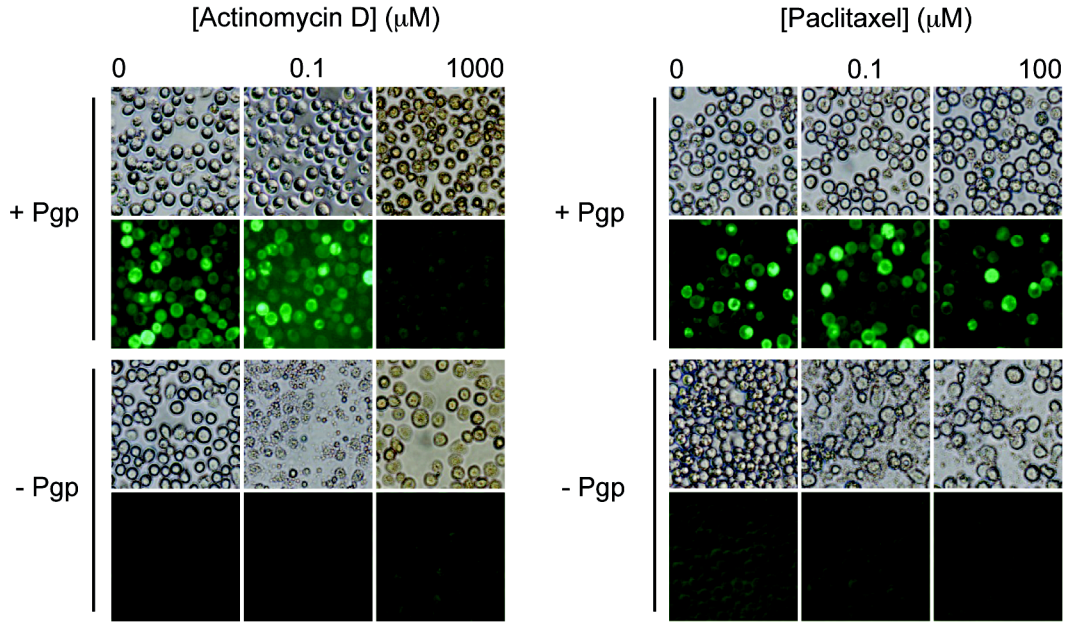
Supplementary Information

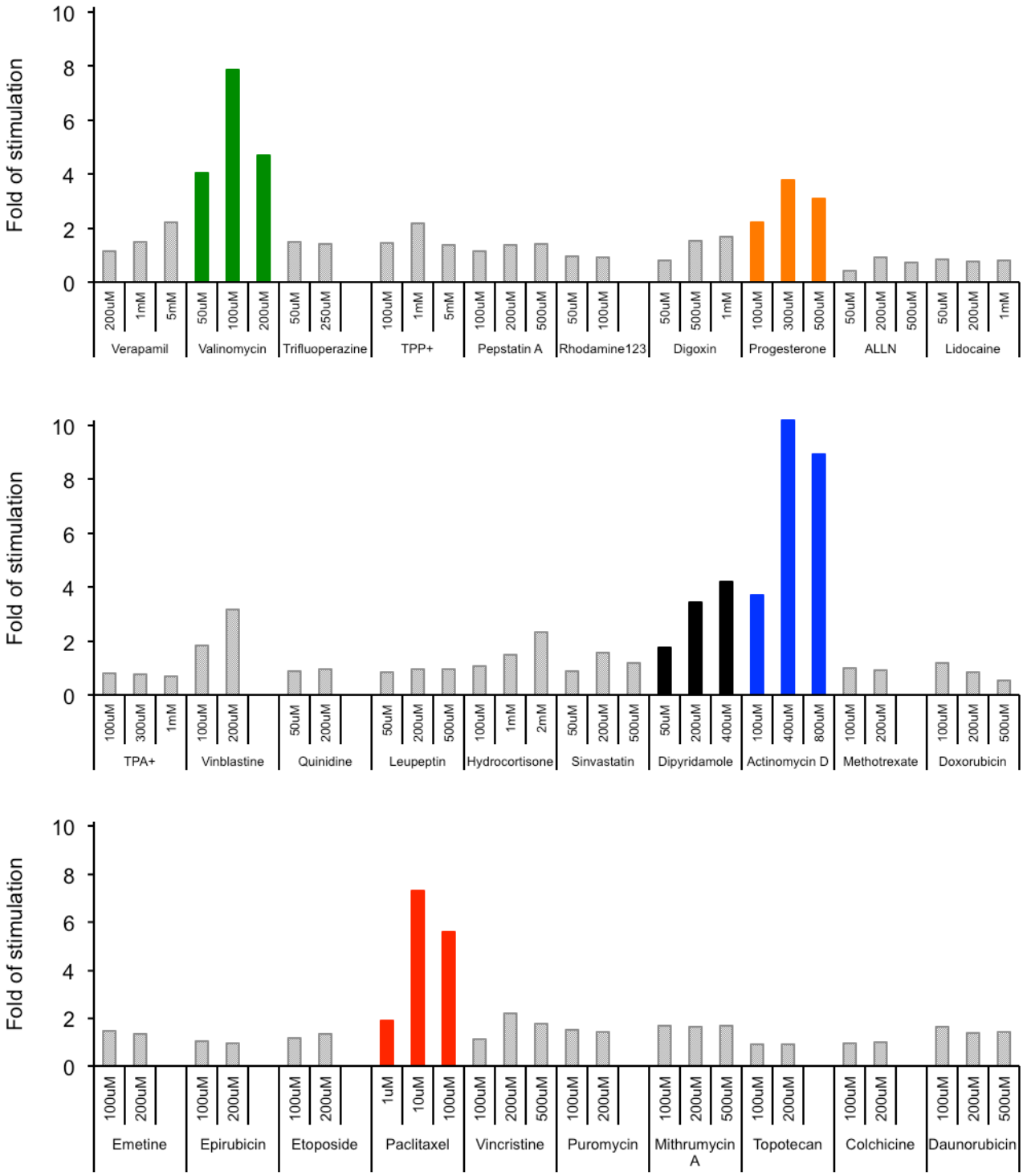
Crystal structure of the multidrug transporter P-glycoprotein from *C. elegans*

Mi Sun Jin[1], Michael L. Oldham[2], Qiuju Zhang[2] & Jue Chen[1,2]

[1]Department of Biological Sciences, Purdue University and [2]Howard Hughes Medical Institute West Lafayette, Indiana 47907, USA.

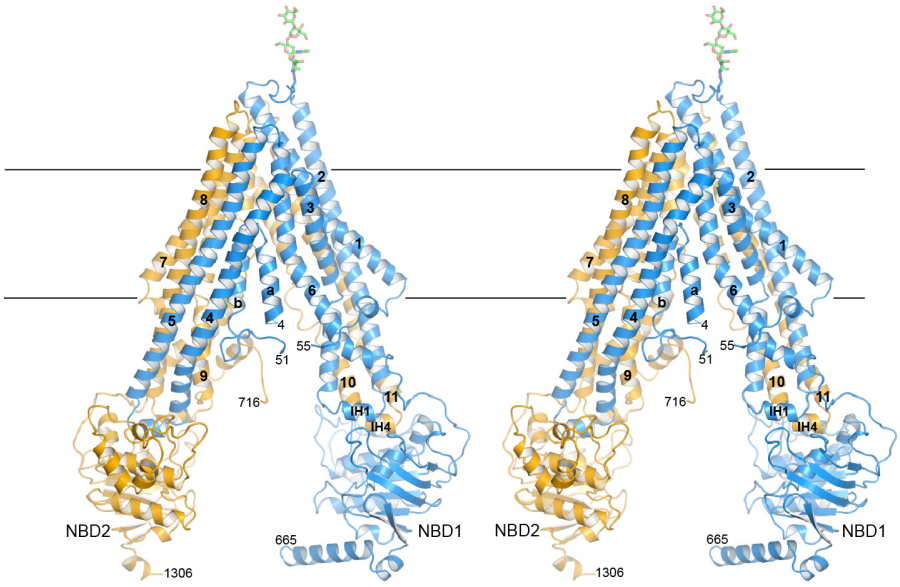


Supplementary Figure 1a. Expression of *C. elegans* P-gp confers cellular resistance to cytotoxic drugs. Sf9 cells expressing P-gp (+Pgp) were cultured in the presence of anticancer drugs actinomycin D (12 hours) or paclitaxel (48 hours) and imaged using light and fluorescence microscopy. Protection from drug-induced cytotoxicity is correlated with the appearance of P-gp in the membrane of infected cells as monitored by green fluorescence protein (GFP) fused on the C-terminus of P-gp. Images of uninfected cells (-Pgp) were also shown as controls.

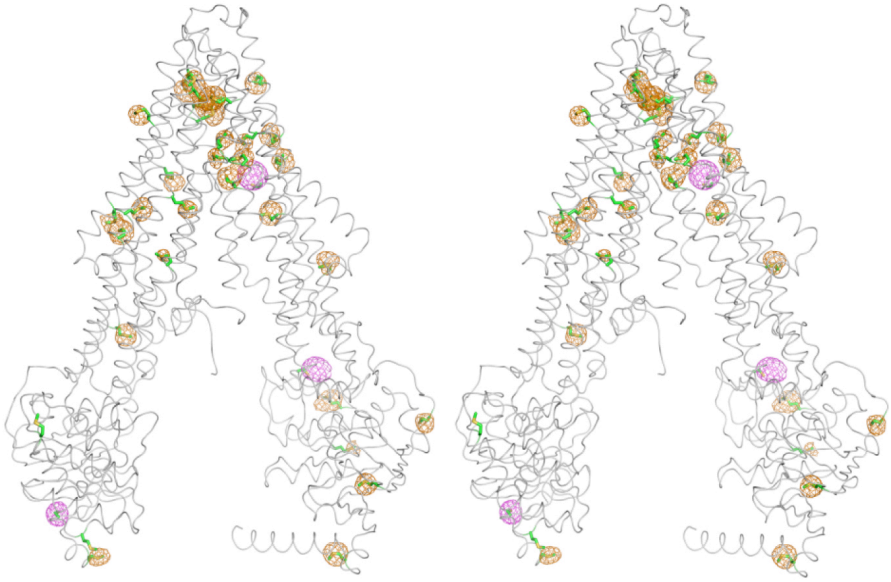


Supplementary Figure 1b. Stimulation of the ATPase activity by different compounds as indicated. The best five substrates are colored as in Fig. 1 in the main text.

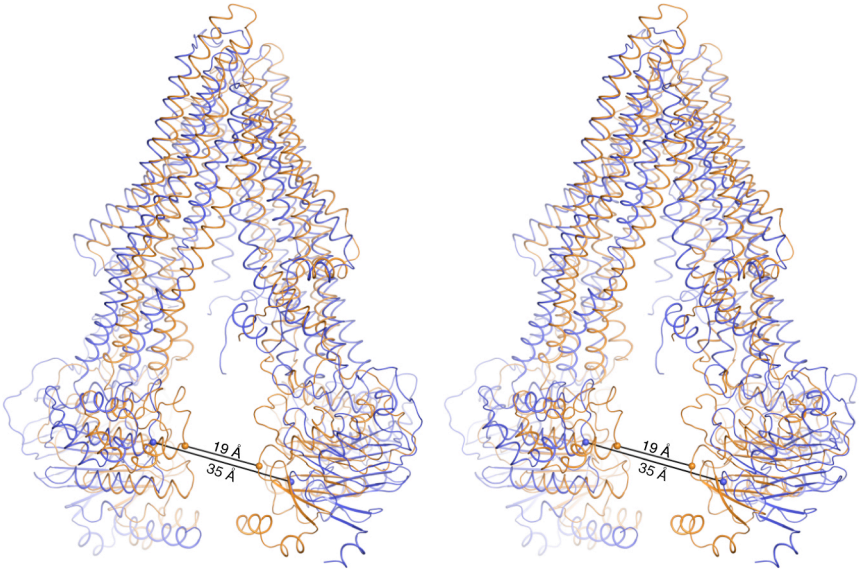
a



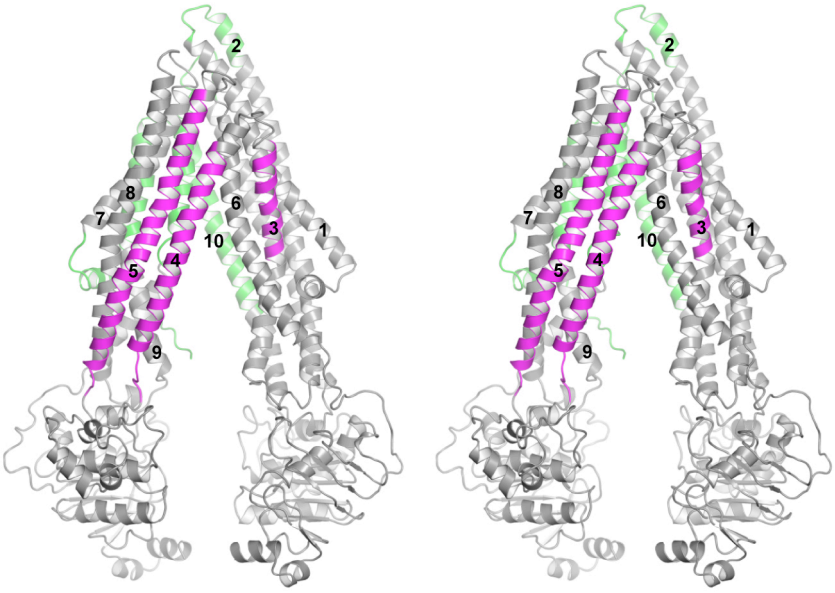
b



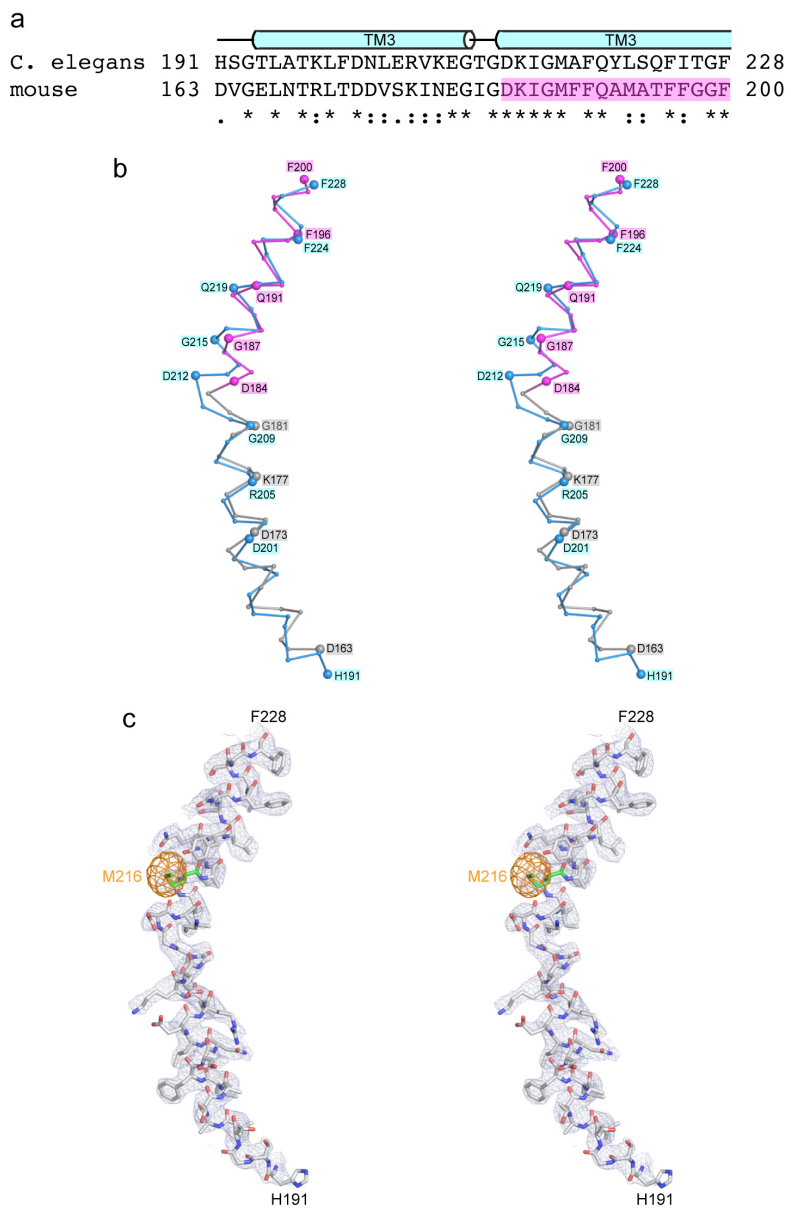
Supplementary Figure 2. The structure of *C. elegans* P-gp. a, Stereo view of a ribbon presentation. N125 and the attached oligosaccharide are shown in stick model. The N- and C- terminal halves of P-gp are shown in blue and gold, respectively. **b,** Stereo view of the anomalous difference Fourier electron density maps. The backbone of P-gp is shown in grey ribbon, methionine and cysteine residues are shown in green stick models. The orange and magenta meshes represent anomalous difference Fourier maps (contoured at 4σ) calculated from data collected at selenium and mercury absorption edges, respectively.



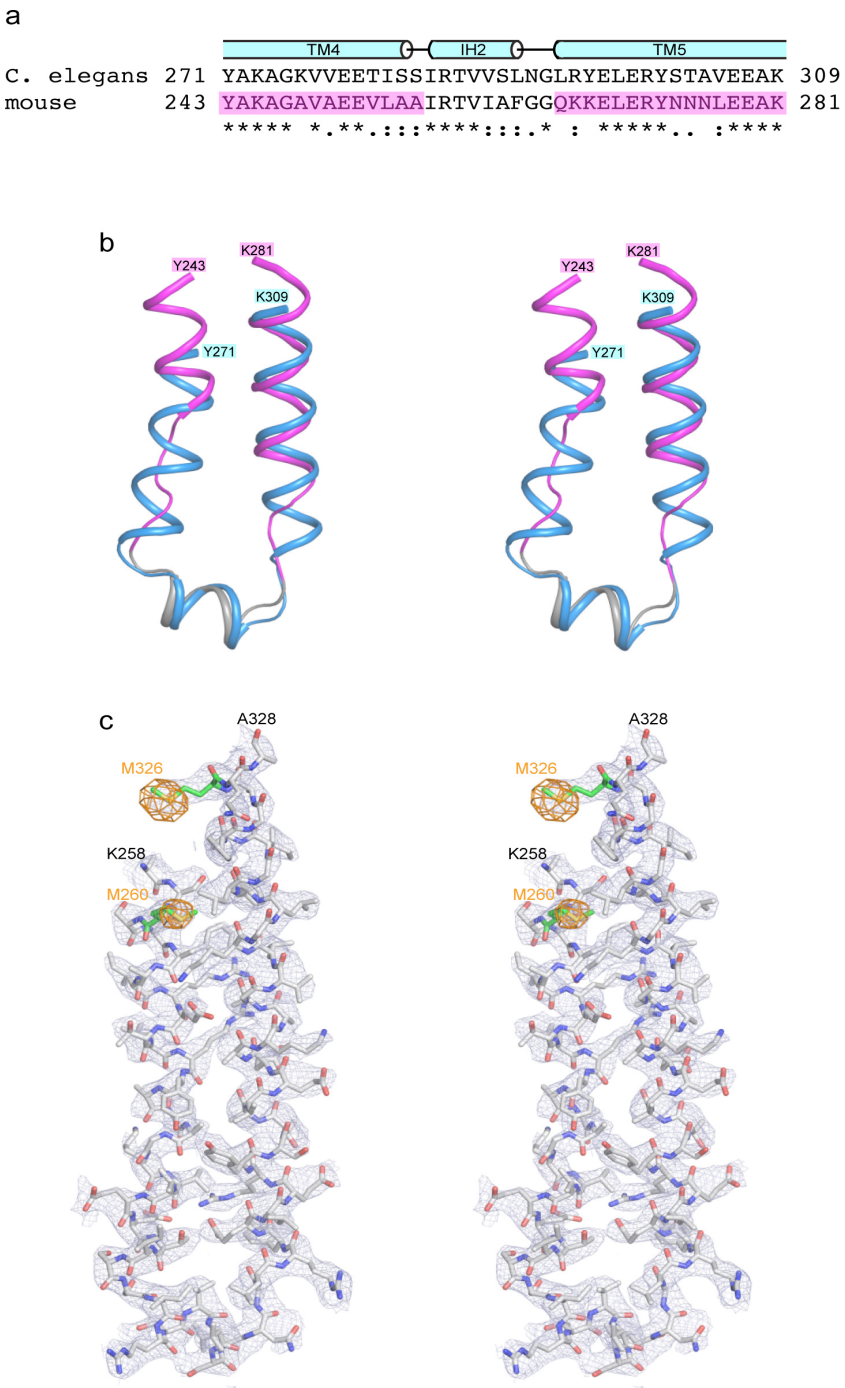
Supplementary Figure 3. Superposition of the mouse P-gp[1] (orange, PDB code: 3G5U) and *C. elegans* P-gp (blue) structures. The distances between two serine residues, one in the ABC signature motif and the other in the Walker A motif, are labeled for both structures.



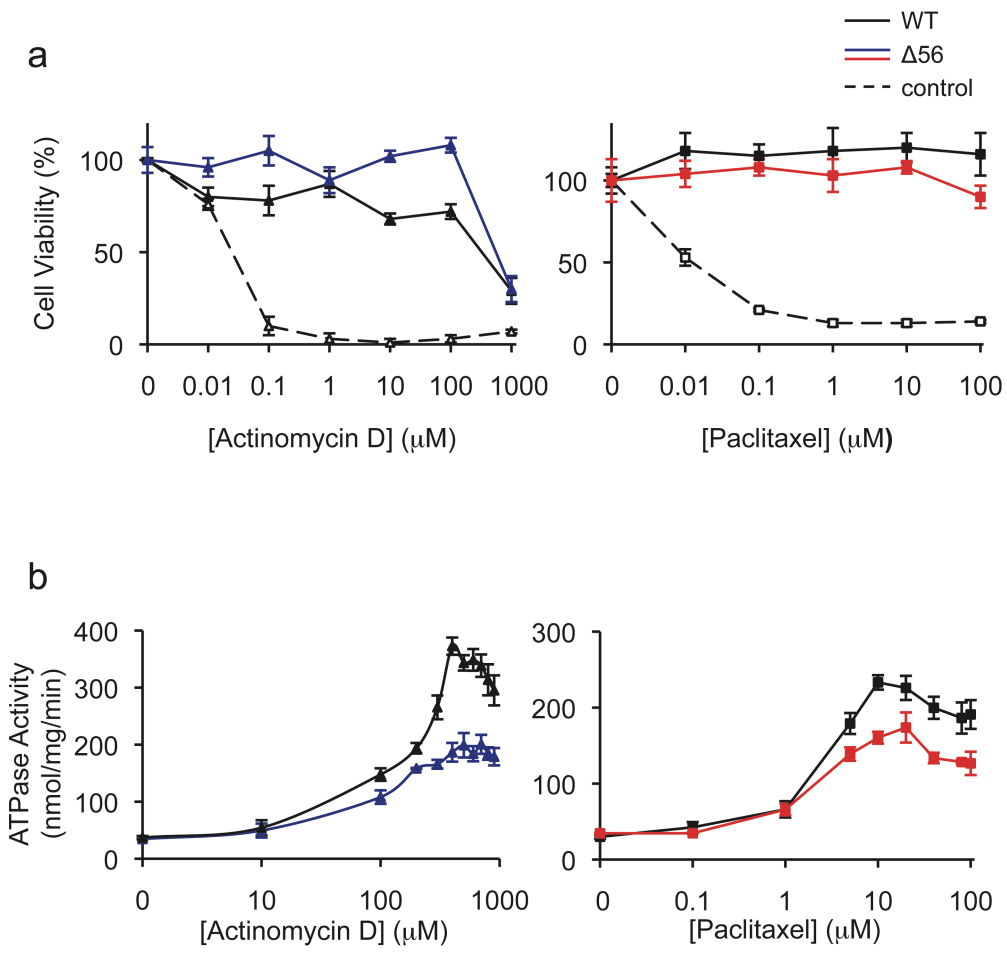
Supplementary Figure 4. Stereo view of the mouse P-gp structure[1] (PDB code: 3G5U). Color codes: magenta, regions containing register errors; green, regions where the structure is not directly comparable with that of the *C. elegans* P-gp; grey, regions where the register assignment is consistent with that of the *C. elegans* P-gp. A region of TM3 (residues 184-200 in mouse P-gp) is shifted by one amino acid, the entire TM4 helix (residues 217-251) is shifted by four amino acids, and TM5 is shifted by three amino acids, although only residues 266-281 are directly comparable between *C. elegans* and mouse P-gp. Furthermore, residues flanking IH2 (251-255 and 265-268) were built incorrectly as loops rather than α-helices in the mouse model.



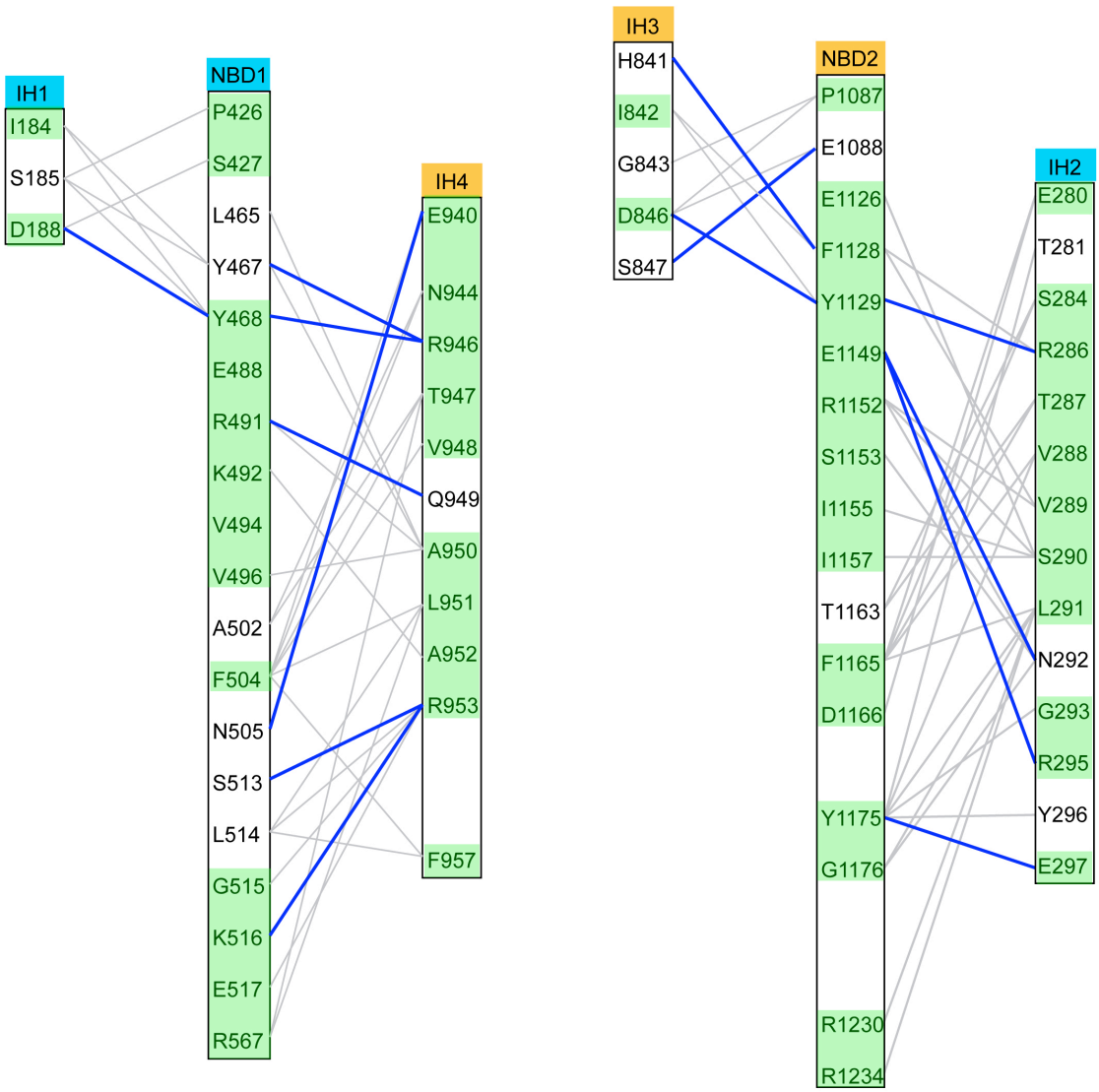
Supplementary Figure 5. Structural comparison of helix TM3. a, sequence alignment. The region in mouse P-gp containing a one-amino acid register shift is highlighted in magenta. **b,** structural alignment. Color codes: blue, *C. elegans* P-gp; magenta, regions in mouse P-gp containing register errors; grey, regions in mouse P-gp where register assignment agrees with that of the *C. elegans* P-gp. The Cα atoms of representative residues aligned based on sequence are shown as larger spheres and labeled: blue, *C. elegans* P-gp; grey/magenta: mouse P-gp. **c,** The electron density maps of the *C. elegans* P-gp in this region: grey, $2F_o$-F_c map (contoured at 1σ); orange: Seleno anomalous difference Fourier map (contoured at 4σ). The structure of *C. elegans* P-gp is shown in stick model.



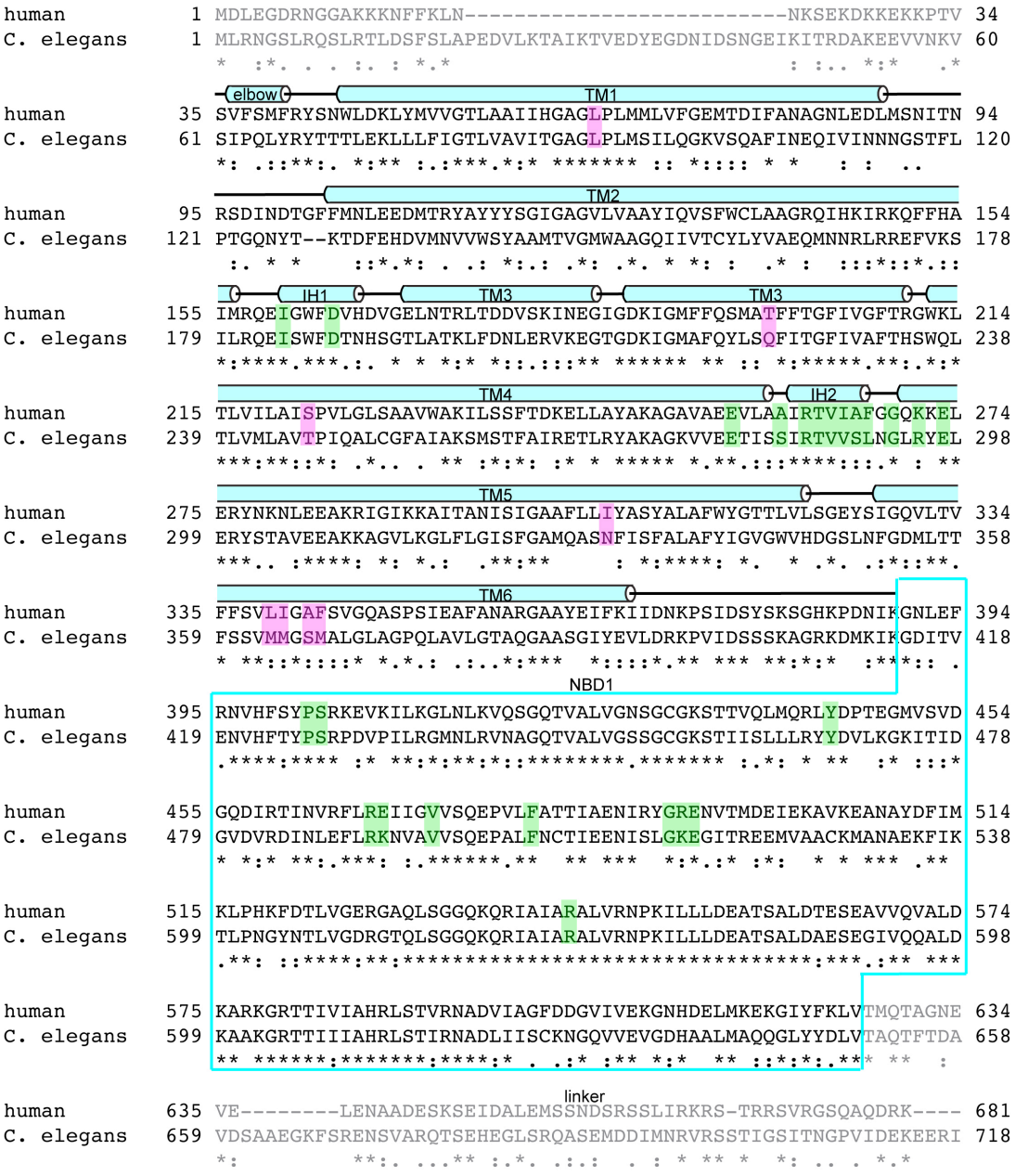
Supplementary Figure 6. Structural comparison of the TM4-TM5 region. For simplicity, only parts of helices TM4 and TM5 are shown. **a,** sequence alignment. **b,** structural alignment. **c,** The electron density maps of the *C. elegans* P-gp in this region: grey, $2F_o$-F_c map (contoured at 1σ); orange: Seleno anomalous difference Fourier map (contoured at 4σ). The structure of *C. elegans* P-gp is shown in stick model. Same color code as Supplementary Figure 5.



Supplementary Figure 7. Cytotoxic (a) and ATPase activity assays (b) of a truncation mutant devoid of the N-terminal 56 residues (Δ56). a, The N-terminal truncation mutant Δ56 confers cellular resistance to actinomycin D (blue) and paclitaxel (red) similarly to the full-length protein (WT). **b,** The truncation mutant has a reduced maximum level of drug-stimulated ATPase activity in detergent, however, the drug concentration dependence is unaltered for the truncation mutant.

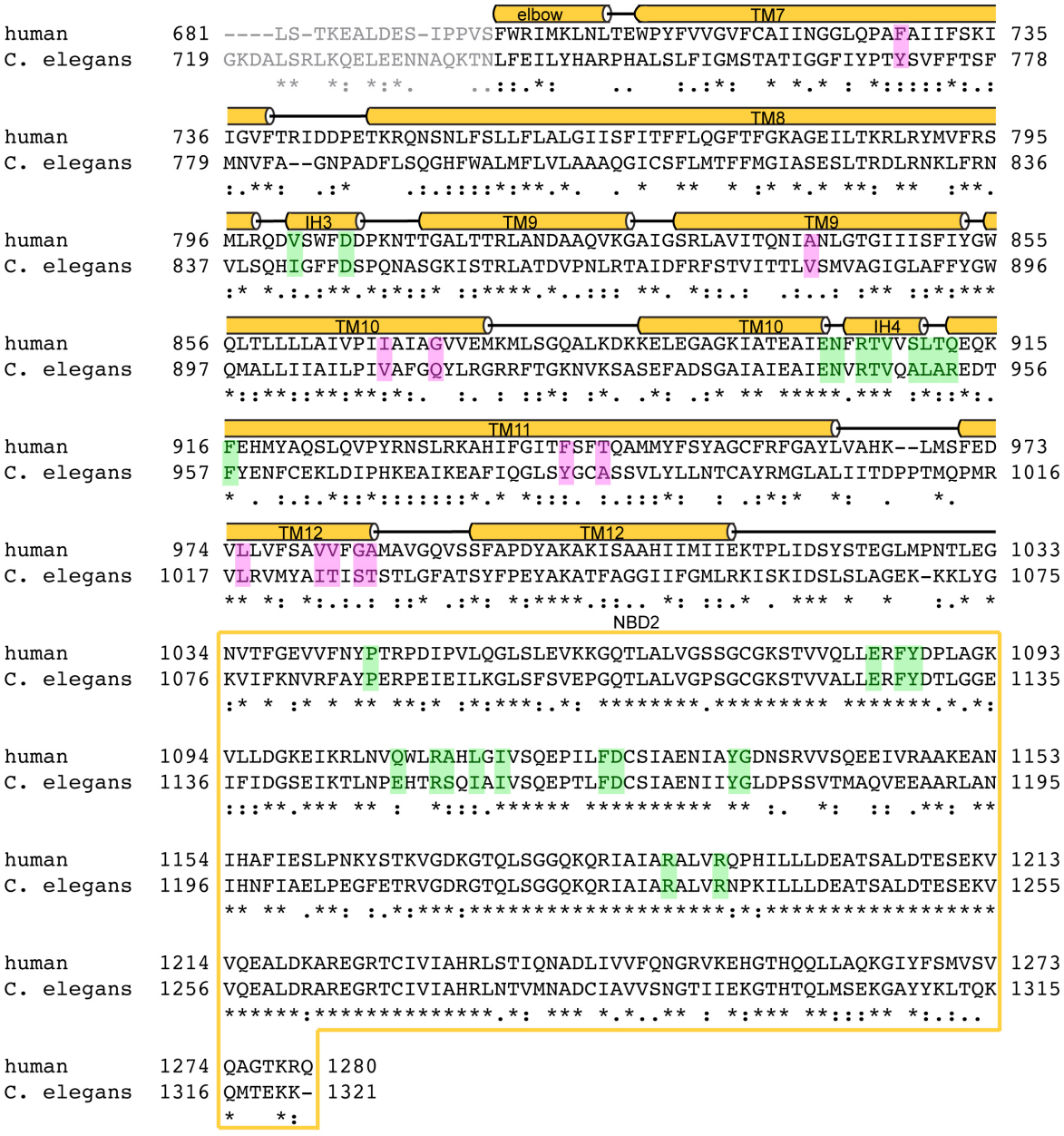


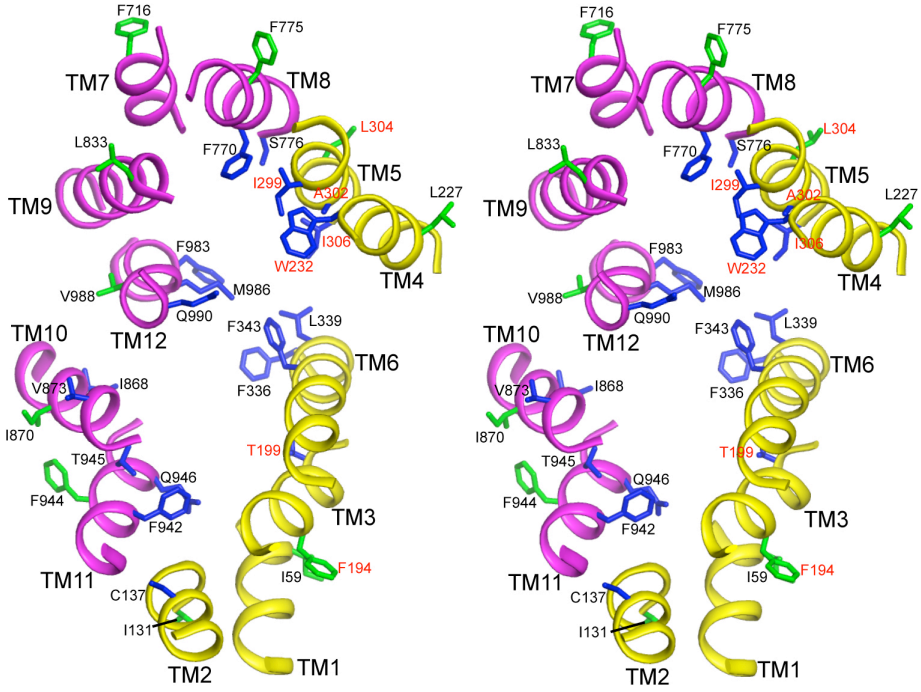
Supplementary Figure 8. Schematic diagram of the interactions between the TMDs and NBDs. Salt bridges and hydrogen bonds are indicated by blue lines; van der Waals interactions are shown as grey lines. Residues conserved in human P-gp are highlighted in green.



Supplementary Figure 9. Sequence alignment between human and *C. elegans* P-gp. Secondary structure is indicated for both the *C. elegans* P-gp structure and the corresponding human P-gp model. Conserved residues at the NBD/TMD interfaces are highlighted in green. Drug-interacting residues labeled in Figure 4f are highlighted in magenta. Residues that are not modeled in the human P-gp structure due to low sequence conservation are shown in grey. Figure is continued on the next page.

Supplementary Figure 9 (continued)





Supplementary Figure 10. Mapping the introduced arginine residues on the modeled human P-gp structure. Residues where arginine mutations enhanced P-gp maturation (shown as blue sticks) are presumed to line the drug-translocation pathway, whereas those inhibited maturation (shown as green sticks) are likely to face the lipids[2]. Residues in TMs 3, 4, and 5 incompatible with the mouse P-gp structure are labeled in red. This figure is modified from Figure 9 of Loo et al., *J Biol Chem*. **284,** 24074 (2009) except that we used our modeled human P-gp structure to interpret the arginine-scanning data.

Supplementary Table 1. Data collection and refinement statistics

	Native	SeMet	Hg
Data collection			
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions			
a, *b*, *c* (Å)	96.8, 155.3, 162.4	95.9, 156.5, 162.0	96.5, 155.8, 163.4
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)	50.0 - 3.4	50.0 – 4.2	50.0 – 4.1
R_{sym}	6.5	10.5	8.5
$I/\sigma I$	32.4 (1.5)	24.6 (1.4)	31.7 (1.5)
Completeness (%)	98.7 (95.5)	99.8 (99.9)	95.5 (91.0)
Redundancy	8.2 (6.1)	11.5 (9.1)	12.0 (7.8)
Refinement			
Resolution (Å)	50.0-3.4		
No. reflections	32319 / 1654		
R_{work}/ R_{free}	24.9 / 28.2		
No. atoms			
Protein	9628		
Carbohydrate	50		
Detergent	68		
B-factors			
Protein	144.0		
Carbohydrate	199.2		
Detergent	173.4		
R.m.s deviations			
Bond lengths (Å)	0.009		
Bond angles (°)	1.200		
Ramachandran plot (%)			
Most favored	90.8		
Allowed	8.0		
Generously allowed	1.2		
Disallowed	0.0		

*Highest resolution shell is shown in parenthesis.

References

1 Aller, S. G. *et al.* Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science (New York, N.Y* **323**, 1718-1722 (2009).

2 Loo, T. W., Bartlett, M. C. & Clarke, D. M. Identification of residues in the drug translocation pathway of the human multidrug resistance P-glycoprotein by arginine mutagenesis. *The Journal of biological chemistry* **284**, 24074-24087, doi:10.1074/jbc.M109.023267 (2009).