*-Boc-(*S*)-amino thioamide **1** or **2** as a yellow solid.

## Synthesis of thiazole ester 3

A solution of thioamide **1** (2.0 g, 9.80 mmol) dissolved in dimethoxyethane (DME, 20 mL) was cooled to –20 °C, followed by addition of $KHCO_3$ (7.84 g, 78.4 mmol) under inert atmosphere. The suspension was stirred for 15 min, followed by addition of ethyl bromopyruvate (3.6 mL, 29.4 mmol). The reaction mixture was further stirred for 1 h while it warmed from –20 °C to room temperature (rt). The reaction mixture was cooled to –20 °C again, and a solution of trifluoroacetic anhydride (5.90 mL, 39.2 mmol) and lutidine (9.7 mL, 83.3 mmol) in DME was added dropwise. The reaction mixture was allowed to warm from 0 °C to rt and was stirred at rt for 12 h. After solvent removal, the residue was redissolved in EtOAc, and the organic layer was washed with brine, dried over anhydrous $Na_2SO_4$, and concentrated under reduced pressure. The resulting residue was purified by silica gel column chromatography using EtOAc/hexane (1:1) as the mobile phase, affording the thiazole ester **3** (2.0 g, 69%) as a yellow solid. MS (ESI) m/z 301 $[M+H]^+$; HRMS: calcd for $C_{13}H_{20}N_2O_4S$ $(M+H)^+$ 301.1143 found 301.1213. $^1H$ NMR (600 MHz, $CDCl_3$) δ 8.07 (s, 1H), 5.22 (s, 1H), 5.11 (s, 1H), 4.41 (q, J = 7.1 Hz, 2H), 1.62 (d, J = 6.9 Hz, 3H), 1.44 (s, 9H), 1.39 (t, J = 7.1 Hz, 3H). $^{13}C$ NMR (150 MHz, $CDCl_3$) δ 161.50, 147.36, 61.58, 28.45, 21.94, 14.51.

## Synthesis of thiazole ester 4

**4** was synthesized from thioamide **2** (1.0 g, 2.80 mmol) according to the same procedure described for the synthesis of compound **3.** Thiazole ester **4** (1.0 g, 79%) was obtained as a yellow solid. MS (ESI) m/z 453 $[M+H]^+$; HRMS: calcd for $C_{21}H_{28}N_2O_5S_2$ $(M+H)^+$ 453.1439 found 453.1524. $^1H$ NMR (600 MHz, $CDCl_3$) δ 8.10 (s, 1H), 7.20 (d, J = 8.6 Hz, 2H), 6.83 (d, J = 8.6 Hz, 2H), 5.58 (s, 1H), 5.21 (s, 1H), 4.41 (q, J = 7.1 Hz, 2H), 3.79 (s, 3H), 3.53 (s, 2H), 3.17 – 2.95 (m, 2H), 1.46 (s, 9H), 1.39 (d, J = 7.1 Hz, 3H). $^{13}C$ NMR (150 MHz, CDCl3) δ 161.38, 158.93, 147.52, 130.23, 129.72, 127.73, 114.16, 61.62, 55.40, 52.43, 36.44, 36.20, 28.43, 14.50, 14.26.

## Synthesis of dipeptide 8

To a solution of compound **3** (0.45g, 1.46 mmol) in dichloromethane (DCM, 10 mL) was added trifluoroacetic acid (TFA, 3.3 mL). The reaction mixture was stirred at rt for 1 h. The solvent was removed under reduced pressure to give the thiazole amine **5** in quantitative yield. Compound **3** (0.50 g, 1.66 mmol) was dissolved in a mixture of THF/MeOH/$H_2O$ (3:1:1). Then NaOH (0.532 g, 13.3 mmol) was added, and the reaction mixture was stirred at rt for 1 h. After removing the solvent at reduced pressure, the residue was redissolved in EtOAc. The organic layer was washed with brine, dried over anhydrous $Na_2SO_4$, and concentrated in vacuo to give thiazole acid **6**. The resulting thiazole acid **6** (0.40 g, 1.46 mmol) was dissolved in anhydrous DMF (5 mL). HOBt (0.670 g, 4.38 mmol), HBTU (1.66 g, 4.38 mmol), and thiazole amine **5** (1.50 mmol, dissolved in DMF) were added sequentially. To this mixture was added $iPr_2EtN$ (1.25 mL, 7.2 mmol) with stirring at rt under $N_2$ for 12 h. Upon completion, the reaction was quenched with aqueous HCl (1 M). The solution was diluted with EtOAc, washed with saturated $NaHCO_3$ solution and brine, dried over $Na_2SO_4$, and filtered. The solvent was then removed in vacuo, and the residue was subjected to column chromatography on silica gel to provide the dipeptide **8** (0.560 g, 84%) as a white solid. MS (ESI) m/z 455 $[M+H]^+$; HRMS: calcd for $C_{19}H_{26}N_4O_5S_2$ $(M+H)^+$ 455.1344 found 455.1422. $^1H$ NMR (600 MHz, $CDCl_3$) δ 8.09 (s,1H), 8.03 (s, 1H), 7.86 (dd, J = 8.5, 3.4 Hz, 1H), 5.59 (tt, J = 8.7, 6.1 Hz, 1H), 5.10 (d, J = 37.7 Hz, 2H), 4.41 (q, J = 7.1 Hz, 2H), 1.79 (d, J = 7.0 Hz, 2H), 1.61 (d, J = 6.8 Hz, 2H), 1.45 (s, 9H), 1.39 (t, J = 7.1 Hz, 3H). $^{13}C$ NMR (150 MHz, $CDCl_3$) δ 173.18, 173.15, 160.60, 149.23, 147.30, 147.28, 127.59, 127.58, 123.98, 61.61, 47.27, 47.22, 28.46, 21.19, 21.16, 14.50.

## Synthesis of linear tripeptide 10

Reaction of dipeptide **8** (0.50 g, 1.10 mmol) with TFA (3.3 mL) in DCM (15 mL) afforded the amine **9** in quantitative yield. Compound **4** (0.50 g, 1.10 mmol) was dissolved in a mixture of THF/MeOH/$H_2O$ (3:1:1). NaOH (0.40 g, 8.84 mmol) was added, and the reaction mixture was stirred at rt for 1 h. After removing the solvent at reduced pressure, the residue was redissolved in EtOAc. The organic layer was washed with brine, dried over anhydrous $Na_2SO_4$, and concentrated in

vacuo to give thiazole acid **7**. The resulting thiazole acid 7 (0.40 g, 0.94 mmol) was dissolved in anhydrous DMF (5 mL), then HOBt (0.431 g, 2.82 mmol), HBTU (1.06 g, 2.82 mmol), and dipeptide amine **9** (1.10 mmol, dissolved in DMF) were added. To this mixture was added $iPr_2EtN$ (1.25 mL, 7.2 mmol) with stirring at rt under $N_2$ for 12 h. Upon completion, the reaction was quenched with aqueous HCl (1 M). The solution was diluted with EtOAc, washed with saturated $NaHCO_3$ solution and brine, dried over $Na_2SO_4$, and filtered. The solvent was then removed in vacuo, and the residue was subjected to column chromatography on silica gel to provide the linear tripeptide **10** (0.430 g, 60%) as a white solid. MS (ESI) m/z 761 $[M+H]^+$; HRMS: calcd for $C_{33}H_{40}N_6O_7S_4$ $(M+H)^+$ 761.1841 found 761.1911. $^1H$ NMR (600 MHz, $CDCl_3$) δ 8.10 (s, 1H), 8.08 (s, 1H), 8.06 (s, 1H), 7.89 (d, J = 8.2 Hz, 1H), 7.70 (d, J = 8.1 Hz, 1H), 7.16 (d, J = 8.6 Hz, 2H), 6.80 (d, J = 8.6 Hz, 2H), 5.58 (tt, J = 15.2, 7.1 Hz, 2H), 5.44 (s, 1H), 5.17 (s, 1H), 4.41 (q, J = 7.1 Hz, 2H), 3.76 (s, 3H), 3.60 – 3.53 (m, 2H), 3.00 (td, J = 12.0, 5.5 Hz, 2H), 1.79 (dd, J = 7.0, 3.9 Hz, 6H), 1.46 (s, 9H), 1.39 (t, J = 7.1 Hz, 3H). $^{13}C$ NMR (150 MHz, $CDCl_3$) δ 173.20, 172.77, 160.54,149.24, 147.28, 130.17, 129.49, 127.59, 124.59, 124.33, 114.17, 61.62, 55.39, 47.31, 47.28, 38.77, 36.27, 28.44, 21.22, 21.17, 14.50.

## Synthesis of cyclic peptide 11

Linear tripeptide **10** (0.20 g, 0.26 mmol) was treated with NaOH (0.084 g, 2.10 mmol) to hydrolyze the ethyl ester, then with TFA (1 mL) to remove the *N-t*-Boc protective group. The residue (0.164 g, 0.25 mmol) was dissolved in a mixture of DMF/DCM (2:1, 30 mL), then a solution of PyBop (0.299 g, 0.57 mmol) and 4-dimethylaminopyridine (0.143g, 1.17 mmol) in DMF/DCM (2:1, 44 mL) was added slowly over 10 h using a syringe pump. The reaction mixture was washed with aqueous HCl (1 M), saturated $NaHCO_3$, and brine, and dried over $Na_2SO_4$. The solvent was then removed in vacuo. The residue was purified by column chromatography on silica gel to give cyclic peptide **11** (0.102 g, 58%) as a white solid. MS (ESI) m/z 615 $[M+H]^+$; HRMS: calcd for $C_{26}H_{26}N_6O_4S_4$ $(M+H)^+$ 615.0898 found 615.0950. $^1H$ NMR (600 MHz, $CDCl_3$) δ 8.70 (d, J = 7.9 Hz, 1H), 8.65 (dd, J = 13.4, 7.7 Hz, 2H), 8.18 (s, 1H), 8.16 (s, 1H), 8.15 (s, 1H), 7.28 (d, J = 8.6 Hz, 2H), 6.85 (d, J = 8.6 Hz, 2H), 5.68 – 5.54 (m, 3H), 3.79 (s, 3H), 3.78 – 3.72 (m, 2H), 3.13 (dd, J = 13.8, 4.8 Hz, 1H), 2.84 (dd, J = 13.9, 8.4 Hz, 1H), 1.73 (d, J = 6.8 Hz, 6H). 13C NMR (150 MHz, $CDCl_3$) δ 171.46, 171.44, 159.74, 158.96, 149.04, 148.92, 130.36, 129.82, 124.78, 124.40, 124.17, 114.17, 114.13, 55.43, 53.57, 51.12, 47.57, 47.39, 38.19, 36.27, 25.13.

## Synthesis of AAC-DNPT

To a solution of cyclic peptide **11** (50 mg, 0.08 mmol) in DCM (2 mL) was added 2,4-dinitrobenzenesulfenyl chloride (22 mg, 0.09 mmol) and TFA (0.018 mL, 0.24 mmol). The reaction mixture was stirred at rt for 1 h. Reaction was quenched with $H_2O$. The organic layer was washed with brine, dried over anhydrous $Na_2SO_4$, and the solvent was then removed in vacuo. The residue was purified by column chromatography on silica gel to give AAC-DNPT (32 g, 57%) as a yellow solid. MS (ESI) m/z 693 $[M+H]^+$; HRMS: calcd for $C_{24}H_{20}N_8O_7S_5$ $(M+H)^+$ 693.0058 found 693.0143. $^1H$ NMR (600 MHz, $CDCl_3$) δ 9.08 (d, J = 2.3 Hz, 1H), 8.71 (d, J = 7.5 Hz, 1H), 8.61 (dd, J = 15.0, 7.5 Hz, 2H), 8.49 (d, J = 9.0 Hz, 1H), 8.42 (dd, J = 9.0, 2.4 Hz, 1H), 8.25 (s, 1H), 8.17 (s, 1H), 8.14 (s, 1H), 5.83 (dt, J = 7.5, 5.9 Hz, 1H), 5.68 – 5.58 (m, 2H), 3.45 (d, J = 5.8 Hz, 2H), 1.75 (t, J = 6.8 Hz, 6H). $^{13}C$ NMR (150 MHz, $CDCl_3$) δ 171.85, 159.99, 159.66, 149.61, 148.87, 148.18, 145.83, 145.53, 145.29, 128.91, 127.63, 125.10, 124.91, 124.34, 121.77, 50.87, 47.55, 44.91, 25.08.

## Expression and purification of single-cysteine mutants of Pgp

Single-Cys mutant constructs of murine Pgp (Mdr1a, accession number NM_011076, GenBank JF83415) were generated on a Cysless Pgp (CL-Pgp) background in the *Pichia pastoris* pPIC-CL-*mdr1a* expression vector[38] by site-directed mutagenesis. For cryoEM structural studies of L335C and V978C, we further substituted the catalytic carboxylates in both NBDs to glutamines, E552Q/E1197Q, to generate ATP hydrolysis-deficient mutants. The Pgp construct used in this study contained C-terminal hexahistidine and Twin-Strep purification tags [40]. Large-scale Pgp biomass production in *Pichia pastoris* and microsomal membrane preparations were conducted according to published protocols[38], [39].

For ATPase activity measurement and MS analysis, we purified Pgp in the presence of n-dodecyl-D-maltopyranoside (DDM) supplemented with the lipid 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE). Briefly, microsomes were resuspended in 20 mM Tris (pH 8.0), 20 mM imidazole, 20% glycerol, 500 mM NaCl, protease inhibitors (10 µg/mL leupeptin and pepstatin A, 2.5 µg /mL chymostatin,1 mM PMSF), 0.2 mM tris(2-carboxyethyl)phosphine with 1% DDM for 60 min at 4 °C. After centrifugation at 38,000 × *g* for 30 min, Pgp was purified from the supernatant using Ni-NTA affinity chromatography in Buffer A (50 mM Tris pH 8.0, 150 mM NaCl, 20% glycerol) supplemented with 0.067% DDM, 0.04% sodium cholate, and 0.1 mg/mL POPE, and with 20 mM imidazole for wash buffer or 200 mM imidazole for elution buffer. The eluate from Ni-NTA was concentrated for further purification by size exclusion chromatography on a Superdex 200 Increase 10/300 column using 20 mM Tris pH 7.5, 150 mM NaCl, 0.067% DDM, 0.04% sodium cholate, and 0.1 mg/ml POPE.

For cryoEM structural determination, we purified Pgp in a mixture of lauryl maltose neopentyl glycol (LMNG) and cholesteryl hemisuccinate (CHS). Briefly, after the solubilization of microsomes in DDM, Pgp was purified from the supernatant using Ni-NTA affinity chromatography in Buffer A supplemented with 0.02% LMNG, and 0.004% CHS. The eluate from Ni-NTA was applied to pre-equilibrated Strep-Tactin Superflow resin and incubated at 4 °C for 1 h. Flow-through was removed, and the resin was washed with Buffer A in the presence of 0.02% LMNG, and 0.004% CHS. Pgp was eluted with the same buffer containing 2.5 mM desthiobiotin. The eluate from Strep-Tactin was concentrated for further purification by size exclusion chromatography on a Superdex 200 Increase 10/300 column using a detergent-free buffer containing 20 mM Tris pH 7.5 and 200 mM NaCl.

## Crosslinking between single cysteine mutants of Pgp and AAC-DNPT

The covalent Pgp complexes were typically prepared by the reaction of Pgp (3-5 mg/mL) with 10-fold excess of AAC-DNPT at room temperature for 30 min. L971C labeling was conducted in the presence of MgATP (10 mM). The DNPT color formation was visualized by eye or monitored by UV-visible spectroscopy ($\lambda_{max}$ = 408 nm, ε = 13800 $M^{-1}$ $cm^{-1}$). For mass spectrometric analysis, AAC-labeled Pgp sample was passed through a PD-10 Sephadex G-25 desalting column to remove excess ligand, and the eluate was further treated by addition of cysteine (5 mM) before performing trypsin digestion.

## ATPase activity assay

ATPase activity of Pgp, with or without AAC labeling, was measured at 37 °C using an enzyme-coupled ATP regeneration system[41]. Briefly, 1 µg Pgp was added to 100 µL of ATP cocktail (50 mM Tris, pH 7.5, 12 mM $MgCl_2$, 6 mM phosphoenolpyruvate, 1 mM NADH, 10 units lactate dehydrogenase, 10 units pyruvate kinase, and 10 mM ATP). The rate of ATP hydrolysis was determined by the decrease in NADH absorbance at 340 nm using a microplate reader (Filtermax F5). Verapamil was added from stocks in water, QZ-Ala, AAC-DNPT, FK506 and valinomycin were added from stocks in DMSO such that the final DMSO concentration was ≤ 1%. ATPase activity was calculated as described previously[42]. To analyze the activities of Pgp mutants (L335C, V978C, and L971C, with CL-Pgp as control) with varying concentrations of AAC-DNPT (0-20 µM), we incubated Pgp with AAC-DNPT for 15 min at room temperature prior to the addition of ATP to initiate the ATPase reaction.

## Trypsin digestion of Pgp

Ten µL of 5 mg/mL Pgp, with or without AAC labeling, was added to an S-Trap micro column (Farmingdale NY). Then 15 µL of 100 mM triethylammonium bicarbonate (TEAB) buffer (pH 7.5) containing 10% sodium dodecyl sulfate (SDS) was drawn and mixed with Pgp protein by pipette. Then 2.5 µL of 10% (v/v) $H_3PO_4$ in water was added to the S-Trap. After 10 min incubation, 165 µL

of binding solution (100 mM TEAB in MeOH/$H_2O$ 9:1 (v/v), pH 7.1) was added to the acidified Pgp protein. After 10 min incubation, the S-Trap was seated in a 1.5 mL tube and centrifuged at 4000 × *g* for 2 min until all solution had passed through the S-Trap membrane. The flow-through liquid was drawn back and centrifuged again. After addition of 150 µL of binding solution, the S-Trap was centrifuged at 4000 × *g* for 2 min to wash the protein. The flow-through liquid was removed, and this washing procedure was repeated three times. Trypsin (4 µg, at an enzyme/protein ratio of 1:12.5, w/w) was added to the S-Trap and mixed well. The S-Trap was capped loosely to limit evaporation, then incubated in a dry incubator at 37°C for overnight digestion. After digestion was completed, the resulting peptides were eluted by adding 40 µL of $H_2O$ containing 0.2% formic acid (FA), and centrifuging at 4000 × *g* for 2 min. The flow-through liquid was transferred back to be centrifuged again. Then 35 µL of $CH_3CN$/$H_2O$/FA (80:20:0.2%) was added, and the S-Trap was centrifuged at 4000 × *g* for 2 min. The eluted peptides were collected for LC-MS analysis.

## LC-MS analysis

The LC-MS setup consisted of an ultra-performance liquid chromatography instrument (UPLC, Waters, Milford, MA) coupled with a high-resolution Orbitrap Q Exactive mass spectrometer (Thermo Scientific, San Jose, CA). A reversed-phase column (BEH C18, 1.0 mm × 100 mm, 1.8 µm) was used for separation. The injection volume was 5 µL per analysis. For gradient elution (0-95% $CH_3CN$ with 0.1% FA in water), the mobile phase flow rate was 30 µL/min. The Orbitrap mass spectrometer was equipped with a heated electrospray ionization (HESI) source. The rate of sheath gas flow was 10 L/h and the applied ionization voltage was +4 kV. The ion transfer inlet capillary temperature was kept at 250 °C. Mass spectra were acquired using Thermo Xcalibur (3.0.63) software. The scan mode was set to full scan $MS^1$, followed by data-dependent $MS^2$ acquisition. The resolution of full scan $MS^1$ was 70k and the automatic gain control (AGC) target was set to 5e5. For $MS^2$ acquisition, the resolution was 17.5k, and AGC target was 2e4. The 20 most abundant ions (+2 to +6 ions) were selected to fragment with a normalized collision energy of 30%.

## MDCK-MDR1 transport assay

The Madin Darby Canine Kidney (MDCK) epithelial cells stably transfected with the human *MDR1* gene forms a confluent monolayer, which is widely adopted to evaluate if a compound is subject to Pgp efflux based on the permeability measurement in both directions[43]. The MDCK-MDR1 permeability assay for QZ-Ala was conducted by Bioduro-Sundia Inc. (San Diego, CA). Briefly, 5 µM QZ-Ala in the absence or presence of 10 µM cyclosporin A was added to either the apical (A) or the basolateral (B) side and the amount of permeation was determined on the other side of the monolayer by LC-MS/MS. The efflux ratio ($R_E$ = $P_{app}$ (B-A) to $P_{app}$(A-B)) for QZ-Ala was determined following standard protocols. An $R_E$ > 2.0 generally indicates a substrate for Pgp.

## G72A mutagenesis and drug resistance assays

G72A mutation was conducted on the mouse CL-Pgp template in the pVT expression vector (pVT-CL-*mdr1a*)[44]. First, the three N-glycosylation sites N82/N87/N90 that were previously substituted by Gln for X-ray crystallography were restored to the original codons 5'-AACgtgtccaagAACagtactAAT-3' by QuickChange site-directed mutagenesis. The G72A mutation was then added by a second round of site-directed mutagenesis, and the full-length open reading frame sequenced to confirm no other unwanted mutation was present. Plasmids from three individual clones, together with WT-Pgp and CL-Pgp as well as pVT "empty" vector controls were transformed into *S. cerevisiae* JPY201 (MATa ura3 Δste6::HIS3) cells for expression and functional assays that were performed essentially as previously described[44]. Briefly, 10 ml yeast cultures were grown overnight in uracil-deficient minimal medium, diluted to $OD_{600}$=0.05 in YPD medium [1% (w/v) yeast extract/2% (w/v) peptone/2% (w/v) glucose], and seeded into 96-well plates containing YPD alone or YPD plus 40 µM doxorubicin, 50 µM FK506 or 100 µM valinomycin. Samples were grown in triplicate wells at 30°C for up to 40 h, and yeast cell growth was monitored by measuring the $OD_{600}$ at 2 h increments in a microplate reader (Benchmark Plus, BioRad). The remainder of the 10

ml cultures was used to assess Pgp expression by Western blot analysis of microsomal membrane preparations using the monoclonal C219 anti-Pgp antibody. For ATPase assays, the three N-glycosylation sites were restored in the *P. pastoris* pPIC-CL-*mdr1a* expression vector and the G72A mutant introduced by site-directed mutagenesis. For cry-EM, L335C was added to the G72A mutant by mutagenesis; the integrity of the open reading frame was confirmed by DNA sequencing after each round of mutagenesis. G72A and G72A/L335C mutant proteins were purified from *P. pastoris* microsomal membranes as described for single Cys mutants. Bars in **Fig. 3** represent the mean of ≥3 independent experiments ± SEM; two way ANOVA with post hoc Bonferroni tests identified those pairs with very highly significant differences ($p < 0.001$).

## EM sample preparation

After shipment, quality control of the samples was performed by collecting negative stain EM images on a Tecnai $G^2$ Spirit TWIN TEM (FEI) [45]. For cryoEM, all samples were adjusted to 3.5 mg/mL in 50 mM Tris, pH 7.5 and 200 mM NaCl. OF335 was obtained after reaction with a 4-fold excess of AAC-DNPT; OF978C and 971C were obtained with 10-fold molar excess for 30 min at room temperature. To trigger NBD-dimerization, the QQ constructs were then incubated with 5 mM MgATP for 1 h at room temperature prior to grid preparation. For OF971, 5 mM $Mg^{2+}$ATP/Vi was added to the sample and incubated for 1 h at room temperature. Freezing protocol was followed as previously described[46]. All samples were vitrified on freshly glow-discharged CF-1.2/1.3 TEM grids (Protochips, USA) with a Vitrobot Mark IV (Thermo Fisher Scientific, Inc., USA) at 100% humidity and 4 °C, with a nominal blot force of –2 and a blotting time of 12 s. Grids were plunged into liquid ethane and stored in liquid nitrogen until further use.

## EM data acquisition and processing

The data sets for OF978 and OF971 were acquired on a Titan Krios G4, operated at 300kV and equipped with a Selectris X imaging filter and a Falcon 4 direct electron detector (all Thermo Fisher Scientific, USA). Datasets were obtained using automation strategies of EPU software v2.13 (Thermo Scientific) at a nominal magnification of 215,000, corresponding to a calibrated pixel size of 0.573 Å. The camera was operated in electron counting mode, and the data were saved in electron-event representation (EER) format. All other data sets were obtained on a Titan Krios G3i (Thermo Fisher Scientific, USA), using automation strategies of EPU 2.9 or newer, equipped with a Gatan BioQuantum K3 Imaging Filter (Gatan, USA) in electron counting mode. The nominal magnification was 105,000 corresponding to a calibrated pixel size of 0.837 Å. The exposure time for all data sets was ~4s, and the total dose was 70 $e^-/Å^2$ (Selectris X - Falcon 4) or 75 $e^-/Å^2$ (BioQuantum K3). Quality of the data was monitored during collection using cryoSPARC live v3.2.0 and v3.3.1[47]. Details about number of collected images and picked particles, as well as number of particles in the final map and resolutions are listed in **Extended Data Table 1**.

For data processing for IF335, the initial model was obtained using cryoSPARC v3.2.0. Further processing, including 3D classifications and refinements, was obtained in Relion 3.1[48]. For the final refinement, particles and map were transferred back to cryoSPARC v3.2.0 to run a Non-Uniform Refinement (NUR)[49]. All other data sets were processed in cryoSPARC v3.3.1. Particles were picked broadly with the blob picker, and the best classes were selected for further processing. Sorting of the particles was achieved by multiple heterogenous refinements and NUR. Global and local CTF refinements[50] were performed toward the end of the processing pipeline, followed by another round of NUR. An exemplary processing pathway is provided in **Extended Data Fig. 10**. Processing results for all structures are shown in **Extended Data Fig. 6**.

For all data sets, the images were repicked with the Topaz[51] picker, and this increased the resolution for all three maps from the OF335 data set. For all other data sets, no improvement of the maps could be achieved with this approach. Density modification with *phenix.resolve_cryo_em*[52] was carried out using two half-maps together with the FSC-based

resolution and the molecular masses of the molecules. This procedure resulted in significant improvement of the map qualities. The density-modified map of OF335-1lig was used for **Extended Data Fig. 7**.

## Model building

For all datasets, the structures of Pgp in the inward and outward conformations (PDBID: 4Q9I and 6C0V, respectively) were used as templates. All structures were manually edited in COOT[53] and refined using *phenix.real_space_refine,* in combination with rigid-body refinement[54] and several rounds of rebuilding in COOT. A quality check of all structures with MolProbity[55] indicated excellent stereochemistry with 93.1% - 98.1% of the non-glycine and non-proline residues found in the most-favored region, and 0.00% - 0.09% outliers (all-atom clashscore: 9.29 – 15.76). Refinement and validation statistics are summarized in **Extended Data Table 1**. Figures were drawn with ChimeraX[56] and PyMOL (The PyMOL Molecular Graphics System, Version 2.0, Schrödinger, LLC).

## Molecular dynamics simulations

Molecular dynamics simulations were performed starting from the structures OF978-1lig and OF335-2lig embedded in a patch of lipid bilayer. The bound cholesterol hemisuccinate molecules in the cryoEM structures were replaced with cholesterol in the same binding poses. Bound magnesium ions and ATP were also retained. The crosslinked ligands were removed and the crosslinking cysteine was replaced by the amino acid originally present at that position. No attempt was made to model the unresolved linker (residue 626-686). Instead the C-terminus of the first subunit (residue 625) was N-methylated and the N-terminus of the second subunit (residue 687) was acetylated to imitate the presence of an unstructured loop region connecting the two subunits. The structures were then inserted into a model plasma membrane and solvated using CHARMM-GUI[57], [58] (composition outer leaflet: 30% CHL, 35% POPC, 35% PSM; composition inner leaflet: 30% CHL, 25% PAPC, 25 POPE, 25% POPS). All protonation states were set according to propka3[59], [60]. The resulting simulation systems were placed in rectangular boxes with a size of approximately 12 × 12 × 17 $nm^3$ and around 260,000 atoms each. The ligands were modeled using the CGenFF server[61], [62] and placed into the solvated systems at their original location as in the pdb structures with the crosslink to the protein removed.

All molecular dynamics simulations were performed with the CHARMM36m forcefield[63] (version july2021) including CGenFF parameters[64] (version 4.6) using gromacs[65], [66] (version 2021.6). Hydrogen bond lengths were constrained using LINCS[67] in all simulations. The simulation systems were energy minimized and subsequently equilibrated in six steps while gradually releasing restraints on the positions of the heavy and backbone atoms of the protein, the lipid atoms and the ligand atoms (see **Extended Data Table 2**).

The production simulations were started from the last configuration of the equilibration with random initial velocities. All production simulations were performed in the NPT ensemble using a semi-isotropic Parrinello-Rahman barostat[68] with a target pressure of 1 bar and a coupling constant of 5 ps. The velocity rescale thermostat[69] with a time constant of 1 ps was used to keep the target temperature of T=400 K. We used a time step of 0.002 ps.

# Acknowledgements

We thank J.-H. Leung, W.-H. Lee, and A. Ward for early contributions, C. Katz for technical assistance, S. Welsch and S. Prinz for support during the cryoEM data collection, W. Kühlbrandt for generous access to the cryoEM facilities in Frankfurt and general support, and W. Kühlbrandt and I. Wilson for critical reading and editing of the manuscript. This work was supported by National Institute of General Medical Sciences R01 GM118594 (QZ) and GM141216 (ILU), the Deutsche

Forschungsgemeinschaft (DFG) Sonderforschungsbereich 944 and 1557 and INST 190/196 (AM), the Bundesministerium für Bildung und Forschung (BMBF) 01ED2010 (AM), the Max Planck Society (HJ, GH,AM & TG), and the South Plains Foundation (ILU).

## Author contributions

TG performed cryoEM and single particle analysis with help from DJ and AM. DW conducted ligand synthesis and crosslinking, identified Pgp single Cys mutants, performed biochemical characterization, and optimized and prepared Pgp samples for mass spectrometry and cryoEM studies under the supervision of QZ. PZ performed mass spectrometric characterization of crosslinking under the supervision of HC. KP built molecular models with help from TG and AM. JZ expressed Pgp mutants, and NT assayed G72A function under the supervision of ILU. HJ conducted molecular dynamics simulations under the supervision of GH. QZ and ILU conceived the Pgp labeling experiments and designed and tested the mutations. AM designed and supervised the cryoEM research. AM, TG and DJ interpreted the cryoEM structural data with significant input from QZ. QZ conceived and initiated the project, designed the research, and coordinated collaborations. The manuscript was written by TG, DW, DJ, ILU, QZ and AM. All authors carefully discussed and interpreted the data.

## Competing interests

Authors declare that they have no competing interests.

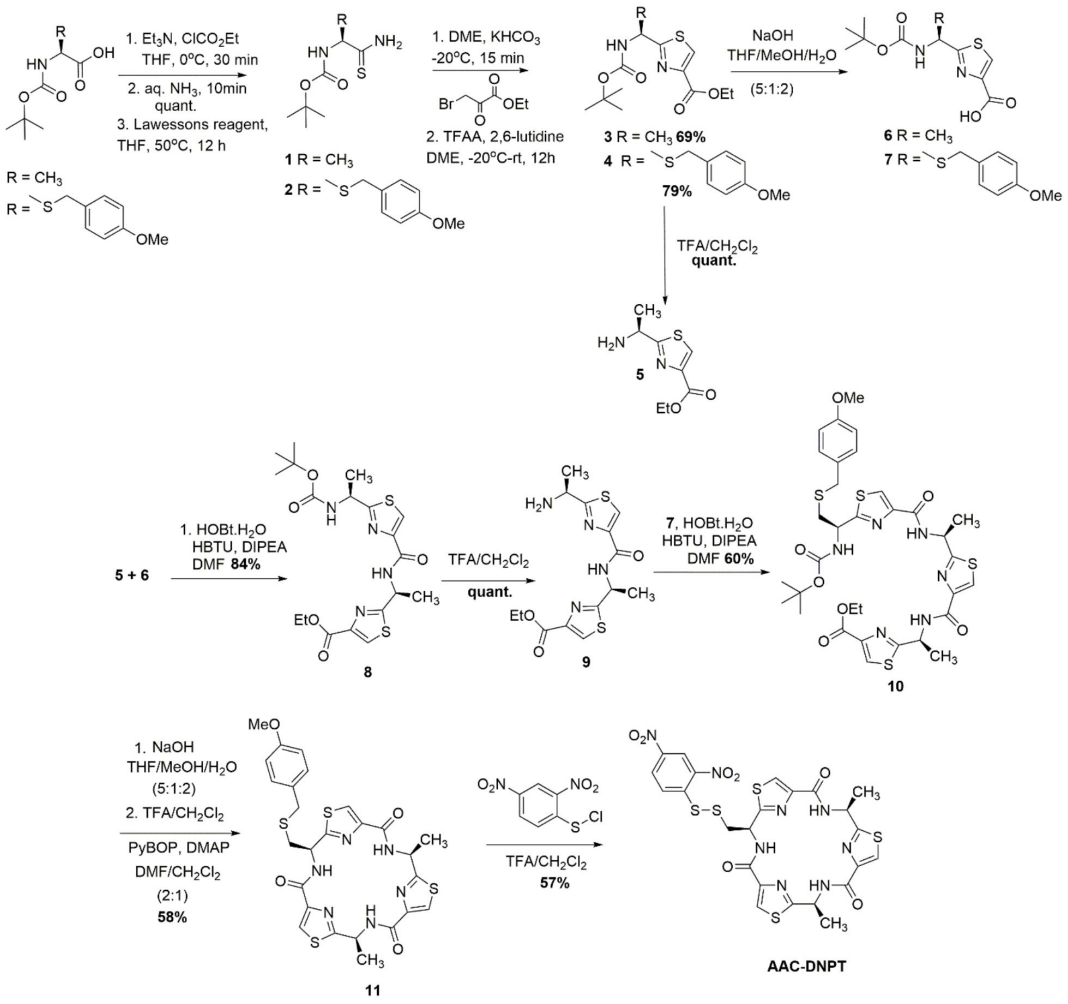
## Data availability

Supplementary Information is available for this manuscript. All cryoEM density maps have been deposited in the Electron Microscopy Data Bank under accession numbers EMD-14754, EMD-14755, EMD-14756, EMD-14758, EMD-14759, EMD-14760, EMD-14761 and EMD-17630.

Atomic coordinates for the atomic models have been deposited in the Protein Data Bank under accession numbers 7ZK4, 7ZK5, 7ZK6, 7ZK8, 7ZK9, 7ZKA, 7ZKB and 8PEE.

## Additional Information

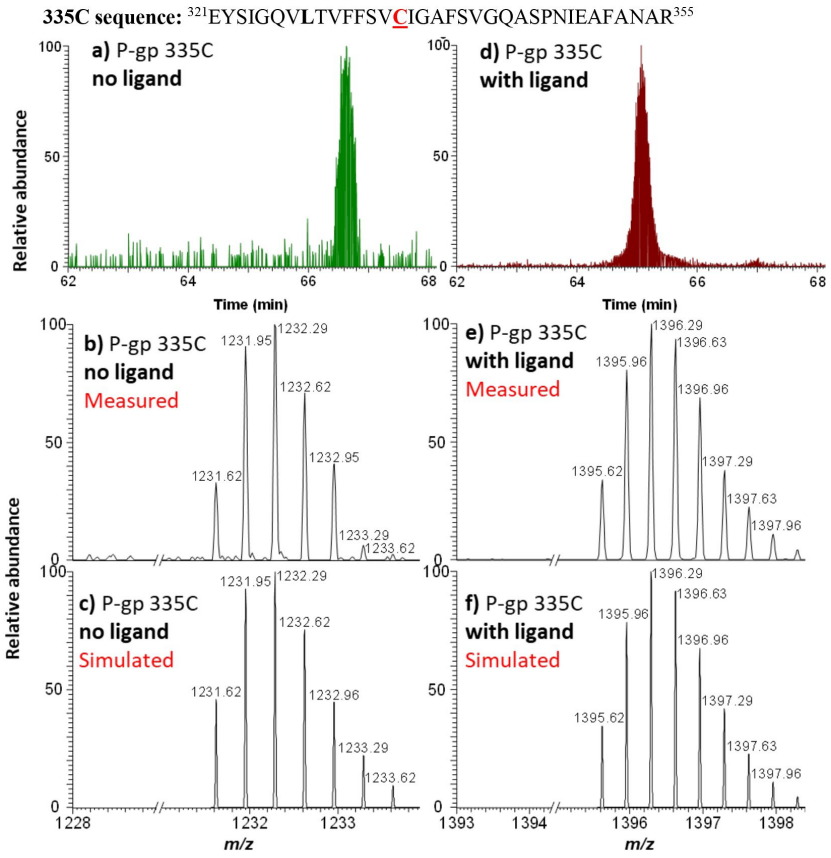
Correspondence and requests for materials should be addressed to Q.Z. or A.M.




## Extended Data Fig. 1.

### Chemical synthesis of AAC-DNPT.

NMR characterization data of synthetic compounds are provided in the Supplementary Information.

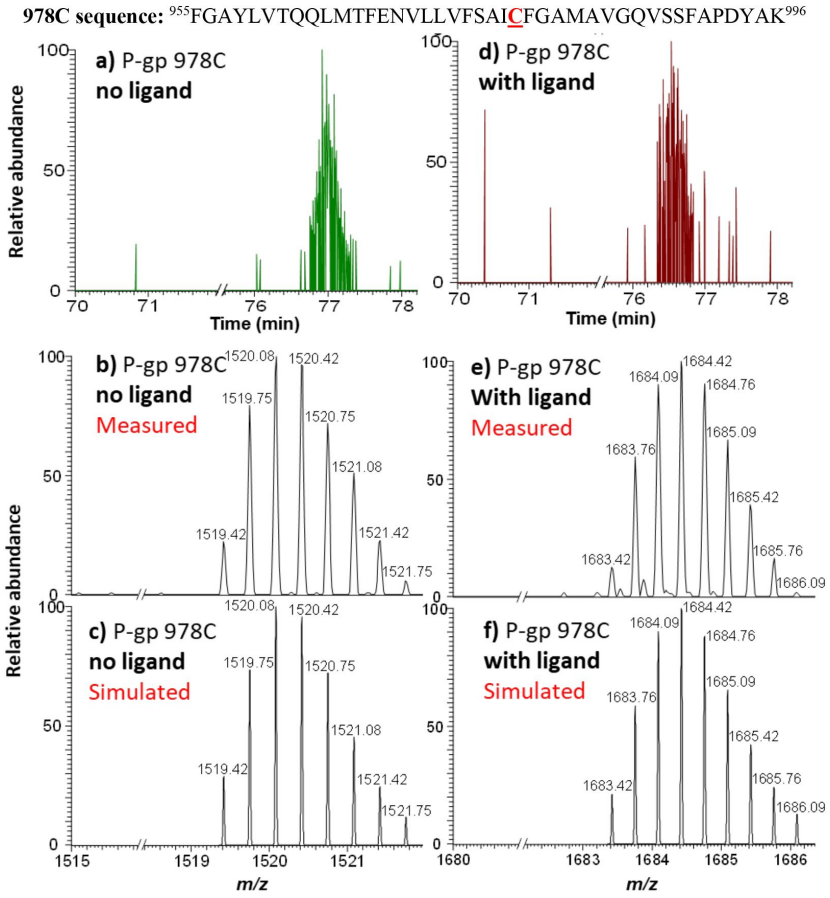



| | Observed *m/z* | Theoretical *m/z* | Mass error (ppm) |
|---|---|---|---|
| **335C** | 1232.28652 | 1232.28728 | 0.62 |
| **335C+L** | 1396.29317 | 1396.29261 | 0.40 |

## Extended Data Fig. 2.

### MS detection of 335C peptide before and after AAC labeling.

Top: 335C peptide sequence. **(a)** LC-MS trace for the 335C peptide, with **(b)** measured and **(c)** simulated m/z ($[335C+3H]^{3+}$). **(d)** LC-MS trace for the AAC-labeled 335C peptide, with **(e)** measured and **(f)** simulated m/z ($[335C+L+3H]^{3+}$). The observed isotope peak distributions for unlabeled and labeled 335C peptides are in good agreement with the simulated peak distribution; mass errors are shown in the table underneath.

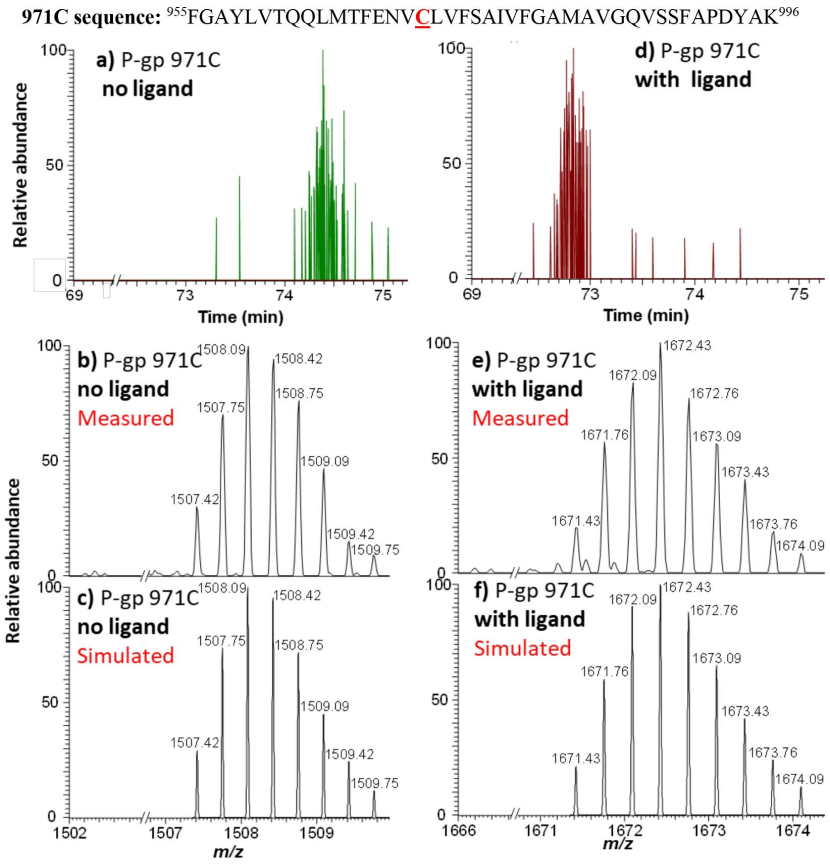



| | Observed *m/z* | Theoretical *m/z* | Mass error (ppm) |
|---|---|---|---|
| **978C** | 1520.08379 | 1520.08515 | 0.89 |
| **978C+L** | 1684.42427 | 1684.42447 | 0.12 |

## Extended Data Fig. 3.

### MS detection of 978C peptide before and after AAC labeling.

Top: 978C peptide sequence. **(a)** LC-MS trace for the 978C peptide, with (**b)** measured and **(c)** simulated m/z ($[978C+2H+Na]^{3+}$). **(d)** LC-MS trace for the AAC-labeled 978C peptide, with **(e)** measured and **(f)** simulated m/z ($[978C+L+2H+Na]^{3+}$). The observed isotope peak distributions for unlabeled and labeled 978C peptides are in good agreement with the simulated peak distributions; mass errors are shown in the table underneath.

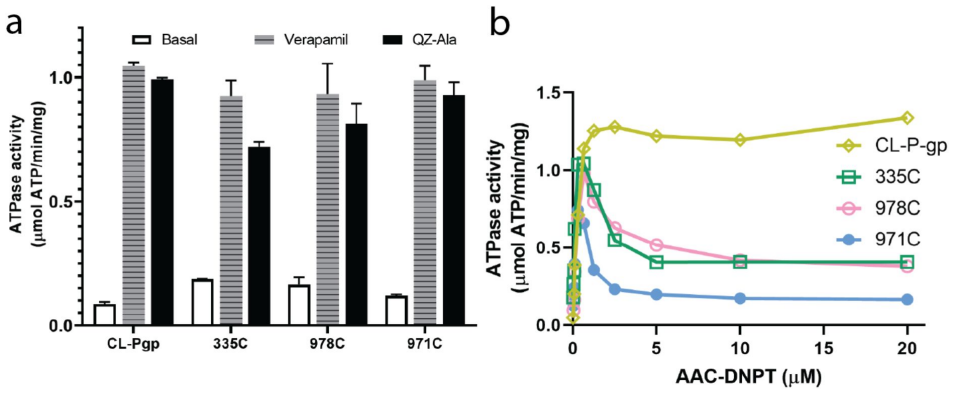



| | Observed *m/z* | Theoretical *m/z* | Mass error (ppm) |
|---|---|---|---|
| **971C** | 1508.08876 | 1508.08595 | 1.86 |
| **971C+L** | 1672.42763 | 1672.42527 | 1.41 |

## Extended Data Fig. 4.

### MS detection of 971C peptide before and after AAC labeling.

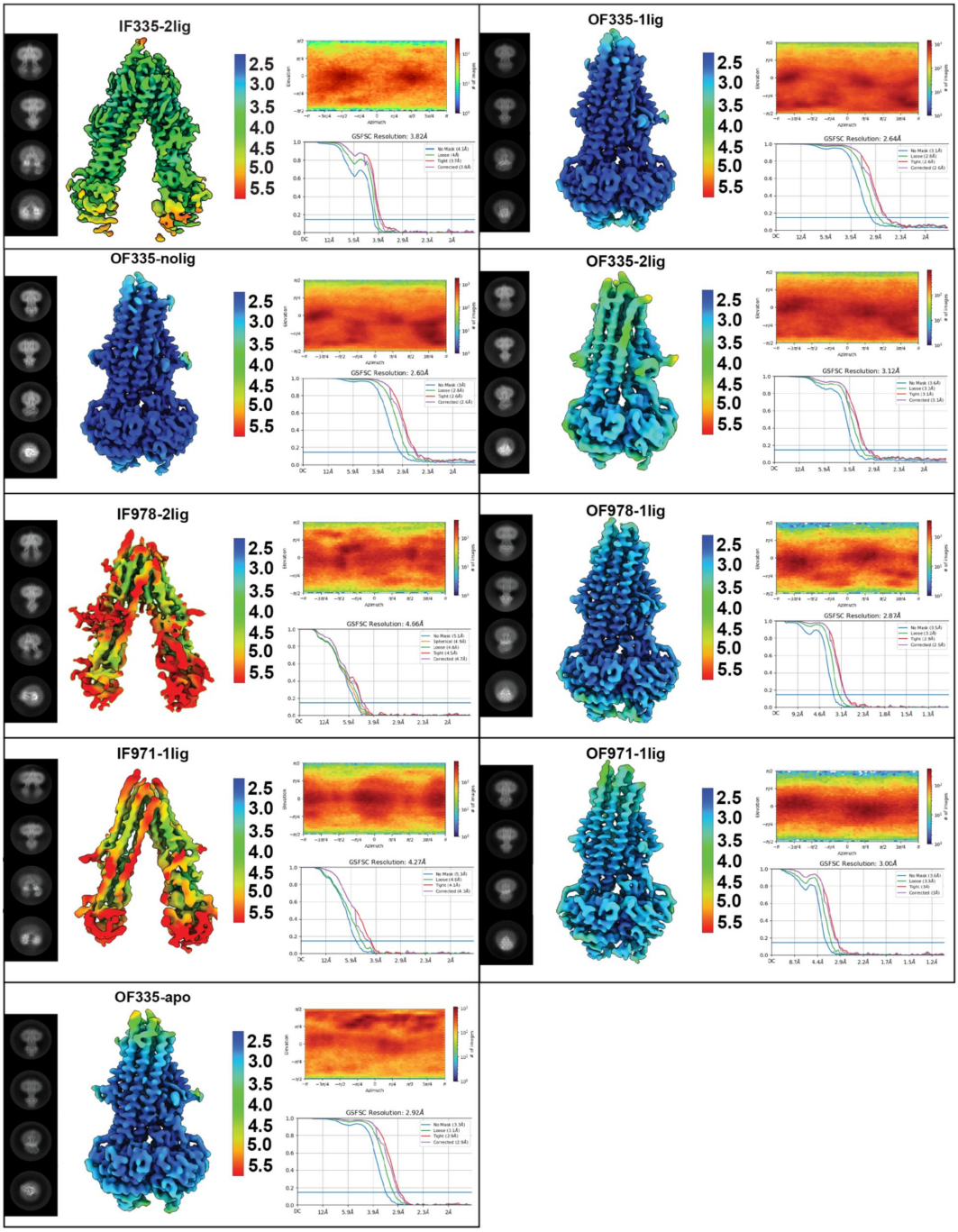
Top: 971C peptide sequence. **(a)** LC-MS trace for the 971C peptide, with **(b)** measured and **(c)** simulated m/z ($[971C+3H]^{3+}$). **(d)** LC-MS trace for the AAC-labeled 971C peptide, with **(e)** measured and **(f)** simulated m/z ($[971C+L+3H]^{3+}$). The observed isotope peak distributions for unlabeled and labeled 971C peptides are in good agreement with the simulated peak distributions; mass errors are shown in the table underneath.




## Extended Data Fig. 5.

### ATPase activities of Pgp mutants L335C, V978C, and L971C.

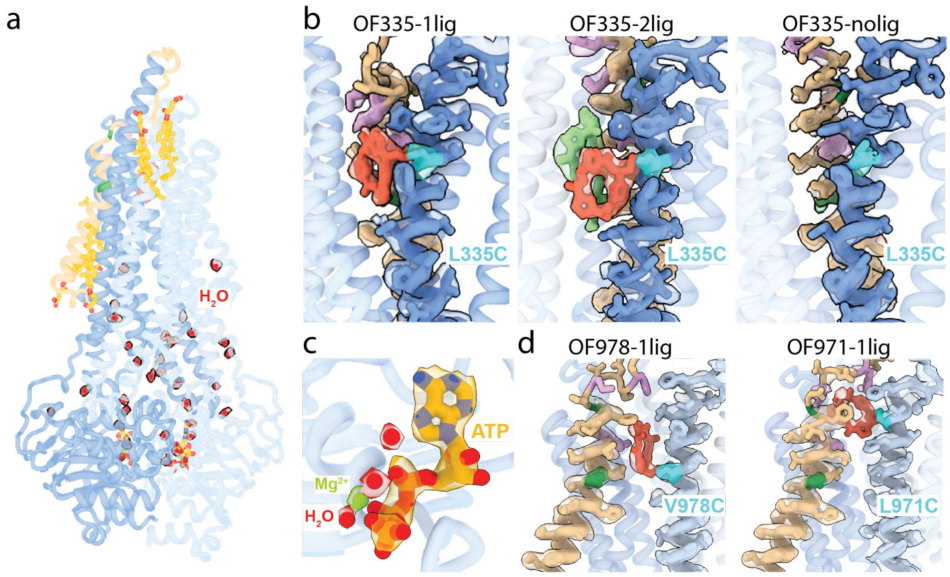
**(a)** ATPase activity was enhanced for each mutant after addition of verapamil (30 $\mu$M) or QZ-Ala (2 $\mu$M). Data are averaged from triplicate measurements with standard error bars. **(b)** ATPase activity after addition of various concentrations of AAC-DNPT. CL-Pgp was included as control. A representative curve was shown.




## Extended Data Fig. 6.

### Summary of cryo-EM data analysis for all structures.

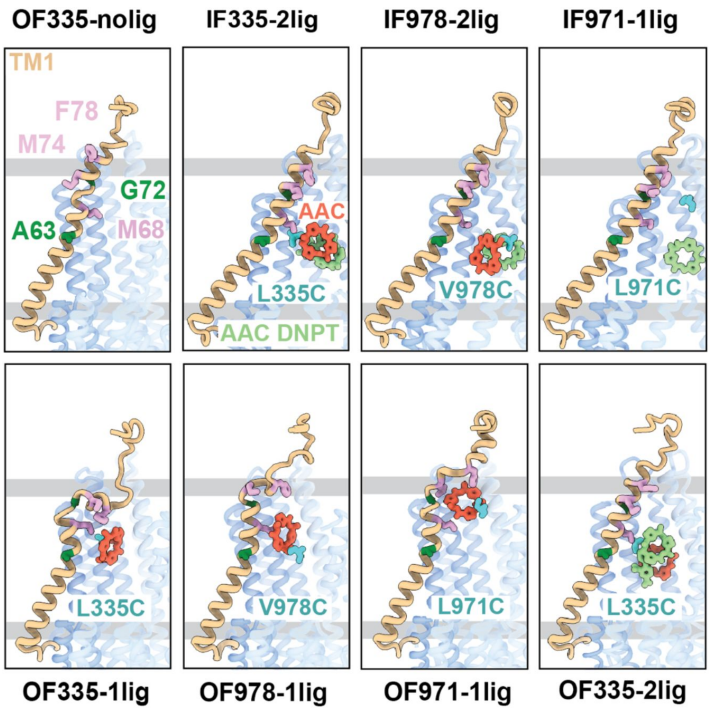
Each panel shows 2D classes, local resolution estimation, viewing distribution and FSC curves for each map. All data shown here were obtained using cryoSPARC.




**Extended Data Fig. 7.**

**High resolution features of the OF structures.**

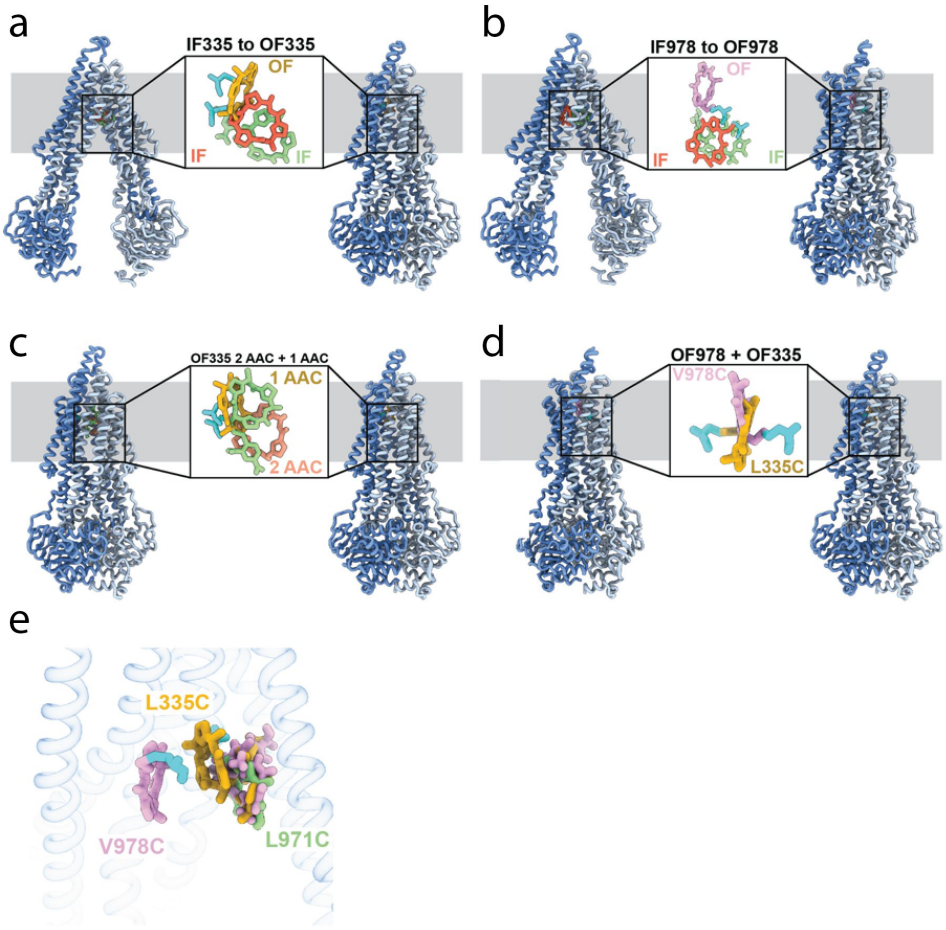
**(a)** Overall structure of OF335-1lig solved at 2.6 Å resolution. High resolution allows for fitting of water (red balls) and the surface-bound cholesterol hemisuccinate molecules (yellow sticks), in addition to the bound nucleotides and the AAC cyclopeptide. Contour level in the density modified map is 0.6. **(b)** Local densities for TM6 (blue) and TM1 (brown) in the OF Pgp335 showing the disulfide bridge in cyan between AAC (red) and TM6. AAC-DNPT is shown in light green in the OF335-2lig. Contour level used is 0.19. **(c)** Cryo-EM densities for ATP (yellow), $Mg^{2+}$ (green) and water molecules (red) in the OF335-1lig model. Contour level in the density modified map is 0.6 (water), 1.1 (ATP) and 1.7 ($Mg^{2+}$). **(d)** Disulfide bridge (cyan) between TM12 (light blue) and AAC (red) for OF978-1lig and OF971-1lig. Contour levels used are 0.19 (OF978-1lig) and 0.05 (OF971-1lig). The densities for A63 and A72 (dark green) and M68,M74 and F78 (plum) are shown in each panel.




**Extended Data Fig. 8.**

**Comparison of the binding cavities of all structures from the crosslink experiments.**

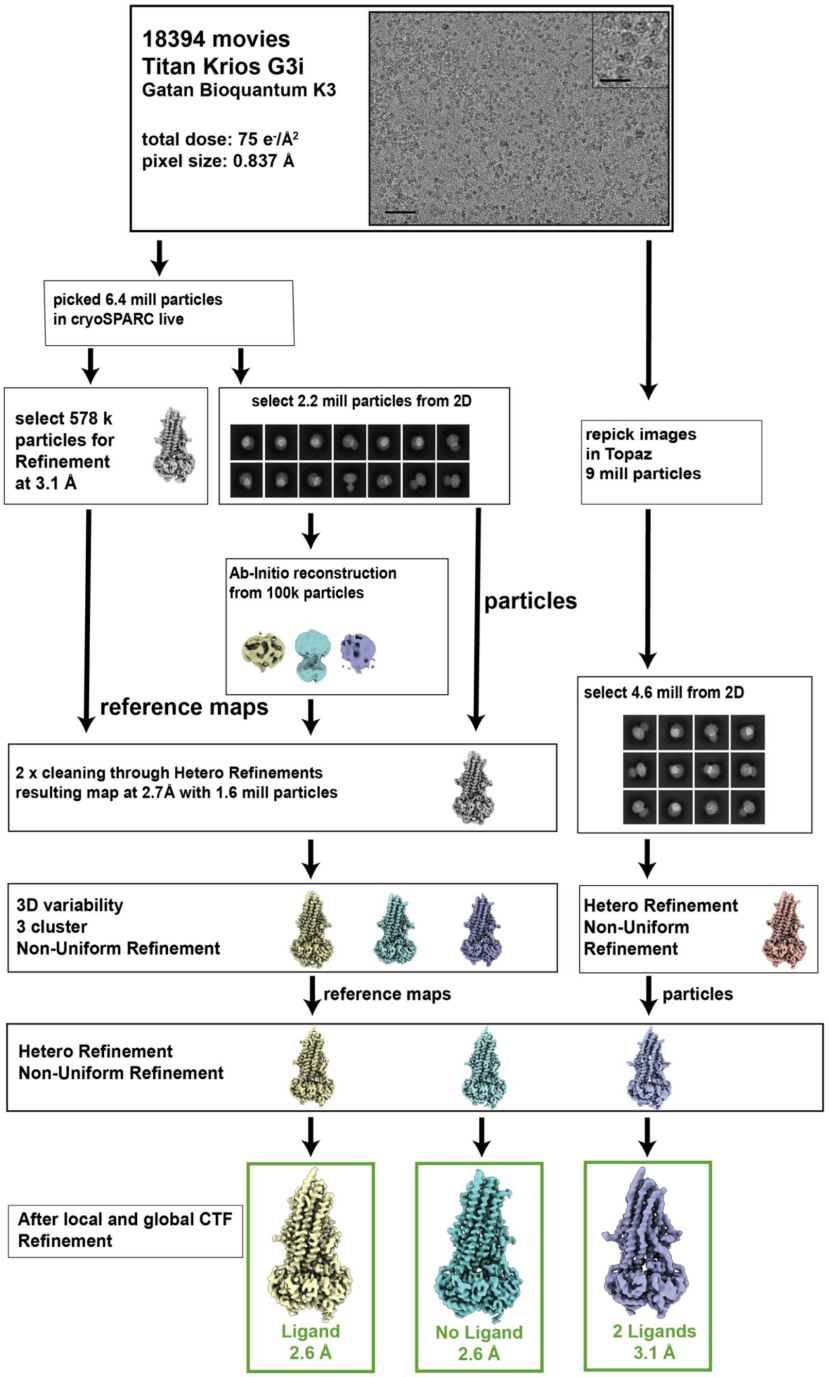
TM1 is shown in brown, the crosslinked AAC in red and the non-crosslinked AAC-DNPT in light green, the crosslinked residue is marked in cyan. A63 and G72 are shown in dark green; M68, M75 and F78 are shown in plum.




## Extended Data Fig. 9.

### Comparison of AAC binding in the different Pgp structures.

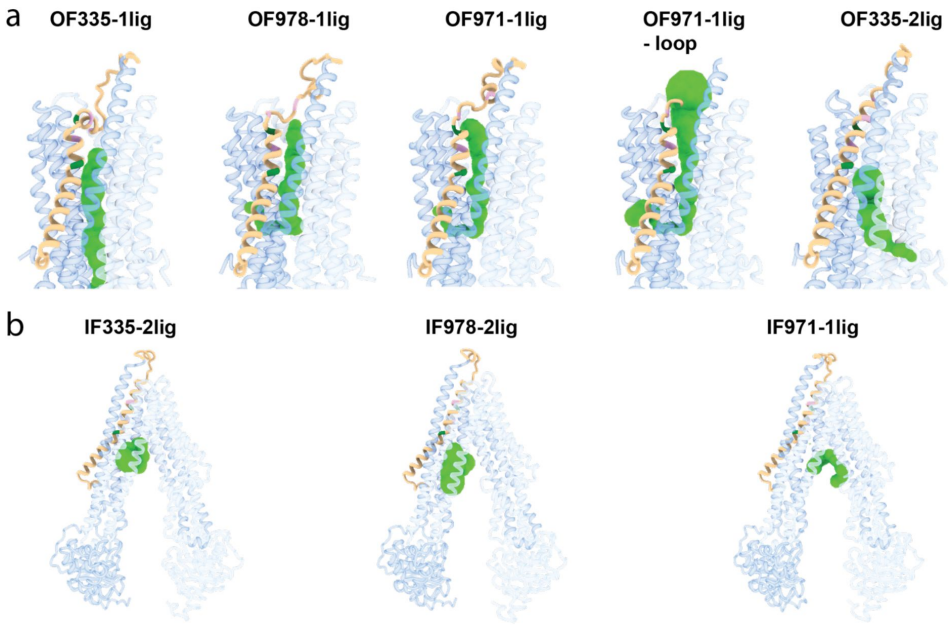
**(a)** Transition from IF (left) to OF (right) conformation for L335C. The IF ligand is shown in red for the crosslinked molecule, and in green for the non-crosslinked. The OF ligand is shown in yellow. **(b)** Transition from IF (left) to OF (right) for V978C. Ligand molecules for the IF conformation are shown in red (crosslinked) and green (non-crosslinked). The OF ligand is shown in plum. **(c)** Comparison between OF335 structures with 1 or 2 ligands bound. The ligands from OF335-1lig, with one molecule bound, is shown in yellow. The two molecules bound in structure OF335-2lig are shown in light pink (crosslinked) or light green (non-crosslinked). **(d)** Comparison of the ligand position between OF335 and OF978. The ligand for OF335 is shown in yellow, for OF978 in plum. **(e)** Overlay of the ligand in the three IF structures. The ligand in IF335 is shown in yellow, in IF978 in plum, and in IF971 in light green. The crosslinked residues L335C and V978C are shown in cyan. An overall alignment in ChimeraX was used for the comparison.




**Extended Data Fig. 10.**

**Example of cryo-EM data processing pathway for Pgp.**

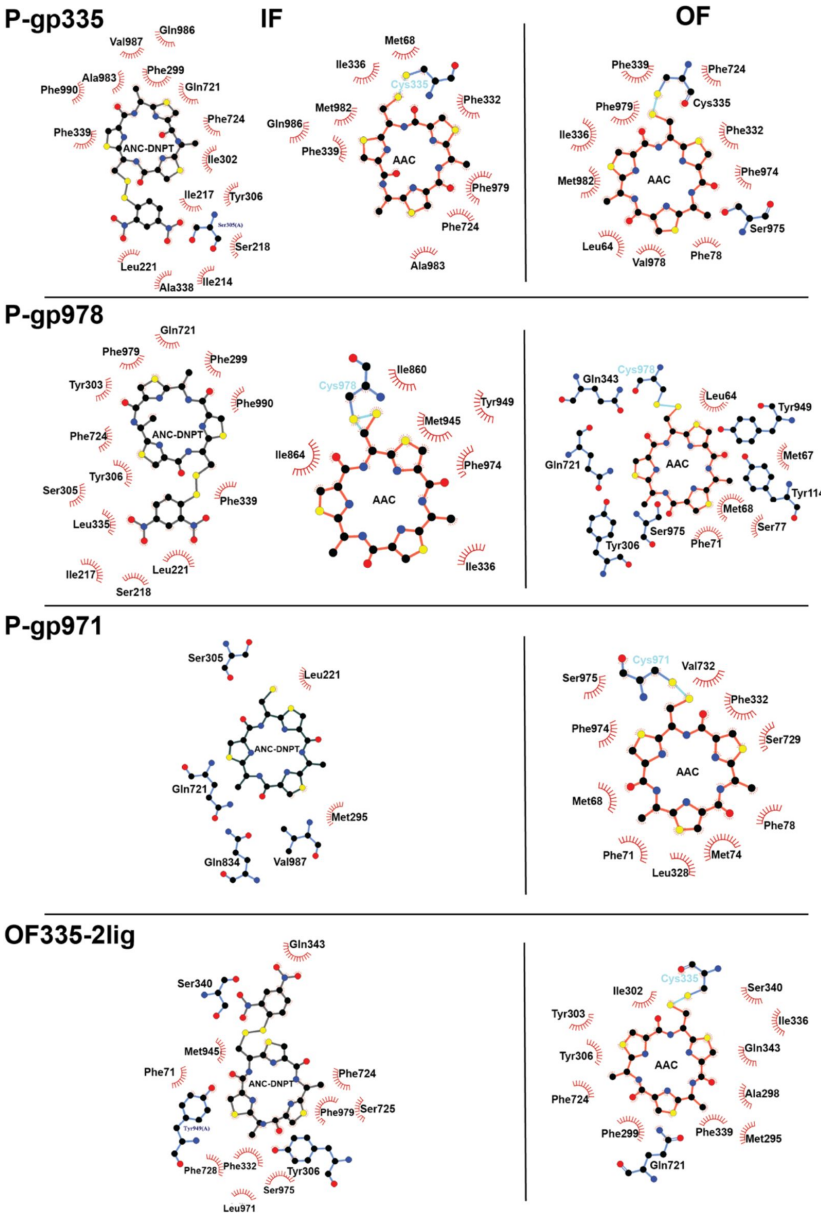
The cryo-EM data for OF335 were collected on a Titan Krios G3i with a Gatan Bioquantum K3. In total, 18394 micrographs were collected. The entire processing pipeline was performed in cryoSPARC and cryoSPARC live. All data sets were processed in a similar manner. The scale bar is 50 nm on the micrographs and 20 nm on the inset.




**Extended Data Fig. 11.**

**Visualization of tunnels in Pgp by the CAVER plugin for PyMOL[70].**

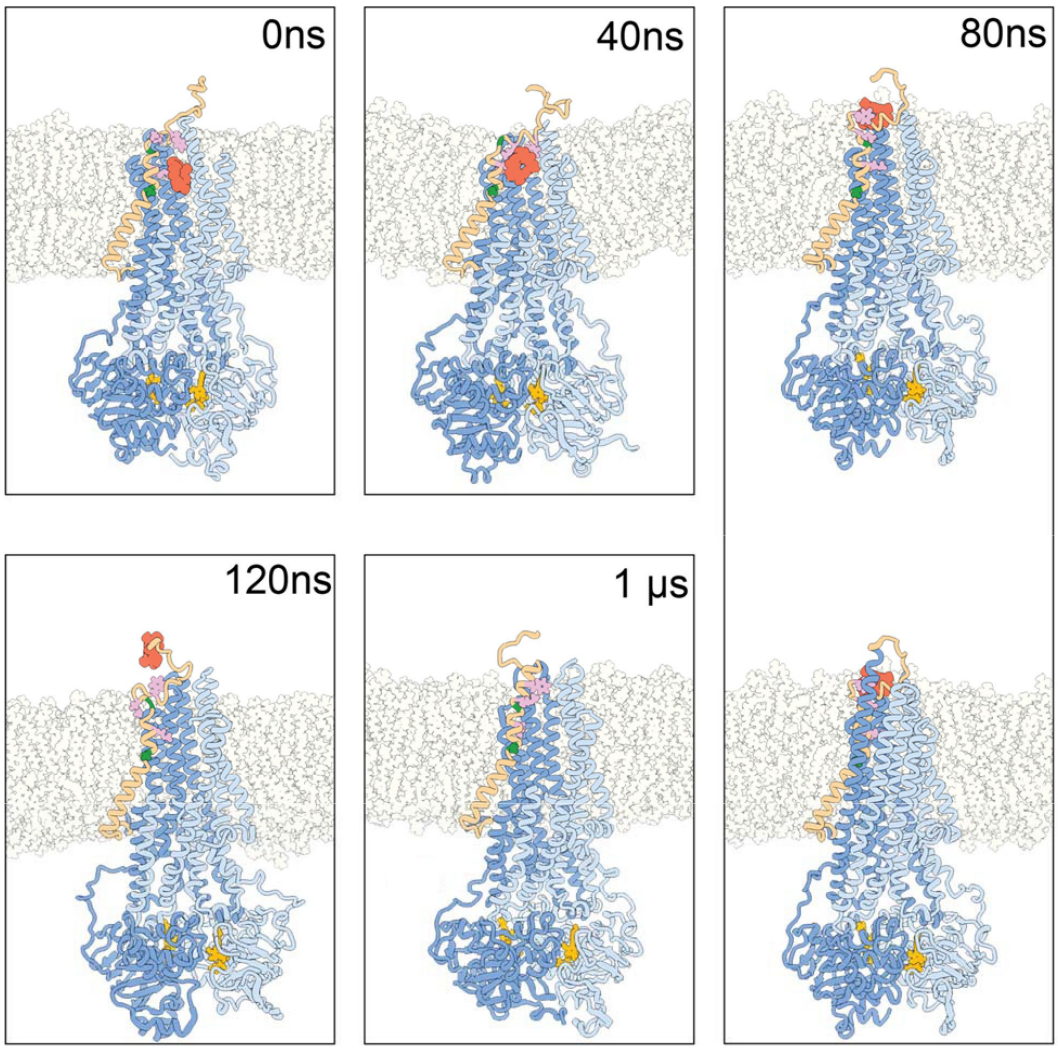
The standard value of 3 Å was chosen as the minimum probe diameter. With OF335-nolig as template, it was not possible to calculate a tunnel. The protein is shown as cartoon and individual tunnels colored as green spheres.




## Extended Data Fig. 12.

### Residues in Pgp mutants interacting with AAC and AAC-DNPT.

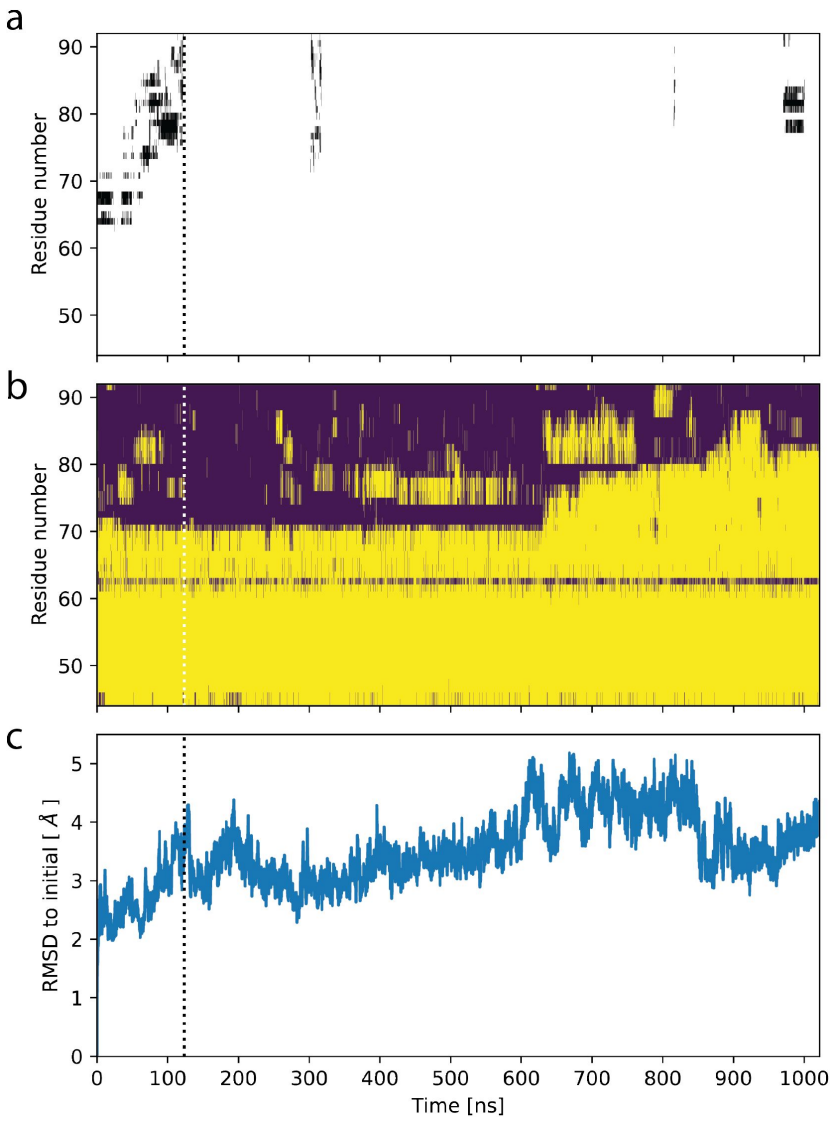
Hydrophobic interactions are shown as red halos, amino acids for hydrogen bonds are shown in ball-and-stick representation. LigPlot+ was used for calculating and visualization of these interactions[71].




**Extended Data Fig. 13.**

**Snapshots for MD simulation started from OF978-1lig at 400K.**

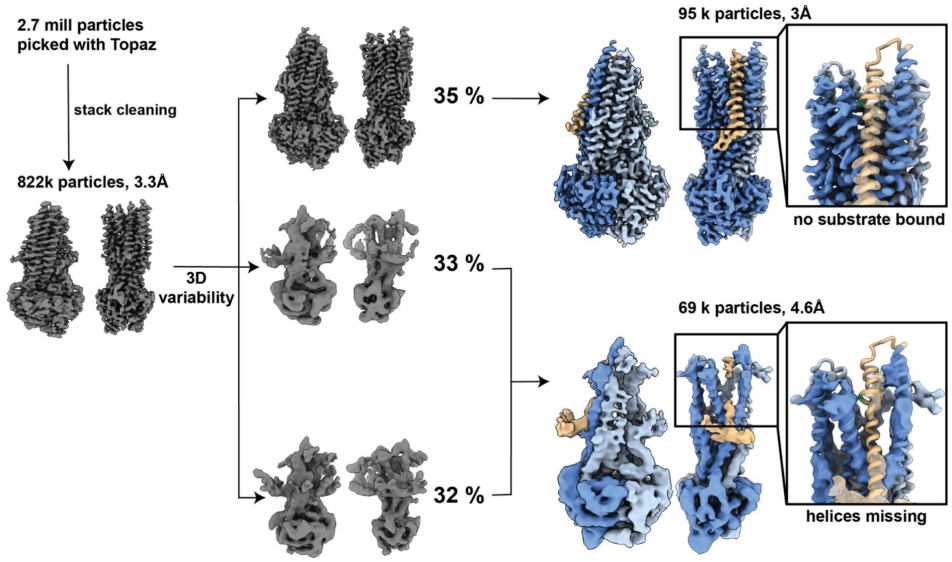
TM1 is shown in brown, the ligand is shown in red spheres. A63 and G72 are shown in green; M68,M74 and P78 are shown in plum. The plasma membrane is shown with grey outlines. Water is not shown. **(0 ns)** Cut through the initial configuration, OF978-1lig. **(40 ns)** TM1 has rearranged while interacting with the ligand from the top. **(80 ns)** The ligand is leaving the transporter while interacting with TM1 from the side. The top panel shows a cut through the structure, the bottom panel shows the full structure. **(120 ns)** The ligand has left the transporter and has its last transient interactions with the top of TM1. **(1us)** the end of the simulation shows a refolding of TM1.




**Extended Data Fig. 14.**

**Ligand release in MD simulation starting from OF978-1lig.**

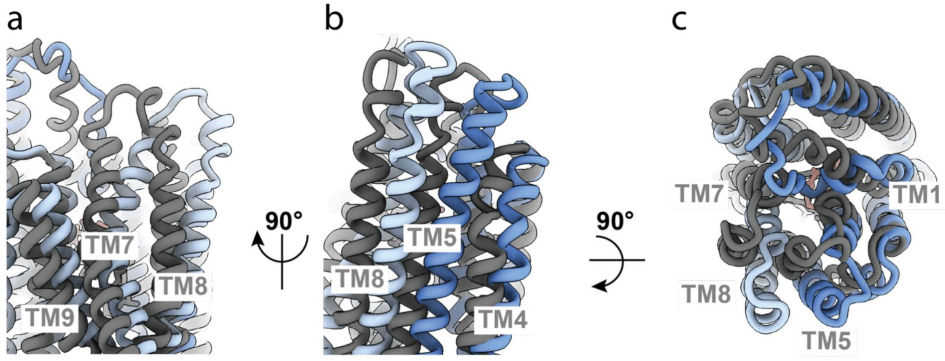
**(a)** Contacts of the ligand with residues in TM1 as function of time. Black bars indicate residues within 5 Å of the ligand. **(b)** Secondary structure in TM1 as function of time. Loop regions are shown in purple, alpha helices in yellow. **(c)** Root-mean-square distance (RMSD) of backbone atoms in Pgp transmembrane domain (residues 34-366 and 691-1010) to initial structure as function of time. Dotted lines indicate the time point of ligand release.




## Extended Data Fig. 15

### Processing overview for OF335-G72.

In total, 2.7 million particles were picked using Topaz. Particle stack was cleaned through several rounds of 3D classification. Stack with 822k particles was used for 3D variability and separated into 3 clusters. Clusters were further processed with iterative rounds of 3D classifications and refinements and revealed 2 distinct conformations in the data set. An apo structure with 3Å resolution with no AAC bound and a low resolution structure of 4.6Å with poor resolution in the transmembrane helices and TM1,2,3,4,5,6,10 and 12 partially missing. Magnified images: atomic model OF335-nolig fit into the structures. TM1 is shown in brown, G72A is marked in green.




## Extended Data Fig. 16.

### Comparison of OF335 with one AAC bound (blue) compared to two AAC molecules bound (grey).

**(a)** Conformational differences between TM7, TM8, and TM9. The tops of helices TM7 and TM8 point in opposite directions in the two structures. **(b)** Conformational differences in TM8, TM5, and TM4. All three helices are shifted significantly with one AAC bound compared to two AAC molecules bound. **(c)** Top view comparing the two structures. The top ends of the helices are all turned further inwards when two substrate molecules are bound compared to one bound molecule. An overall alignment in ChimeraX-1.3 was used for the comparison.

| | OF335-apo (EMD-15687) (PDB 8AVY) | OF335-nolig (EMD-14754) (PDB 7ZK4) | OF335-1lig (EMD- 14755) (PDB 7ZK5) | OF335-2lig (EMD- 14756) (PDB 7ZK6) | IF335-2lig (EMD-17630) (PDB 8PEE) |
|---|---|---|---|---|---|
| **Data collection** | | | | | |
| Microscope | Titan Krios G3i | Titan Krios G3i | Titan Krios G3i | Titan Krios G3i | Titan Krios G3i |
| Energy filter and camera | Gatan BioQuantum K3 | Gatan BioQuantum K3 | Gatan BioQuantum K3 | Gatan BioQuantum K3 | Gatan BioQuantum K3 |
| Energy filter slit width | 30 eV | 30 eV | 30 eV | 30 eV | 30 eV |
| Voltage (kV) | 300 | 300 | 300 | 300 | 300 |
| Nominal magnification | 105,000 | 105,000 | 105,000 | 105,000 | 105,000 |
| Pixel size (Å) | 0.837 | 0.837 | 0.837 | 0.837 | 0.837 |
| Electron exposure ($e^{-}$/Å$^2$) | 75 | 75 | 75 | 75 | 75 |
| Total exposure time (s) | 4 | 4 | 4 | 4 | 4 |
| Number of frames per image | 50 | 50 | 50 | 50 | 50 |
| Total number of images | 6594 | 18394 | 18394 | 18394 | 13999 |
| Defocus range (μm) | -1.2 to -2.5 | -1 to -2.5 | -1 to -2.5 | -1 to -2.5 | -1 to -2.5 |
| **Image processing** | | | | | |
| Processing software | cryoSPARC | cryoSPARC | cryoSPARC | cryoSPARC | RELION 3.1 |
| Motion correction software | cryoSPARC | cryoSPARC | cryoSPARC | cryoSPARC | RELION3.1 |
| CTF estimation software | cryoSPARC | cryoSPARC | cryoSPARC | cryoSPARC | Gctf |
| Particle selection software | Topaz | Topaz | Topaz | Topaz | RELION 3.1 |
| Initial particle images (no.) | 2,262,305 | 9,065,013 | 9,065,013 | 9,065,013 | 2,975,426 |
| Final particle images (no.) | 790,159 | 1,437,110 | 1,184,253 | 939,924 | 218,872 |
| Final refinement software | cryoSPARC | cryoSPARC | cryoSPARC | cryoSPARC | cryoSPARC |
| Symmetry imposed | C1 | C1 | C1 | C1 | C1 |
| Map resolution (Å) | 2.9 | 2.6 | 2.6 | 3.1 | 3.8 |
| B-factor (Å$^2$) | 138.2 | 119.3 | 118.6 | 154.8 | 155.2 |
| **Refinement statistics** | | | | | |
| Initial model (PDB code) | 6C0V | 6C0V | 6C0V | 6C0V | 4Q9I |
| Modeling software | Coot, PHENIX | Coot, PHENIX | Coot, PHENIX | Coot, PHENIX | Coot, PHENIX |
| Model composition | | | | | |
| Non-hydrogen atoms | 9433 | 9371 | 9436 | 9267 | 9217 |
| Protein residues | 1175 | 1176 | 1177 | 1177 | 1157 |
| Water molecules | 4 | 4 | 33 | 4 | - |
| Ligands | 12 | 9 | 10 | 7 | 7 |
| Mean *B* factors (Å$^2$) | | | | | |
| Protein | 27.15 | 27 | 39 | 59 | 71 |
| Water molecules | - | - | 37 | 30 | - |
| Ligand | 37.03 | 59 | 44 | 50 | 57 |
| R.m.s. deviations | | | | | |
| Bond lengths (Å) | 0.030 | 0.024 | 0.024 | 0.019 | 0.024 |
| Bond angles (°) | 1.265 | 1.04 | 1.16 | 0.91 | 1.040 |
| Validation | | | | | |
| MolProbity score | 1.50 | 1.50 | 1.56 | 1.49 | 2.00 |
| Clashscore | 9.39 | 9.29 | 9.43 | 9.33 | 12.52 |
| Poor rotamers (%) | 0.74 | 0.10 | 0.10 | 0.10 | 0.32 |

**Extended Data Table 1.**

**Cryo-EM data collection, refinement, and validation statistics.**

| | | | | | |
|---|---|---|---|---|---|
| Ramachandran plot | | | | | |
| Favored (%) | 98.04 | 97.95 | 97.70 | 98.12 | 94.19 |
| Allowed (%) | 1.96 | 2.05 | 2.22 | 1.79 | 5.81 |
| Disallowed (%) | 0.00 | 0.00 | 0.08 | 0.09 | 0.00 |

| | OF971-1lig (EMD- 14758) (PDB 7ZK8) | IF971-1lig (EMD- 14759) (PDB 7ZK9) | OF978-1lig (EMD- 14760) (PDB 7ZKA) | IF978-2lig (EMD- 14761) (PDB 7ZKB) |
|---|---|---|---|---|
| **Data collection** | | | | |
| Microscope | Titan Krios G4 | Titan Krios G3i | Titan Krios G4 | Titan Krios G3i |
| Energy filter and camera | Selectris X Falcon 4 | Gatan BioQuantum K3 | Selectris X Falcon 4 | Gatan BioQuantum K3 |
| Energy filter slit width | 10 eV | 30 eV | 10 eV | 30 eV |
| Voltage (kV) | 300 | 300 | 300 | 300 |
| Nominal magnification | 215,000 | 105,000 | 215,000 | 105,000 |
| Pixel size (Å) | 0.537 | 0.837 | 0.537 | 0.837 |
| Electron exposure ($e^-/Å^2$) | 70 | 75 | 70 | 75 |
| Total exposure time (s) | 4 | 4 | 4 | 4 |
| Number of frames per image | 70 | 50 | 70 | 50 |
| Total number of images | 12128 | 6758 | 14998 | 6204 |
| Defocus range (μm) | -1 to -2.5 | -1 to -2.5 | -1 to -2.5 | -1 to -2.5 |
| **Image processing** | | | | |
| Processing software | cryoSPARC | cryoSPARC | cryoSPARC | cryoSPARC |
| Motion correction software | cryoSPARC | cryoSPARC | cryoSPARC | cryoSPARC |
| CTF estimation software | cryoSPARC | cryoSPARC | cryoSPARC | cryoSPARC |
| Particle selection software | Topaz | cryoSPARC | cryoSPARC | cryoSPARC |
| Initial particle images (no.) | 1,644,112 | 1,664,538 | 1,761,489 | 1,744,204 |
| Final particle images (no.) | 298,192 | 443,812 | 271,467 | 343,873 |
| Final refinement software | CryoSPARC | CryoSPARC | CryoSPARC | CryoSPARC |
| Symmetry imposed | C1 | C1 | C1 | C1 |
| Map resolution (Å) | 3.0 | 4.3 | 2.9 | 4.7 |
| B-factor ($Å^2$) | 128.1 | 285 | 116.8 | 347.4 |
| **Refinement statistics** | | | | |
| Initial model (PDB code) | 6C0V | 4Q9I | 6C0V | 4Q9I |
| Modeling software | Coot, PHENIX | Coot, PHENIX | Coot, PHENIX | Coot, PHENIX |
| Model composition | | | | |
| Non-hydrogen atoms | 9331 | 9004 | 9407 | 9035 |
| Protein residues | 1181 | 1157 | 1177 | 1157 |
| Water molecules | 4 | - | 4 | - |
| Ligands | 7 | 1 | 9 | 2 |
| Mean *B* factors ($Å^2$) | | | | |
| Protein | 56 | 181 | 31 | 121 |
| Water molecules | 33 | - | 12 | - |
| Ligand | 55 | 154 | 37 | 71 |
| R.m.s. deviations | | | | |
| Bond lengths (Å) | 0.016 | 0.005 | 0.003 | 0.004 |
| Bond angles (°) | 1.00 | 0.92 | 0.59 | 0.80 |
| Validation | | | | |
| MolProbity score | 1.74 | 2.07 | 1.69 | 2.08 |
| Clashscore | 9.69 | 15.76 | 9.86 | 13.26 |
| Poor rotamers (%) | 0.10 | 0.21 | 0.00 | 0.00 |
| Ramachandran plot | | | | |
| Favored (%) | 96.52 | 94.62 | 97.02 | 93.06 |

**Extended Data Table 1.** (continued)

| | | | | |
|---|---|---|---|---|
| Allowed (%) | 3.48 | 5.29 | 2.90 | 6.94 |
| Disallowed (%) | 0.00 | 0.09 | 0.08 | 0.00 |

| | OF971-1lig (EMD- 14758) (PDB 7ZK8) | IF971-1lig (EMD- 14759) (PDB 7ZK9) | OF978-1lig (EMD- 14760) (PDB 7ZKA) | IF978-2lig (EMD- 14761) (PDB 7ZKB) |
|---|---|---|---|---|
| **Data collection** | | | | |
| Microscope | Titan Krios G4 | Titan Krios G3i | Titan Krios G4 | Titan Krios G3i |
| Energy filter and camera | Selectris X Falcon 4 | Gatan BioQuantum K3 | Selectris X Falcon 4 | Gatan BioQuantum K3 |
| Energy filter slit width | 10 eV | 30 eV | 10 eV | 30 eV |
| Voltage (kV) | 300 | 300 | 300 | 300 |
| Nominal magnification | 215,000 | 105,000 | 215,000 | 105,000 |
| Pixel size (Å) | 0.537 | 0.837 | 0.537 | 0.837 |
| Electron exposure ($e^{-}$/Å$^2$) | 70 | 75 | 70 | 75 |
| Total exposure time (s) | 4 | 4 | 4 | 4 |
| Number of frames per image | 70 | 50 | 70 | 50 |
| Total number of images | 12128 | 6758 | 14998 | 6204 |
| Defocus range (μm) | -1 to -2.5 | -1 to -2.5 | -1 to -2.5 | -1 to -2.5 |
| **Image processing** | | | | |
| Processing software | cryoSPARC | cryoSPARC | cryoSPARC | cryoSPARC |
| Motion correction software | cryoSPARC | cryoSPARC | cryoSPARC | cryoSPARC |
| CTF estimation software | cryoSPARC | cryoSPARC | cryoSPARC | cryoSPARC |
| Particle selection software | Topaz | cryoSPARC | cryoSPARC | cryoSPARC |
| Initial particle images (no.) | 1,644,112 | 1,664,538 | 1,761,489 | 1,744,204 |
| Final particle images (no.) | 298,192 | 443,812 | 271,467 | 343,873 |
| Final refinement software | CryoSPARC | CryoSPARC | CryoSPARC | CryoSPARC |
| Symmetry imposed | C1 | C1 | C1 | C1 |
| Map resolution (Å) | 3.0 | 4.3 | 2.9 | 4.7 |
| B-factor (Å$^2$) | 128.1 | 285 | 116.8 | 347.4 |
| **Refinement statistics** | | | | |
| Initial model (PDB code) | 6C0V | 4Q9I | 6C0V | 4Q9I |
| Modeling software | Coot, PHENIX | Coot, PHENIX | Coot, PHENIX | Coot, PHENIX |
| Model composition | | | | |
| Non-hydrogen atoms | 9331 | 9004 | 9407 | 9035 |
| Protein residues | 1181 | 1157 | 1177 | 1157 |
| Water molecules | 4 | - | 4 | - |
| Ligands | 7 | 1 | 9 | 2 |
| Mean *B* factors (Å$^2$) | | | | |
| Protein | 56 | 181 | 31 | 121 |
| Water molecules | 33 | - | 12 | - |
| Ligand | 55 | 154 | 37 | 71 |
| R.m.s. deviations | | | | |
| Bond lengths (Å) | 0.016 | 0.005 | 0.003 | 0.004 |
| Bond angles (°) | 1.00 | 0.92 | 0.59 | 0.80 |
| Validation | | | | |
| MolProbity score | 1.74 | 2.07 | 1.69 | 2.08 |
| Clashscore | 9.69 | 15.76 | 9.86 | 13.26 |
| Poor rotamers (%) | 0.10 | 0.21 | 0.00 | 0.00 |
| Ramachandran plot | | | | |
| Favored (%) | 96.52 | 94.62 | 97.02 | 93.06 |
| Allowed (%) | 3.48 | 5.29 | 2.90 | 6.94 |
| Disallowed (%) | 0.00 | 0.09 | 0.08 | 0.00 |

**Extended Data Table 1.** (continued)

| Equilibration step | Time [$ps$] | Timestep [$ps$] | Ensemble | Thermostat | Barostat | Force constant backbone $\left[\frac{kJ}{molnm^2}\right]$ | Force constant side chain $\left[\frac{kJ}{molnm^2}\right]$ | Force constant lipids $\left[\frac{kJ}{molnm^2}\right]$ | Force constant dihedrals $\left[\frac{kJ}{molrad^2}\right]$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 125 | 0.001 | NVT | Berendsen | | 4000 | 2000 | 1000 | 1000 |
| 2 | 125 | 0.001 | NVT | Berendsen | | 2000 | 1000 | 400 | 400 |
| 3 | 125 | 0.001 | NPT | Berendsen | Berendsen (semi-isotropic) | 1000 | 500 | 400 | 200 |
| 4 | 500 | 0.002 | NPT | Berendsen | Berendsen (semi-isotropic) | 500 | 200 | 200 | 200 |
| 5 | 500 | 0.002 | NPT | Berendsen | Berendsen (semi-isotropic) | 200 | 50 | 40 | 100 |
| 6 | 500 | 0.002 | NPT | Berendsen | Berendsen (semi-isotropic) | 50 | 0 | 0 | 0 |

## Extended Data Table 2

**Equilibration protocol for MD simulations.**

# References


1 Juliano R. L., Ling V (1976) **A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants** *Biochim. Biophys. Acta* **455**:152–162 https://doi.org/10.1016/0005-2736(76)90160-7

2 Ueda K., et al. (1986) **The mdr1 gene, responsible for multidrug-resistance, codes for P-glycoprotein** *Biochem. Biophys. Res. Commun* **141**:956–962 https://doi.org/10.1016/s0006-291x(86)80136-x

3 Sharom F. J (2011) **The P-glycoprotein multidrug transporter** *Essays Biochem* **50**:161–178 https://doi.org/10.1042/bse0500161

4 Loo T. W., Clarke D. M (2000) **Identification of residues within the drug-binding domain of the human multidrug resistance P-glycoprotein by cysteine-scanning mutagenesis and reaction with dibromobimane** *J. Biol. Chem* **275**:39272–39278 https://doi.org/10.1074/jbc.M007741200

5 Loo T. W., Bartlett M. C., Clarke D. M (2009) **Identification of residues in the drug translocation pathway of the human multidrug resistance P-glycoprotein by arginine mutagenesis** *J. Biol. Chem* **284**:24074–24087 https://doi.org/10.1074/jbc.M109.023267

6 Aller S. G., et al. (2009) **Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding** *Science* **323**:1718–1722 https://doi.org/10.1126/science.1168750

7 Alam A., Kowal J., Broude E., Roninson I., Locher K. P (2019) **Structural insight into substrate and inhibitor discrimination by human P-glycoprotein** *Science* **363**:753–756 https://doi.org/10.1126/science.aav7102

8 Nosol K., et al. (2020) **CryoEM structures reveal distinct mechanisms of inhibition of the human multidrug transporter ABCB1** *Proc. Natl. Acad. Sci. U. S. A* **117**:26245–26253 https://doi.org/10.1073/pnas.2010264117

9 Chufan E. E., Sim H. M., Ambudkar S. V (2015) **Molecular basis of the polyspecificity of P-glycoprotein (ABCB1): recent biochemical and structural studies** *Adv. Cancer Res* **125**:71–96 https://doi.org/10.1016/bs.acr.2014.10.003

10 Thiebaut F., et al. (1987) **Cellular localization of the multidrug-resistance gene product P-glycoprotein in normal human tissues** *Proc. Natl. Acad. Sci. U. S. A* **84**:7735–7738 https://doi.org/10.1073/pnas.84.21.7735

11 Cascorbi I (2011) **P-glycoprotein: tissue distribution, substrates, and functional consequences of genetic variations** *Handb. Exp. Pharmacol* :261–283 https://doi.org/10.1007/978-3-642-14541-4_6

12 Cordon-Cardo C., et al. (1990) **Expression of the multidrug resistance gene product (P-glycoprotein) in human normal and tumor tissues** *J. Histochem. Cytochem* **38**:1277–1287 https://doi.org/10.1177/38.9.1974900

13 Schinkel A. H (1999) **P-Glycoprotein, a gatekeeper in the blood-brain barrier** *Adv. Drug Deliv. Rev* **36**:179–194

14 Borst P., Schinkel A. H (2013) **P-glycoprotein ABCB1: a major player in drug handling by mammals** *J. Clin. Invest* **123**:4131–4133 https://doi.org/10.1172/JCI70430

15 Fromm M. F (2004) **Importance of P-glycoprotein at blood-tissue barriers** *Trends Pharmacol. Sci* **25**:423–429 https://doi.org/10.1016/j.tips.2004.06.002

16 Szakacs G., Paterson J. K., Ludwig J. A., Booth-Genthe C., Gottesman M. M (2006) **Targeting multidrug resistance in cancer** *Nature reviews. Drug Discov* **5**:219–234 https://doi.org/10.1038/nrd1984

17 Gottesman M. M., Ling V (2006) **The molecular basis of multidrug resistance in cancer: the early years of P-glycoprotein research** *FEBS Lett* **580**:998–1009 https://doi.org/10.1016/j.febslet.2005.12.060

18 Giacomini K. M., et al. (2010) **Membrane transporters in drug development** *Nat. Rev. Drug Discov* **9**:215–236 https://doi.org/10.1038/nrd3028

19 Waghray D., Zhang Q (2018) **Inhibit or Evade Multidrug Resistance P-Glycoprotein in Cancer Treatment** *J. Med. Chem* **61**:5108–5121 https://doi.org/10.1021/acs.jmedchem.7b01457

20 Rees D. C., Johnson E., Lewinson O (2009) **ABC transporters: the power to change. Nature reviews** *Mol. Cell Biol* **10**:218–227 https://doi.org/10.1038/nrm2646

21 Hofmann S., et al. (2019) **Conformation space of a heterodimeric ABC exporter under turnover conditions** *Nature* **571**:580–583 https://doi.org/10.1038/s41586-019-1391-0

22 Januliene D., Moeller A (2020) **CryoEM of ABC transporters: an ice-cold solution to everything?** *FEBS Lett* **594**:3776–3789 https://doi.org/10.1002/1873-3468.13989

23 Verhalen B., et al. (2017) **Energy transduction and alternating access of the mammalian ABC transporter P-glycoprotein** *Nature* **543**:738–741 https://doi.org/10.1038/nature21414

24 Szewczyk P., et al. (2015) **Snapshots of ligand entry, malleable binding and induced helical movement in P-glycoprotein** *Acta Crystallogr. Section D: Biol. Crystallogr* **71**:732–741 https://doi.org/10.1107/S1399004715000978

25 Alam A., et al. (2018) **Structure of a zosuquidar and UIC2-bound human-mouse chimeric ABCB1** *Proc. Natl. Acad. Sci. U. S. A* **115**:E1973–E1982 https://doi.org/10.1073/pnas.1717044115

26 Kim Y., Chen J (2018) **Molecular structure of human P-glycoprotein in the ATP-bound, outward-facing conformation** *Science* **359**:915–919 https://doi.org/10.1126/science.aar7389

27 Goddeke H., Schafer L. V (2020) **Capturing substrate translocation in an ABC exporter at the atomic level** *J. Am. Chem. Soc* **142**:12791–12801 https://doi.org/10.1021/jacs.0c05502

28 McCormick J. W., Vogel P. D., Wise J. G (2015) **Multiple drug transport pathways through human P-Glycoprotein** *Biochemistry* **54**:4374–4390 https://doi.org/10.1021/acs.biochem.5b00018

29 Loo T. W., Clarke D. M (2001) **Defining the drug-binding site in the human multidrug resistance P-glycoprotein using a methanethiosulfonate analog of verapamil, MTS-verapamil** *J. Biol. Chem* **276**:14972–14979 https://doi.org/10.1074/jbc.M100407200

30 Urbatsch I. L., Sankaran B., Bhagat S., Senior A. E (1995) **Both P-glycoprotein nucleotide-binding sites are catalytically active** *J. Biol. Chem* **270**:26956–26961 https://doi.org/10.1074/jbc.270.45.26956

31 Moeller A., et al. (2015) **Distinct Conformational spectrum of homologous multidrug ABC transporters** *Structure* **23**:450–460 https://doi.org/10.1016/j.str.2014.12.013

32 Parveen Z., et al. (2011) **Molecular dissection of dual pseudosymmetric solute translocation pathways in human P-glycoprotein** *Mol. Pharmacol* **79**:443–452 https://doi.org/10.1124/mol.110.067611

33 Lopez-Llano J., Campos L. A., Sancho J (2006) **Alpha-helix stabilization by alanine relative to glycine: roles of polar and apolar solvent exposures and of backbone entropy** *Proteins* **64**:769–778 https://doi.org/10.1002/prot.21041

34 Ward A., Reyes C. L., Yu J., Roth C. B., Chang G (2007) **Flexibility in the ABC transporter MsbA: Alternating access with a twist** *Proc. Natl. Acad. Sci. U. S. A* **104**:19005–19010 https://doi.org/10.1073/pnas.0709388104

35 Dawson R. J., Locher K. P (2006) **Structure of a bacterial multidrug ABC transporter** *Nature* **443**:180–185 https://doi.org/10.1038/nature05155

36 Fan C., Kaiser J. T., Rees D. C (2020) **A structural framework for unidirectional transport by a bacterial ABC exporter** *Proc. Natl. Acad. Sci. U. S. A* **117**:19228–19236 https://doi.org/10.1073/pnas.2006526117

37 Chaptal V., et al. (2022) **Substrate-bound and substrate-free outward-facing structures of a multidrug ABC exporter** *Sci Adv* **8** https://doi.org/10.1126/sciadv.abg9215

38 Swartz D. J., et al. (2014) **Directed evolution of P-glycoprotein cysteines reveals site-specific, non-conservative substitutions that preserve multidrug resistance** *Biosci. Rep* **34**:297–310 https://doi.org/10.1042/BSR20140062

39 Bai J., et al. (2011) **A gene optimization strategy that enhances production of fully functional P-glycoprotein in Pichia pastoris** *PloS one* **6** https://doi.org/10.1371/journal.pone.0022577

40 Zoghbi M. E., et al. (2017) **Substrate-induced conformational changes in the nucleotide-binding domains of lipid bilayer-associated P-glycoprotein during ATP hydrolysis** *J. Biol. Chem* **292**:20412–20424 https://doi.org/10.1074/jbc.M117.814186

41 Urbatsch I. L., Sankaran B., Weber J., Senior A. E (1995) **P-glycoprotein is stably inhibited by vanadate-induced trapping of nucleotide at a single catalytic site** *J. Biol. Chem* **270**:19383–19390

42 Urbatsch I. L., Tyndall G. A., Tombline G., Senior A. E (2003) **P-glycoprotein catalytic mechanism: studies of the ADP-vanadate inhibited state** *J. Biol. Chem* **278**:23171–23179 https://doi.org/10.1074/jbc.M301957200

43 Brouwer K. L., et al. (2013) **In vitro methods to support transporter evaluation in drug discovery and development** *Clin. Pharmacol. Ther* **94**:95–112 https://doi.org/10.1038/clpt.2013.81

44 Swartz D. J., et al. (2014) **Directed evolution of P-glycoprotein cysteines reveals site-specific, non-conservative substitutions that preserve multidrug resistance** *Biosci. Rep* **34** https://doi.org/10.1042/BSR20140062

45 Gewering T., Januliene D., Ries A. B., Moeller A (2018) **Know your detergents: A case study on detergent background in negative stain electron microscopy** *J. Struct. Biol* **203**:242–246 https://doi.org/10.1016/j.jsb.2018.05.008

46 Januliene D., Moeller A (2021) **Single-particle cryoEM of membrane proteins** *Methods Mol. Biol* **2302**:153–178 https://doi.org/10.1007/978-1-0716-1394-8_9

47 Punjani A., Rubinstein J. L., Fleet D. J., Brubaker M. A (2017) **cryoSPARC: algorithms for rapid unsupervised cryoEM structure determination** *Nat. Methods* **14**:290–296 https://doi.org/10.1038/nmeth.4169

48 Scheres S. H (2012) **RELION: implementation of a Bayesian approach to cryoEM structure determination** *J. Struct. Biol* **180**:519–530 https://doi.org/10.1016/j.jsb.2012.09.006

49 Punjani A., Zhang H., Fleet D. J (2020) **Non-uniform refinement: adaptive regularization improves single-particle cryoEM reconstruction** *Nat. Methods* **17**:1214–1221 https://doi.org/10.1038/s41592-020-00990-8

50 Rubinstein J. L., Brubaker M. A (2015) **Alignment of cryoEM movies of individual particles by optimization of image translations** *J. Struct. Biol* **192**:188–195 https://doi.org/10.1016/j.jsb.2015.08.007

51 Bepler T., Kelley K., Noble A. J., Berger B (2020) **Topaz-Denoise: general deep denoising models for cryoEM and cryoET** *Nat. Commun* **11** https://doi.org/10.1038/s41467-020-18952-1

52 Terwilliger T. C., Ludtke S. J., Read R. J., Adams P. D., Afonine P. V (2020) **Improvement of cryoEM maps by density modification** *Nat. Methods* **17**:923–927 https://doi.org/10.1038/s41592-020-0914-9

53 Emsley P., Cowtan K (2004) **Coot: model-building tools for molecular graphics** *Acta Crystallogr. Section D: Biol. Crystallogr* **60**:2126–2132 https://doi.org/10.1107/S0907444904019158

54 Afonine P. V., et al. (2018) **Real-space refinement in PHENIX for cryoEM and crystallography** *Acta Crystallogr. D Struct. Biol* **74**:531–544 https://doi.org/10.1107/S2059798318006551

55 Chen V. B., et al. (2010) **MolProbity: all-atom structure validation for macromolecular crystallography** *Acta Crystallogr. Section D: Biol. Crystallogr* **66**:12–21 https://doi.org/10.1107/S0907444909042073

56 Goddard T. D., et al. (2018) **UCSF ChimeraX: Meeting modern challenges in visualization and analysis** *Protein Sci* **27**:14–25 https://doi.org/10.1002/pro.3235

57 Wu E. L., et al. (2014) **CHARMM-GUI Membrane Builder toward realistic biological membrane simulations** *J. Comput. Chem* **35**:1997–2004 https://doi.org/10.1002/jcc.23702

58 Lee J., et al. (2016) **CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field** *J. Chem. Theory Comput* **12**:405–413 https://doi.org/10.1021/acs.jctc.5b00935

59 Sondergaard C. R., Olsson M. H., Rostkowski M., Jensen J. H (2011) **Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pKa values** *J. Chem. Theory Comput* **7**:2284–2295 https://doi.org/10.1021/ct200133y

60 Olsson M. H., Sondergaard C. R., Rostkowski M., Jensen J. H (2011) **PROPKA3: Consistent treatment of internal and surface residues in empirical pKa predictions** *J. Chem. Theory Comput* **7**:525–537 https://doi.org/10.1021/ct100578z

61 Vanommeslaeghe K., MacKerell A. D. (2012) **Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing** *J. Chem. Inf. Model* **52**:3144–3154 https://doi.org/10.1021/ci300363c

62 Vanommeslaeghe K., Raman E. P., MacKerell A. D. (2012) **Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges** *J. Chem. Inf. Model* **52**:3155–3168 https://doi.org/10.1021/ci3003649

63 Huang J., et al. (2017) **CHARMM36m: an improved force field for folded and intrinsically disordered proteins** *Nat. Methods* **14**:71–73

64 Vanommeslaeghe K., et al. (2010) **CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields** *J. Comput. Chem* **31**:671–690 https://doi.org/10.1002/jcc.21367

65 Pall S., et al. (2020) **Heterogeneous parallelization and acceleration of molecular dynamics simulations in GROMACS** *J. Chem. Phys* **153** https://doi.org/10.1063/5.0018516

66 Abraham M. J., et al. (2015) **GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers** *SoftwareX* **1**:26–31

67 Hess B., Bekker H., Berendsen H. J. C., Fraaije J. G. E. M (1997) **LINCS: A linear constraint solver for molecular simulations** *J. Comput. Chem* **18**:1463–1472

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

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

70. Chovancova E., et al. (2012) **CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures** *PLoS Comput. Biol* **8**

71. Wallace A. C., Laskowski R. A., Thornton J. M. (1995) **LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions** *Protein Eng* **8**:127–134

## Article and author information

### Theresa Gewering

Osnabrück University, Department of Biology/Chemistry, Structural Biology section, 49076 Osnabrück, Germany, Department of Structural Biology, Max Planck Institute of Biophysics, 60438 Frankfurt am Main, Germany
ORCID iD: 0000-0002-4896-6256

**Deepali Waghray**
Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, USA

**Kristian Parey**
Osnabrück University, Department of Biology/Chemistry, Structural Biology section, 49076 Osnabrück, Germany, Department of Structural Biology, Max Planck Institute of Biophysics, 60438 Frankfurt am Main, Germany, Osnabrück University, Center of Cellular Nanoanalytic Osnabrück (CellNanOs), 49076 Osnabrück, Germany
ORCID iD: 0000-0002-4842-6479

**Hendrik Jung**
Department of Theoretical Biophysics, Max Planck Institute of Biophysics, 60438 Frankfurt am Main, Germany

**Nghi N.B. Tran**
Department of Cell Biology and Biochemistry, Texas Tech University Health Sciences Center, Lubbock, TX 79430, USA

**Joel Zapata**
Department of Cell Biology and Biochemistry, Texas Tech University Health Sciences Center, Lubbock, TX 79430, USA

**Pengyi Zhao**
Department of Chemistry & Environmental Science, New Jersey Institute of Technology, Newark, New Jersey 07102, USA

**Hao Chen**
Department of Chemistry & Environmental Science, New Jersey Institute of Technology, Newark, New Jersey 07102, USA

**Dovile Januliene**
Osnabrück University, Department of Biology/Chemistry, Structural Biology section, 49076 Osnabrück, Germany, Department of Structural Biology, Max Planck Institute of Biophysics, 60438 Frankfurt am Main, Germany, Osnabrück University, Center of Cellular Nanoanalytic Osnabrück (CellNanOs), 49076 Osnabrück, Germany
ORCID iD: 0000-0002-3279-7590

**Gerhard Hummer**
Department of Theoretical Biophysics, Max Planck Institute of Biophysics, 60438 Frankfurt am Main, Germany, Institute for Biophysics, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany
ORCID iD: 0000-0001-7768-746X

**Ina L. Urbatsch**
Department of Cell Biology and Biochemistry, Texas Tech University Health Sciences Center, Lubbock, TX 79430, USA

**Arne Moeller**
Osnabrück University, Department of Biology/Chemistry, Structural Biology section, 49076 Osnabrück, Germany, Department of Structural Biology, Max Planck Institute of Biophysics, 60438 Frankfurt am Main, Germany, Osnabrück University, Center of Cellular Nanoanalytic Osnabrück (CellNanOs), 49076 Osnabrück, Germany
**For correspondence:** arne.moeller@uos.de
ORCID iD: 0000-0003-1101-5366

**Qinghai Zhang**
Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, USA
**For correspondence:** qinghai@scripps.edu

## Copyright



## Editors

Reviewing Editor
**David Drew**
Stockholm University, Stockholm, Sweden

Senior Editor
**Merritt Maduke**
Stanford University, Stanford, United States of America

**Reviewer #1 (Public Review):**

Summary:
Here the authors have tethered a Pgp substrate to strategically place cysteine residues in the protein. Notably, the cysteine-linked substrate (ANC-DNPT)- stimulates ATP hydrolyse and so is able to undergo IF to OF transitions. The authors then determined cryo-EM structures of these complexes and MD simulations of bound states. By capturing unforeseen OF conformations with substate they propose that TM1 undergoes local conformational changes that are sufficient to translocate substrates, rather than large bundle movements.

Strengths:
This paper provides the first substrate (ANC-DNPT)- bound conformations of PgP and a new mechanistic model of how substrates are translocated.

Weaknesses:
Although the cross-links stimulate ATP hydrolysis, further controls are needed to convince me that the TM1 conformations observed in the structures are physiologically relevant, since they have been trapped by "large" substrates covalently-tethered by cross-links.

- https://doi.org/10.7554/eLife.90174.1.sa2

**Reviewer #2 (Public Review):**

Summary:
The manuscript by Gewering and coworkers is an elegant mechanistic investigation of the mammalian multidrug transporter Pgp. I will not elaborate on the significance of this protein except to point out its clinical involvement in cancer resistance to chemotherapy.

Strengths:
The strengths of the investigation are partly in the combination of sophisticated chemical synthesis, state-of-the-art cryoEM in a well-established biochemical context. What is more exciting is the tackling of a long-standing question in the field: namely how do drugs make their way through the structure to be exported across the membrane? Unfortunately, the field has been stuck in hand waving model based on structures that in the outward-facing conformations are devoid of substrates. The work challenges the dogma that emerged from this hand-waving model and presents an alternative model that appears to be supported by the data.

Weaknesses:
There is much to like about the experimental work here but I am less sanguine on the interpretation. The main idea is to covalently link via disulfide bonds a model tripeptide substrate under different conditions that mimic transport and then image the resulting conformations. The choice of the Pgp cysteine mutants here is critical but also poses questions regarding the interpretation. What seems to be missing, or not reported, is a series of control experiments for further cysteine mutations.

- https://doi.org/10.7554/eLife.90174.1.sa1

**Reviewer #3 (Public Review):**

Summary:
The authors employed a new strategy, covalent substrate-labeling, to address the open issue of the substrate transport mechanism by single particle cryogenic-EM. A cyclic peptide (QZ-Ala), which was already used in the past as a substrate for structural purposes, was modified and covalently attached to ABCB1 at strategic positions in the transmembrane domain via Cys-specific coupling chemistry. Overall, four mutations (two per TMD) were generated and functionally analyzed. These residues are in proximity to the QZ-Ala binding site and are labeled by verapamil. Interestingly, two mutants could only be labeled if ATP and $Mg^{2+}$ were present.

Strengths:
Three of the four mutants were structurally analyzed by single particle cryo-EM with structures in both, the inward- and outward-facing conformation and overall resolutions ranging from 2.6 to 4.3 Å. Applying multi-model analyses allowed for the extraction of additional structures from one data set. These structures formed the basis for a detailed analysis of the substrate translocation pathway. This enabled the researcher to compare the IF and OF states of the same mutant and same substrate, which is pivotal for their conclusions. The mutations 335/978 trap the substrate at different points of the translocation pathway, while 971 located two helical turns away from the first set, trapped the system at a later stage. The described strategy revealed a cascade of conformational changes during substrate transport which focus on TMH1, which is straight in the IF state, but swings out in the OF state. Pivotal for such a change is G72. This residue was mutated to Ala and also structurally analyzed. These structures were supported by MD simulations and functional data. Thus, the new prosed kinking and straightening mechanisms are different from the so far accepted wide-open OF state observed in bacterial transporters. This clearly suggests a different mechanism for ABCB1.

Weaknesses:
I have a couple of minor issues that I have listed in the section recommendations for authors Overall, the manuscript is very well written, sheds new light on the molecular mechanism of substrate translocation by ABCB1, and might even provide a new starting point for inhibitor design. I know that it is unusual, but I like the manuscript in its current version and recommend acceptance in its current form.

- https://doi.org/10.7554/eLife.90174.1.sa0

**Author Response**

> ***Reviewer 1 (Public Review):***
>
> *Weakness: Although the cross-links stimulate ATP hydrolysis, further controls are needed to convince me that the TM1 conformations observed in the structures are physiologically relevant, since they have been trapped by "large" substrates covalently-tethered by crosslinks.*

Reviewer 1 raised concerns about the relatively large size of our covalently attached AAC substrate that would potentially distort TM1 in Pgp. We would like to clarify that AAC has a molecular weight of 462 Da, which, in comparison to many known Pgp substrates ranging from 250 to over 1,000 Da, is not a large compound. For instance, the few other Pgp substrates mentioned in our manuscript all have a comparable or larger size: verapamil, 455 Da; doxorubicin, 544 Da; FK506, 804 Da; valinomycin, 1,111 Da; cyclosporin A, 1,203 Da.

Furthermore, AAC was strategically attached to a site distant from TM1 in the inwardfacing Pgp conformation. After it was exported to the outward-facing state, several TM helices accommodate the compound. The observation that only TM1 exhibited significant conformational changes suggests its potential role in the transport mechanism. This hypothesis is supported by our findings, where a conservative substitution (G72A) in TM1 resulted in a dramatic loss of transport function for various drug substrates and impaired verapamil-stimulated ATPase activity.

> ***Reviewer 1 (Recommendations for the Authors):***
>
> *I understand the need for an unconventional approach to understanding the translocation pathway. What would help to support this model is to cross-link a much smaller substrate, as the one used is quite large and could potentially distort TM1 in the outward-state when cross-linked.*

We thank the reviewer for this recommendation, and we have outlined plans for future experiments involving other substrates, including smaller ones, to further investigate our proposed model. However, it is important to acknowledge that conducting these studies will require a significant amount of effort and resources, which we believe extend beyond the scope of our current manuscript.

> *In unbiased MD simulations starting from the IF state are there any simulations where the substrate follows the same path as proposed here?*

All our MD simulations were performed in the outward-facing state to focus on potential substrate release pathways. Starting MD simulations from the inwardfacing state would introduce complexities in capturing the necessary domain motions and nucleotide binding and hydrolysis required for substrate translocations. Therefore, we opted not to perform MD studies starting from the inward-facing state.

> ***Reviewer 2 (Public Review):***
>
> *Weakness: There is much to like about the experimental work here but I am less sanguine on the interpretation. The main idea is to covalently link via disulfide bonds a model tripeptide substrate under different conditions that mimic transport and then image the resulting conformations. The choice of the Pgp cysteine mutants here is critical but also poses questions regarding the interpretation. What seems to be missing, or not reported, is a series of control experiments for further cysteine mutations.*

Reviewer 2 raised concerns about the interpretation of our results and suggested the need for additional mutant designs to validate our proposed TM1 mechanism. Firstly, we believe that the observed TM1 conformational changes are valid in our cryoEM structures, despite the use of different conditions and several mutants to capture Pgp in the outward-facing state.

Regarding the G72A mutant, we consider it conclusive that this single point mutation in the TM1 has a profound effect. Importantly, the G72A mutant was readily expressed and purifiable as a stable protein. We were able to resolve a high-resolution structure of the G72A mutant (without the substrate), confirming that the protein is not generally destabilized but properly folded.

Above all, we appreciate the Reviewer's suggestion to explore additional mutations and intend to do so in future studies.

> ***Reviewer 2 (Recommendations for the Authors):***
>
> *I am sold on the results regarding TM1 conformational changes as they are evident in the cryoEM structures. However, the set of states compared between mutants are not biochemically equivalent: for 335 and 978 they used an ATP-impaired Pgp whereas for 971 they used what appears to be WT, and the conformation was imaged presumably subsequent to ATP hydrolysis and Vanadate trapping. This is significant if the authors were unable to trap the OF in the impaired mutant background and should be highlighted. I have to believe that they tried that condition but I could be wrong.*

We acknowledge the point made by the Reviewer about the biochemical equivalence of mutant states and the potential significance of using an ATP-impaired mutant for trapping the outward-facing conformation of 971. We have not yet attempted to use the ATPase-deficient 971C mutant for crosslinking and intend to address this question in future studies.

In our current approach, we used the ATPase-active 971C for two specific reasons:

1. Our biochemistry data, as shown in Fig 1C, indicates that 971C only crosslinks in the presence of ATP hydrolysis conditions. Vanadate trapping was employed to stabilize the outward-facing conformation.
2. Based on our experience, we have observed that the conformations of ATP-bound (mutant) and vanadate-trapped states of an ABC transporter are structurally equivalent at this resolution level of our study (see ref. 21: Hoffmann et al. NATURE 2019).

> *The authors propose a new model for substrate translocation. It is based on three mutants and a number of structures. If the authors were not challenging the current dogma I would not have written the next comment. Considering the impact of the findings, I would have designed a couple more cysteine mutants based on their model. For instance, this pathway has a number of stabilizing interactions, can't they make a mutant that preserves conformational switching but eliminates substrate translocation? I*

> *like the G97A mutant result but I am worried that the effect could just be a general destabilization or misfolding as part of the cryoEM particles seem to suggest. The authors advance one interpretation of the disorder observed in this mutant but it could easily be my interpretation.*

We thank the reviewer for the suggestion to design additional mutants to further validate our proposed model for substrate translocation. We agree that this would be highly valuable, considering the potential impact of our findings. However, given the time-intensive nature of our approach, we believe that presenting these additional designs in a future study is a reasonable course of action.

Regarding the G72A mutation, we believe that our current data fully supports our model and the role of TM1 in regulating the Pgp activity. Importantly, we would like to emphasize that the G72A mutant was readily expressed and purifiable as a stable protein. Additionally, our cryoEM structural determination of the G72A mutant at high resolution confirmed that the protein is not generally destabilized but properly folded.

> *There are a couple of troubling methodological questions that I want the authors to address or clarify:*
>
> *1- In the methods they report that the final sample for cryoEM was prepared on a SEC devoid of detergent. It is obvious that the sample was folded but I was wondering why the detergent was removed? Was that critical for observing these structures with multiple ligands? Did they observe any lipids in their cryoEM?*

We avoid detergent in the buffer on final SEC purification. This step is to remove free detergent from the background which helps during cryoEM imaging. Of course, this cannot be done with every detergent but due to the very low CMC of LMNG it is possible. By now, we have verified this method for several other transporters with the same success. While this procedure helps us to obtain better images it is not necessary to obtain specific conformations or ligand bound states, nor does it affect these states or conformations.

In our cryoEM structures , we did observe multiple cholesterol hemisuccinate (CHS) molecules on the outer transmembrane surface of Pgp.

> *2- Can the authors comment on why labeling was carried out in the presence of ATP? Does it matter if the substrate was added prior to ATP and incubated for a few minutes?*

For every dataset, we first added the substrate to be cross-linked and afterwards added the ATP. In the cases of 335C and 978C, labeling was successful before ATP was added, as evidenced by the inward-facing structures with cross-linked substrate.

However, for 971C, cross-linking only occurred after the addition of ATP. We interpret this data to suggest that the 971 site is inaccessible to the substrate in the inward-facing state, and cross-linking can only occur after the transporter transitions to outward-facing state. This is in line with our inward-facing structure which does not show a cross-linked substrate, and our biochemical data shown in Fig 1C, where 971C only crosslinked in the presence of ATP.

> *3- I am not an expert on MD simulations and I understand that carrying out simulations at higher temperatures used to be a trick to accelerate the process. Is this still necessary? Why didn't the author use approaches such as WESTPA?*

Most so-called enhanced sampling methods, including WESTPA, explicitly define a reaction coordinate for the process of interest, usually based on intuition or prior studies. If this coordinate is chosen poorly, enhanced sampling usually fails, either because the sampling

becomes inefficient or because the sampling biases the transition pathway (or both). Lacking reliable intuition or prior knowledge on which motions would result in substrate release, we chose temperature to speed up the process. High temperature largely avoids the introduction of an any bias through the definition of a progress coordinate. By contrast, the weighted ensemble method underlying WESTPA is a great method to simulate unbiased dynamics of a process with a known progress coordinate, but unfortunately requires to choose a progress coordinate prior to the simulation and will then mostly sample the process along this progress coordinate, because this is the only direction in which sampling is improved. High temperature MD on the other hand accelerates all processes in the system under study. Indeed, we have now confirmed that the pathway found at high temperature is also feasible at near-ambient conditions.

In new simulations, we have now observed a similar release pathway at T=330 K. As the only difference, the substrate has not fully dissociated from the protein after 2.5 us, with weak interactions persisting at the top part of TM1 from the extracellular side. Importantly, this is a configuration observed also in higher temperature simulations but with much shorter lifetime.

In response, we will include these new findings in the revised manuscript.

> *4- One way to show that the two substrates binding mode is biochemically relevant is to measure Vmax at different substrate concentrations. One would expect a cooperative transition if that interaction is mechanistically important.*

We have measured Vmax as a function of QZ-Ala concentration in a previous report (ref. 24), supporting positive cooperativity for binding to two sites.

We thank Reviewer 3 for recommending the acceptance of our manuscript as is. We will address all minor comments from Reviewer 3 in the revised manuscript.