Bioscience Reports

www.bioscirep.org

Directed evolution of P-glycoprotein cysteines reveals site-specific, non-conservative substitutions that preserve multidrug resistance

Douglas J. SWARTZ*†[1], Leo MOK*†[1], Sri K. BOTTA*†, Anukriti SINGH*†, Guillermo A. ALTENBERG†‡ and Ina L. URBATSCH*†[2]

*Department of Cell Biology and Biochemistry, Texas Tech University Health Sciences Center, Lubbock, Texas, U.S.A.
†Center for Membrane Protein Research, Texas Tech University Health Sciences Center, Lubbock, Texas, U.S.A.
‡Department of Cell Physiology and Molecular Biophysics, Texas Tech University Health Sciences Center, Lubbock, Texas, U.S.A.

Synopsis

Pgp (P-glycoprotein) is a prototype ABC (ATP-binding-cassette) transporter involved in multidrug resistance of cancer. We used directed evolution to replace six cytoplasmic Cys (cysteine) residues in Pgp with all 20 standard amino acids and selected for active mutants. From a pool of 75000 transformants for each block of three Cys, we identified multiple mutants that preserved drug resistance and yeast mating activity. The most frequent substitutions were glycine and serine for Cys^{427} (24 and 20%, respectively) and Cys^{1070} (37 and 25%) of the Walker A motifs in the NBDs (nucleotide-binding domains), Cys^{1223} in NBD2 (25 and 8%) and Cys^{638} in the linker region (24 and 16%), whereas close-by Cys^{669} tolerated glycine (16%) and alanine (14%), but not serine (absent). Cys^{1121} in NBD2 showed a clear preference for positively charged arginine (38%) suggesting a salt bridge with Glu^{269} in the ICL2 (intracellular loop 2) may stabilize domain interactions. In contrast, three Cys residues in transmembrane α-helices could be successfully replaced by alanine. The resulting CL (Cys-less) Pgp was fully active in yeast cells, and purified proteins displayed drug-stimulated ATPase activities indistinguishable from WT (wild-type) Pgp. Overall, directed evolution identified site-specific, non-conservative Cys substitutions that allowed building of a robust CL Pgp, an invaluable new tool for future functional and structural studies, and that may guide the construction of other CL proteins where alanine and serine have proven unsuccessful.

Key words: ABC transporter, multidrug transporter, non-conservative cysteine substitutions, protein evolution site-saturation mutagenesis, yeast drug resistance.

INTRODUCTION

Pgp (P-glycoprotein), also known as ABCB1 or MDR1, is a member of the ABC (ATP-binding-cassette) transporter super-family that functions as a multidrug efflux pump [1,2]. It can contribute to multidrug resistance in cancers and other human diseases by preventing intracellular accumulation of cytotoxic agents, such as anticancer drugs [3,4]. Polyspecific binding of transported drugs and Pgp inhibitors occurs in a central cavity formed by two TMDs (transmembrane domains), but the precise locations and nature of these binding sites remain poorly defined [5,6]. Pgp and related ABC transporters have been crystallized in inward-facing and outward-facing conformations, which suggest
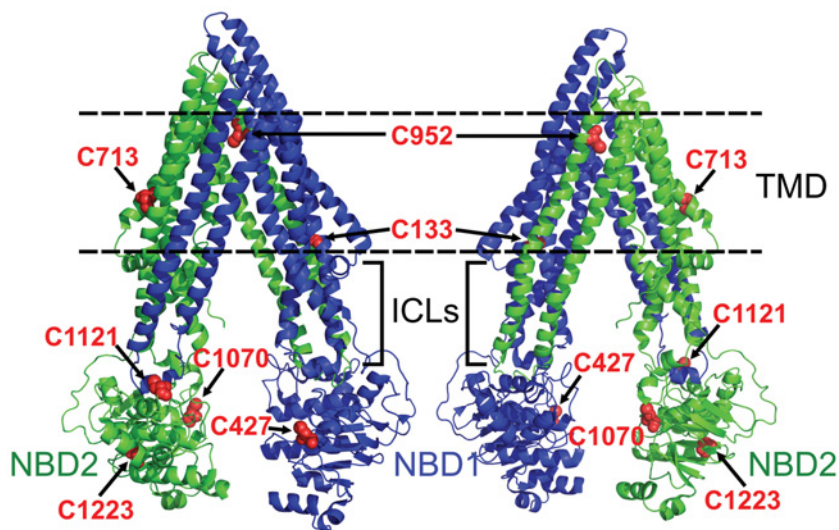

Figure 1   Location of the nine native Cys residues in Pgp

Cys residues (red) are shown in two views of the Pgp crystal structure (3G61). Cys^{427} is located in the N-terminal NBD (NBD1, blue), while Cys^{1070}, Cys^{1121} and Cys^{1223} are in the C-terminal NBD (NBD2, green). Cys^{133}, Cys^{713} and Cys^{952} are located in the membrane bound portion of the TMDS. Cys^{638} and Cys^{669} are in the flexible linker between NBD1 and TMD2 that was not resolved in the Pgp crystal structure. Dotted lines delineate the approximate location of the membrane bilayer. ICL, intracellular loop.

an alternating access transport mechanism that is powered by association/dissociation of the two intracellular NBDs (nucleotide-binding domains) [6–9]. However, these conformations have not been well defined in solution. Identifying conformational changes that occur during the Pgp drug transport cycle, and understanding the molecular details of the drug/inhibitor binding sites will provide information that is necessary for structure-based design of novel Pgp inhibitors that could then be evaluated for clinical efficacy [10].

A variety of biophysical techniques have been applied to studying conformational changes and protein ligand interactions in membrane proteins including EPR and NMR spectroscopy, and RET (resonance energy transfer) using optical probes [11–14]. Such studies in Pgp would require a CL (Cys-less) protein for the introduction of site-specific Cys residues to avoid labelling of native Cys that are susceptible to covalent modification [15,16]. Previously, two CL human Pgp proteins were generated by replacing each native Cys with Ala or Ser [16,17]. These two CL proteins were found to be functional in mammalian cells, but did exhibit differences in drug-binding and drug-stimulated ATPase activity when compared with each other or WT (wild-type) human Pgp. CL human Pgps have been extensively used for covalent crosslinking and substrate-binding studies using native cell membranes and purified protein [18,19]. However, mammalian expression is not well suited for large-scale purification, limiting its application to biochemical and biophysical studies that require significant amounts of highly purified protein. In addition, human Pgp protein yield and stability is lower than that of the mouse Pgp orthologue expressed in heterologous yeast expression systems [20,21]. To avoid this limitation, a CL mouse Pgp where

all Cys were replaced with Ala was constructed and expressed in yeast [22]. While this catalytically active CL mouse Pgp can be purified from *Pichia pastoris* yeast in reasonable quantities, its function is diminished. In addition, drug resistance that it confers to *Saccharomyces cerevisiae*, *in vivo*, is greatly impaired. Here, we undertook steps to generate a new fully functional, highly expressed mouse CL Pgp that will serve as an improved tool for biophysical studies of the Pgp transport mechanism.

CL proteins have been created for a variety of membrane transporters by the common strategy of replacing native Cys residues with Ala or Ser, but replacement of native Cys can often result in partial or complete loss of protein expression and/or function, as with Pgp [23–28]. A study by Ullman and co-workers demonstrated that site-saturation mutagenesis could be used to identify functional Cys substitutes, for individual Cys residues that could not be successfully replaced by Ala or Ser [29]. Their work suggests that Ala and Ser are not always the ideal Cys replacement. Mouse Pgp contains nine native Cys residues (2 more than human Pgp), with three located in the TMDs, three in the N-terminal NBD/linker region, and three in the C-terminal NBD (Figure 1). Our study uses a directed evolutionary approach to replace the six cytoplasmic Cys residues with all 20 standard amino acids and selecting for active mutants. Mutants were constructed using the codon optimized Pgp gene that we previously found to increase yield and quality of Pgp expressed in yeast [30]. We identified multiple Cys mutants that were as active as WT Pgp and maintained Pgp substrate polyspecificity, but contained non-conservative substitutions at six sites. The resulting CL-Pgp maintained the expression and full activity in yeast functional assays, and the purified protein displayed drug-stimulated
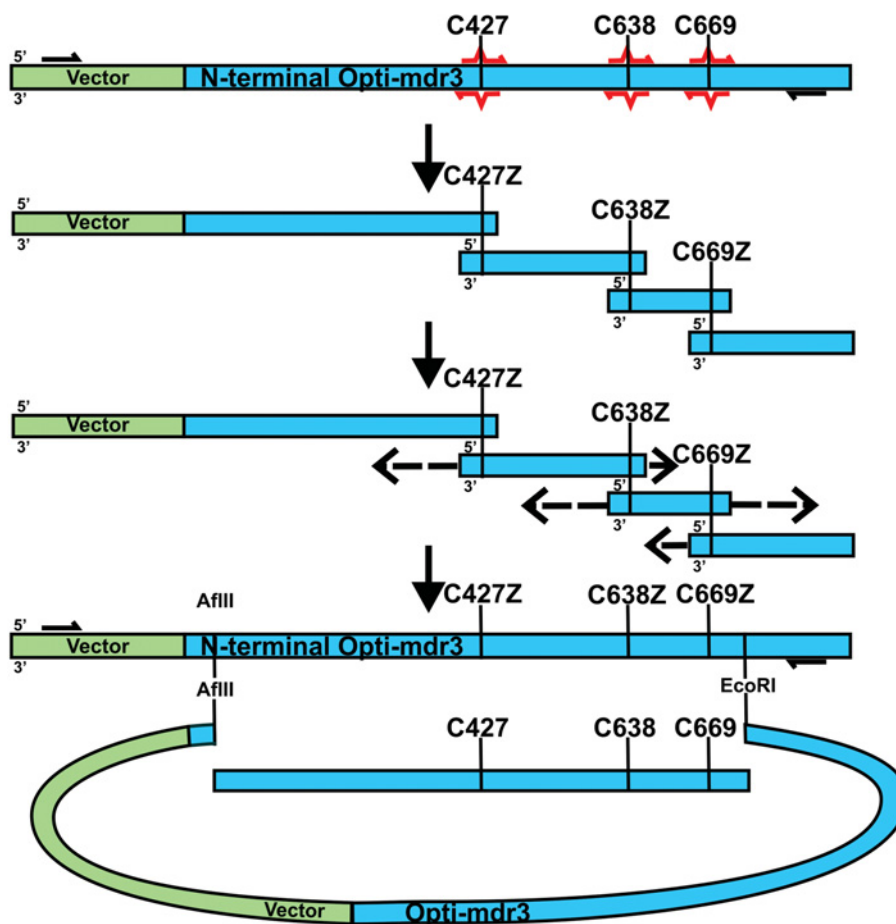
Figure 2   Strategy for construction of the C427Z, C638Z and C669Z mutant block by site-saturation mutagenesis
An overlap extension PCR approach is used to amplify initial fragments using mutagenic (red arrows) and flanking primers (black arrows) with pVT-opti-mdr3 as template. Gel-purified fragments are annealed and PCR assembled with the flanking primers to generate a fragment containing appropriate unique restriction sites (in this case 5′- AflII and 3′- EcoRI). Gel purified PCR fragments and vector (with the AflII-EcoRI fragment excised) are then cotransformed into yeast for assembly via homologous recombination.

ATPase activity indistinguishable from that of WT Pgp. Overall, directed evolution identified novel non-conservative cysteine substitutions, revealed important domain interactions and produced a CL protein that, because it is much more robust and active than previous versions, will be an invaluable new tool for studying the Pgp mechanism.

EXPERIMENTAL

Materials

FK506, cyclosporine A, doxorubicin and valinomycin were purchased from A.G. Scientific. Fluconazole was from LKT Laboratories. Verapamil, nystatin and ATP were from Sigma Aldrich. *Escherichia coli* lipids (Polar Extract) were pur-

chased from Avanti Polar Lipids, and DDM (n-dodecyl-β-D-maltopyranoside) was from Inalco. CP-MTS [7-diethylamino-3-(4′-maleimidylphenyl)-4-methylcoumarin] was from Biotium.

Mutant construction

The C427Z/C638Z/C669Z (NBD1/linker) and C1070Z/C1121Z/C1223Z (NBD2) mutant blocks were generated by site-saturation mutagenesis using oligonucleotides that contained a fully degenerate codon at each Cys position (NNN, where N represents an equal mixture of A, T, G and C nucleotides), as outlined in Figure 2 (Supplementary Table S1 at http://www.bioscirep.org/bsr/034/bsr034e116add.htm). A codon-optimized mouse Pgp (*mdr1a*, *mdr3*, GenBank JF834158) in the pVT expression vector (pVT-Opti-mdr3) served as the PCR template to enhance yeast expression and production of high-quality protein [30]. This optimized Pgp has the same amino acid sequence and is

functionally identical to non-optimized Pgp, and is referred to as WT Pgp. Degenerate codons were introduced by overlap extension PCR using Phusion Hotstart II polymerase (Thermofisher) [31]. First, mutagenic and non-mutagenic flanking primers were used by PCR to amplify four individual fragments that overlapped at the three Cys sites in each block. The four fragments were resolved on a 2% (w/v) agarose gel and purified by electroeluting DNA from excised agarose fragments in 12–14 kDa MWCO (molecular-mass cut-off) Spectrapor dialysis membrane (Spectrum) in TAE (Tris/acetate/EDTA) buffer. Each purified fragment (100 ng) was combined and extended in a second PCR reaction without additional PCR primers; then, flanking primers were added to amplify the full-length block. The flanking primers extended the NBD1/linker block by approximately 100 bp beyond an upstream AflII restriction site and a downstream EcoRI restriction site, to allow for homologous recombination with digested, gel-purified pVT-Opti-mdr3 plasmid in yeast [32]. Similarly, the NBD2 cysteine block containing C1070Z/C1121Z/C1223Z was generated with flanking primers that extended ∼100 bp beyond the upstream SpeI site and the downstream AgeI site for homologous recombination. The NBD1/linker and NBD2 mutant blocks were co-transformed with their respective digested pVT-Opti-mdr3 plasmids into competent *S. cerevisiae* JPY201 (MATa ura3 Δste6::HIS3) cells by electroporation [33,34]. Usually, ∼1 μg of cut plasmid was co-transformed with a 2-fold excess of PCR fragment per electroporation sample, yielding >3000 colonies; a total of 15 electroporations were performed for each block to obtain approximately 75000 transformants.

Selection and screening methods

Transformants were plated on uracil-deficient medium (uracil is the selection marker for the pVT vector [35] containing 150 μM fluconazole to start the selection process. Parallel samples were plated on uracil-deficient medium (without fluconazole) to estimate transformation efficiencies. Colonies from fluconazole containing plates were combined into mass population stocks, and a 1/20 volume of the stocks was inoculated into 15 ml liquid cultures containing 2.5 μM nystatin. The cultures were then grown to density in a shaking incubator at 30°C for 3–4 days. The resulting cultures were diluted 30-fold into 15 ml of fresh medium containing 150 μM fluconazole and again grown to density, followed by successive dilutions and growths in nystatin and a combination of doxorubicin (15 μM) and valinomycin (50 μM). Vector controls (transformed with empty pVT vector), and WT strains were grown in parallel to assess selection efficiency. Finally, the cells were allowed to recover for 3 h by growth in the drug-free medium and mated with an α-type yeast strain DC17 to identify those that restore mating in the otherwise sterile ste6Δ yeast strain JPY201 [34]. To identify the Cys mutations in yeast that successfully mated, DNA from single colonies was amplified by PCR in 96-well plates using the appropriate flanking primers for each block (Supplementary Table S1) and the PCR fragments sequenced (Beckman Coulter Genomics) [31]. Colonies picked for sequencing were maintained on uracil-deficient agar plates (one-well plates to track the PCR well numbers) and stored at 4°C. Plasmids of unique mutants were recovered from 10 ml yeast cultures by disrupting the cells with glass beads, purifying the plasmid DNA using a Wizard Plus miniprep Kit (Promega) and propagating extracted plasmids in XL10 Gold *E. coli* cells (Stratagene), followed by full-length sequencing.

Combining mutant blocks

Homologous recombination was also used to combine blocks of NBD1/linker and NBD2 Cys mutants. Plasmids containing NBD1/linker Cys mutants were digested with SpeI and AgeI and gel-purified. NBD2 Cys mutant blocks were PCR amplified with the flanking primers using single colonies as templates. The gel-purified fragments were pooled and co-transformed into yeast with the digested NBD1/linker plasmids. To confirm the presence of NBD2 mutants, single yeast colonies were amplified by PCR and digested with PstI; a unique PstI site is removed in all C1070Z mutants (Z≠C). A total of 42 single colonies were screened in yeast for biological activity. Then, plasmids were extracted again, and unique mutant combinations were confirmed by full-length sequencing before being retransformed into naïve *S. cerevisiae* for detailed analysis. C133A, C713A and C952A were generated by site-directed mutagenesis (QuikChange II XL, Stratagene) using the oligonucleotide primers given in Supplementary Table S1 to generate the fully CL pVT-CL-Opti-mdr3.

Analysis of Cys mutant expression and function

To assess mutant function, yeast fungicide resistance assays were performed as previously described [30,36]. Briefly, 10 ml yeast cultures were grown overnight in uracil-deficient medium containing 7.5% (v/v) glycerol, diluted to OD600 = 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 50 μM FK506, 100 μM valinomycin or 40 μM doxorubicin. Samples were grown in triplicate at 30°C for up to 30 h, and yeast cell growth was monitored by measuring the OD600 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 the Western blot analysis of microsomal membrane preparations as described [20].

WT and Cys mutant purification and analysis

For large-scale protein purification, JPY201 cells expressing WT and mutant Pgps were grown in 14 litre fermentor cultures (BioFlow IV) of uracil-deficient medium containing 7.5% glycerol, and harvested during log-phase growth (OD600 ∼4.0). Microsomal membranes were processed as previously described except that the membranes were washed with 500 mM NaCl to remove peripheral proteins and increase final protein purity [20,30]. Pgp was extracted with 0.6% DDM in the presence of 250 mM NaCl and purified by nickel affinity chromatography and SEC (size exclusion chromatography) as described for *P. pastoris* expressed Pgp [20,30], except that the NaCl concentration of all buffers was increased from 50 to 250 mM to increase protein stability [20,30].

Pgp concentrations were initially determined from the absorbance at 280 nm (A280) using a calculated molar extinction coefficient of 109,750 M-1 cm-1 (1 A280 unit = 1.29 mg/ml). WT and Cys mutants were resolved side-by-side on SDS/PAGE gels, stained with Coomassie Brilliant Blue and Pgp was quantified using ImageJ (http://rsbweb.nih.gov) to compare mutant Pgp levels; Pgp purified from *P. pastoris* served as standard. For ATPase assays, purified Pgp was activated with 10 mM DTT (dithiothreitol) and 1% *E. coli* polar lipids and ATPase activity was measured at 37 °C in a coupled assay utilizing an ATP-regenerating system, as previously described [30,37]. Drug-stimulated ATPase activity was analysed by the equation $V = Vbas + (Vmax \times S^b)/(S^b + Ks^b)$, where S is the concentration of stimulatory drug, V is the rate of ATP hydrolysis, Vbas is the ATPase activity in the absence of drug, Vmax is the maximum drug stimulated ATPase activity, Ks is the concentration of drug required for half-maximal stimulation, and b is the Hill coefficient. Cyclosporine A inhibition was analysed using the equation $V = Vmax-(Imax \times I^b)/(Ki^b + I^b)$, where I is the inhibitor concentration, Imax is the maximal inhibited ATPase activity, and Ki is the inhibitor concentration required for half-maximal inhibition.

RESULTS

Substitution of intracellular Cys residues

Each of the nine native Cys residues in mouse Pgp is highly conserved among Pgp orthologues except Cys^{638} and Cys^{669}, both located in the flexible linker that connects the two homologous halves of the protein (Supplementary Table S2 at http://www.bioscirep.org/bsr/034/bsr034e116add.htm). Four of the six cytoplasmic Cys residues are located in the highly conserved NBDs that contain motifs necessary for ABC transporter function (Figure 1) [2,15]. In contrast, the three Cys in the TMDs are located in highly ordered transmembrane helices where the small, hydrophobic methyl Ala side chain would likely be a good substitution because it supports the α-helical structure [38]. Therefore our first objective in this study was to replace the six cytoplasmic Cys by site-saturation mutagenesis, allowing every possible amino acid substitution at every Cys position, and determine which amino acids were the most likely to retain activity. Ideally, all six native Cys would be replaced simultaneously to account for potential structural distortions within the domains as well as interdomain communication (Figure 1). However the large number of possible combinations ($20^6 = 64 \times 10^6$) is impractical to screen. Instead, we simultaneously replaced the three Cys in NBD1 and the adjacent linker region (Cys^{427}, Cys^{638}, Cys^{669}) in one construct, and the three Cys in NBD2 (Cys^{1070}, Cys^{1121} and Cys^{1223}) in a second construct, reducing the required number of mutants (20^3 amino acids = 8000 per block of three, or 64^3 codons = 262000). For site-saturation mutagenesis, an overlap extension PCR approach was used to replace the native Cys codons with fully degenerate (NNN) codons that encode all 20 amino acids (64 codons), including Cys (Figure 2). Mutant PCR

libraries were directly transformed into *S. cerevisiae* by homologous recombination (see the Materials and Methods section for details) to select for active mutants by successive passages in different fungicides (nystatin, fluconazole, valinomycin and doxorubicin) followed by a mating assay. This selection scheme is based on the ability to convey resistance of the yeast to fungicides as well as export a-factor pheromone and restore mating in the otherwise sterile yeast by complementing for Ste6, a Pgp yeast homologue [34,39]. Mating, by itself, is a stringent test of mutant function as we previously demonstrated [40]. The customized selection scheme was designed to identify mutants that could transport multiple, structurally diverse substrates and preserve polyspecificity, an important quality of Pgp.

In total, approximately 75000 transformants carrying the NBD1/linker Cys mutants entered the selection process and 25 unique triple Cys mutant combinations were recovered after yeast mating. Likewise, 75000 transformants carrying the NBD2 Cys mutants passed through the selection process and 24 unique mutants were recovered. Unique mutants were defined as a distinct combination of amino acids as well as distinct codons across the Cys sites in a mutant block, or mutants with the same combination of codons that were derived from different transformation/selection samples. Analysis of the codon usage revealed an overall bias (>70%) towards low-frequency codons (defined as those occurring at <10% frequency in highly expressed yeast genes [30,41]) for all of the six Cys positions, possibly because Cys is an amino acid that occurs at low frequency (nine Cys per 1276 amino acids in Pgp). Frequencies of amino acid substitutions for each Cys site are given in Figures 3 and 4, and are summarized below according to the locations of the Cys in the protein.

Cysteine substitutions found by directed evolution

For Cys^{427}, located in the Walker A motif of NBD1, Gly was the most prominent substitution occurring at a frequency of 24%, followed by Ser at 20% (Figure 3A). Other small hydrophobic amino acids, Ala (12%) and Val (8%) and the polar Thr (8%) were also present. A similar pattern was observed for the Walker A motif Cys^{1070} of NBD2, with Gly (37%) being the most prominent substitution, followed by Ser (25%). A difference was that Ala and Thr occurred at very low frequencies of just 4% (Figure 4A). The Walker A motif is also called glycine-rich loop (or P-loop) because it wraps around the phosphates of bound ATP and positions the γ-phosphate for hydrolysis (Figure 3B). Thus small hydrophobic Gly or polar Ser substitutions may support better the flexibility of this loop.

For the two Cys residues in the linker region, Cys^{638} and Cys^{669}, many revertants were observed (12 and 20%, respectively). The most prominent substitutions for Cys^{638} were Gly (24%), Ser (16%) and Pro (12%) while Ala (8%) and Val (8%) substituted at lower frequencies. A different profile was observed for Cys^{669} with Gly, Ala, Pro and Thr all occurring at 12 to 16%, whereas Ser was absent. The linker is considered to be a highly flexible region that connects NBD1 to the TMD of the C-terminal Pgp half, which may explain a preference for small residues. No
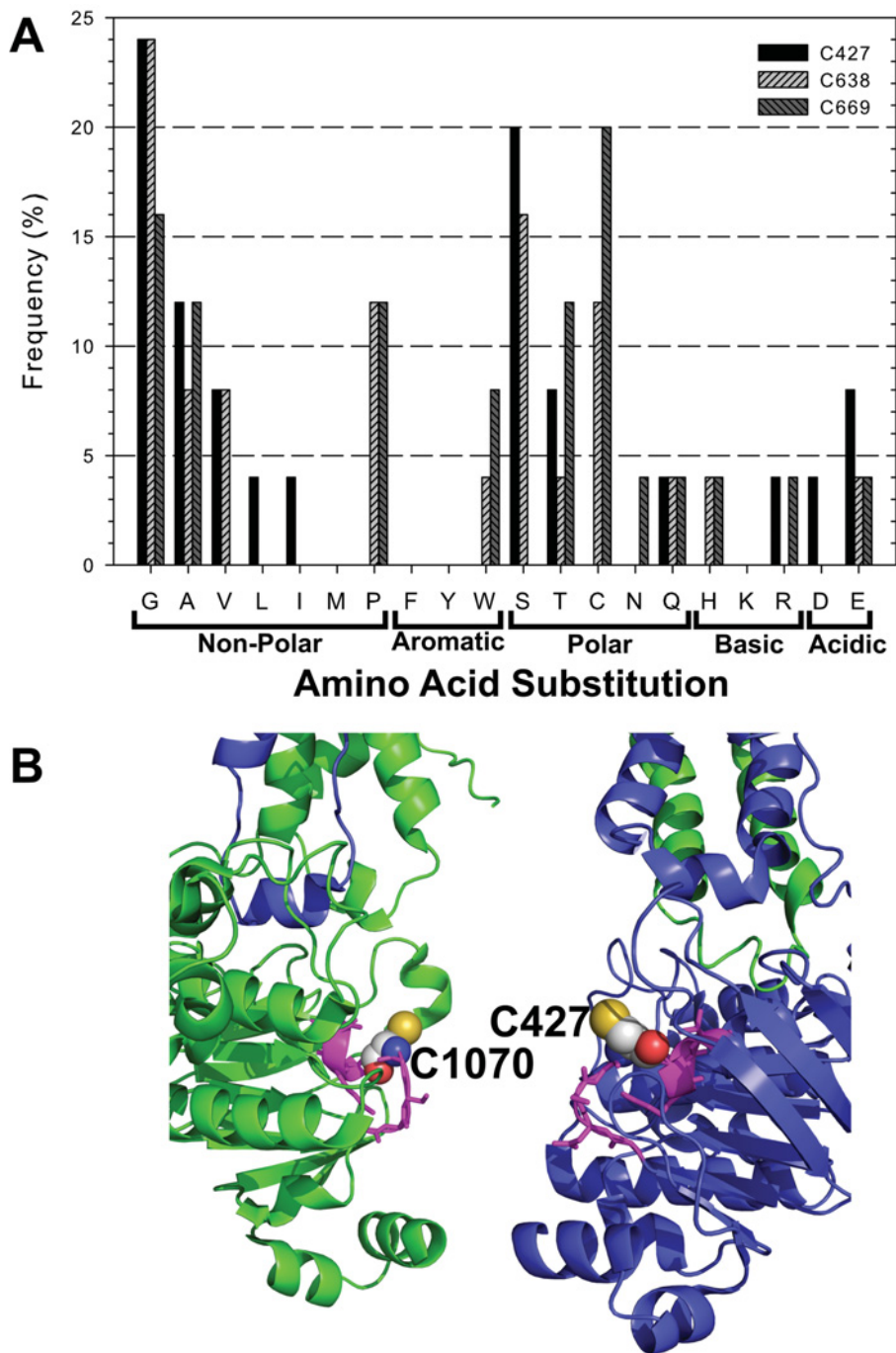
Figure 3   Frequency of NBD1/linker amino acid substitutions
(**A**) The frequency of each amino acid found at the three Cys positions in 25 unique NBD1/linker Cys mutants identified from sequencing of yeast colonies after selection. (**B**) Location of Cys[427] in NBD1 (blue) and Cys[1070] in NBD2 (green) of Pgp. The other amino acids in the Walker A motifs are shown in magenta.

structural information is available on those Cys since this region was not resolved in the crystal structure (Figure 1).

For Cys[1121], we found a clear preference for Arg (38%) while Thr (13%), Ser (8%), Gly (8%) and Asn (8%) substituted but at low frequencies. In the crystal structure of Pgp, Cys[1121] is located at the surface of NBD2 in close proximity to the ICL2 (intracellular loop 2) that extends from TMD1 and interfaces with NBD2 (blue, Figures 1 and 4B). In this arrangement, Cys[1121] in NBD2 would face Glu[269] and Gln[266] in ICL2. This may provide an explanation for the preference of a positively charged Cys
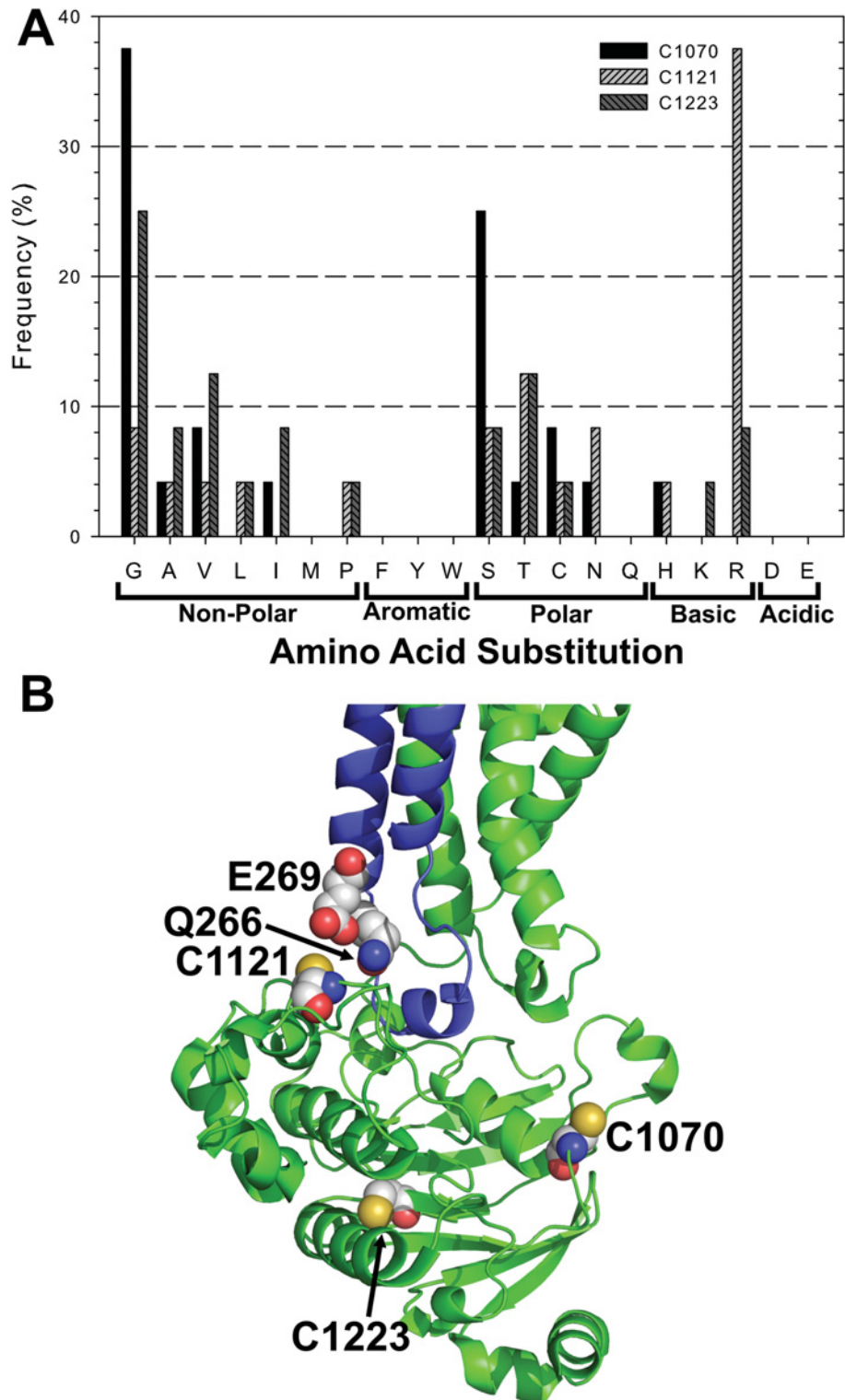
Figure 4   Frequency of NBD2 amino acid substitutions

(**A**) The frequency of each amino acid substitution found at the three Cys sites in 24 unique NBD2 Cys mutants. (**B**) Location of the Cys residues in NBD2 (green). A Glu and Gln, from the N-terminal ICL (blue), face Cys[1121] which was frequently replaced by the positively charged amino acid Arg.

substitution here. For Cys[1223], Gly was the most common replacement (25%), while Ala, Val, Ile, Ser, Thr and positively charged Arg all substituted at frequencies of 8 to 12%. Cys[1223] lies within a β-sheet in the core of NBD2 where space restrictions may favour a preference for small residue substitutions, with Arg posing an exception.

Screening and combining NBD1/linker and NBD2 Cys mutations

After identifying mutants that successfully mated, a random set of 15 unique NBD1/linker and 17 unique NBD2 mutants were tested for their ability to convey resistance against three structurally diverse fungicides, FK506 (804 g/mol), valinomycin (1111 g/mol) and doxorubicin (544 g/mol) (Figure 5A). We did observe differences in drug resistance profiles between the three compounds even though all of the mutated Cys positions are located in the intracellular portions of the protein, distal to the drug-binding sites. In general, most mutants were highly resistant against FK506, which was not part of the drug selection regimen (⩾80% growth compared with WT after 24 h, they did reach confluence after 26 h), as well as valinomycin, with a few exceptions (Figure 5A). The most pronounced differences were seen with doxorubicin as several mutants displayed enhanced growth compared with WT, while others grew slower, reaching only ∼50% of the WT Pgp value after 24 h (they did grow to confluency after 30 h). Importantly, all mutants were active even against a drug that was not part of the selection (FK506), validating our selection scheme.

Eight NBD1/linker and eight NBD2 mutants that maintained the highest levels of resistance against all three fungicides (highlighted in Figure 5A) were chosen to be randomly assembled into mutants with all six intracellular Cys removed (C6Z). Forty two mutant yeast colonies were initially tested for fungicide resistance to screen for incompatibilities between the NBD1/linker and NBD2 mutations. Then, plasmid DNA for the 19 most active mutants was extracted and retransformed into naïve yeast for detailed analysis. Overall, expression levels of the combination mutants were similar to WT (Figure 5B). Most mutants retained high levels of drug resistance against valinomycin and FK506 (at least 80% growth compared with WT) (Figure 5C). The most pronounced differences were again seen with doxorubicin. Several mutants reached only 40–60% of the WT value after 24 h, whereas others, E2, E7 and particularly H7, were as active against the three drugs as WT Pgp. Several mutants were identified twice in our screen of randomly recombined mutants (A9, B1, B7, D7, E2 and E7), and F1 and H7 occurred three-times. H7 was named C6Z-Pgp and chosen as template for further mutagenesis. These results demonstrate that random recombination of NBD1/linker and NBD2 Cys mutations, identified by directed evolution successfully yielded several Pgp proteins that maintain high levels of activity with six native Cys residues removed.

Removal of TMD cysteines and construction of a CL Pgp – Our second objective was to remove the remaining Cys residues in the TMDs (Cys[133], Cys[732], Cys[952]; Figure 1), individually and successively, from the C6Z-Pgp mutant. We chose Cys to Ala substitutions because Ala is known to stabilize α-helical structure [38].

Western blots revealed that all TMD Cys-to-Ala mutants maintained expression when compared with WT and C6Z (Figure 6A). All single and double TMD Cys mutants also maintained high levels of resistance to fungicides (Figure 6B), suggesting that Ala is a good substitution for Cys in transmembrane α-helices of Pgp. When all three TMD Cys mutations were added to the C6Z mutant to form a fully CL protein (CL-Pgp), the resulting protein exhibited at least 75% resistance against all three drugs relative to WT. For comparison, the mouse CL Pgp previously made by replacing all nine Cys with Ala in the non-codon-optimized Pgp background ('all-Ala-CL') was also expressed in *S. cerevisiae* yeast and assayed in parallel (Figure 6B) [22]. The all-Ala-CL Pgp only conferred partial resistance against valinomycin and none against FK506 and doxorubicin, demonstrating the value of the new CL-Pgp that was generated by directed evolution.

Purified CL Pgp as background for thiol-selective labelling

CL-Pgp, WT and selected Cys mutants were purified by affinity and SEC as described, except that NaCl (250 mM) was included in all buffers, as higher salt concentration have been shown to improve purification of human Pgp (see the Materials and Methods section for details) [30,42]. Proteins eluted in a major peak (at ∼15.1 ml), consistent with an apparent size of ∼200 kDa (monomeric protein), and a minor peak at ∼13 ml that corresponded to a Pgp oligomer (Figure 7A). In contrast, when CL-Pgp (or WT, not shown) were chromatographed in 50 mM NaCl buffer the proteins eluted as several broad peaks in early fractions (9–14 ml) as well as in the void volume (Figures 7A and 7B), suggesting Pgp oligomerization and aggregation in low-salt buffers, contrary to *P. pastoris*-purified Pgp [30]. Purity of all Cys mutants was comparable with that of WT -Pgp, with minimal degradation products observed on Coomassie Brilliant Blue-stained SDS-gels (Figure 7C) or Western blots (not shown). The yield of the SEC-purified proteins was about 2.4 mg/100 g yeast cells from 14 litres fermentor cultures.

One major use of a the functional CL-Pgp is that it can be employed as framework to introduce single- and double-Cys mutants at selected positions for labelling with optical probes for spectroscopic studies. To demonstrate the usefulness of the CL Pgp as background for labelling studies, we compare it with a single (Cys[713]) and a double -Cys mutant (T492C/S1137C) that, as previously shown, has two very accessible cysteines in the NBDs [43]. Cys[713] was generated as part of our site-directed mutagenesis aimed at substituting the last three TMD Cys (see above). For the labelling studies, Cys[713] was reacted with the thiol-reactive coumarin dye, CP-MTS (Figure 7D). CL-Pgp and mutant T492C/S1137C were subjected to the same labelling protocol. Minimal fluorescence was detected in the CL-Pgp band on SDS-gels while Cys[713] clearly showed fluorescence, and fluorescence increased in the T492C/S1137C as expected from the number of cysteines present (Figure 7D). The data indicate that the transmembrane Cys at position 713 is accessible to thiol reagents. In the crystal structure, the Cys[713] sulfhydryl is oriented towards neighbouring α-helices (Figure 1) and it is therefore likely that
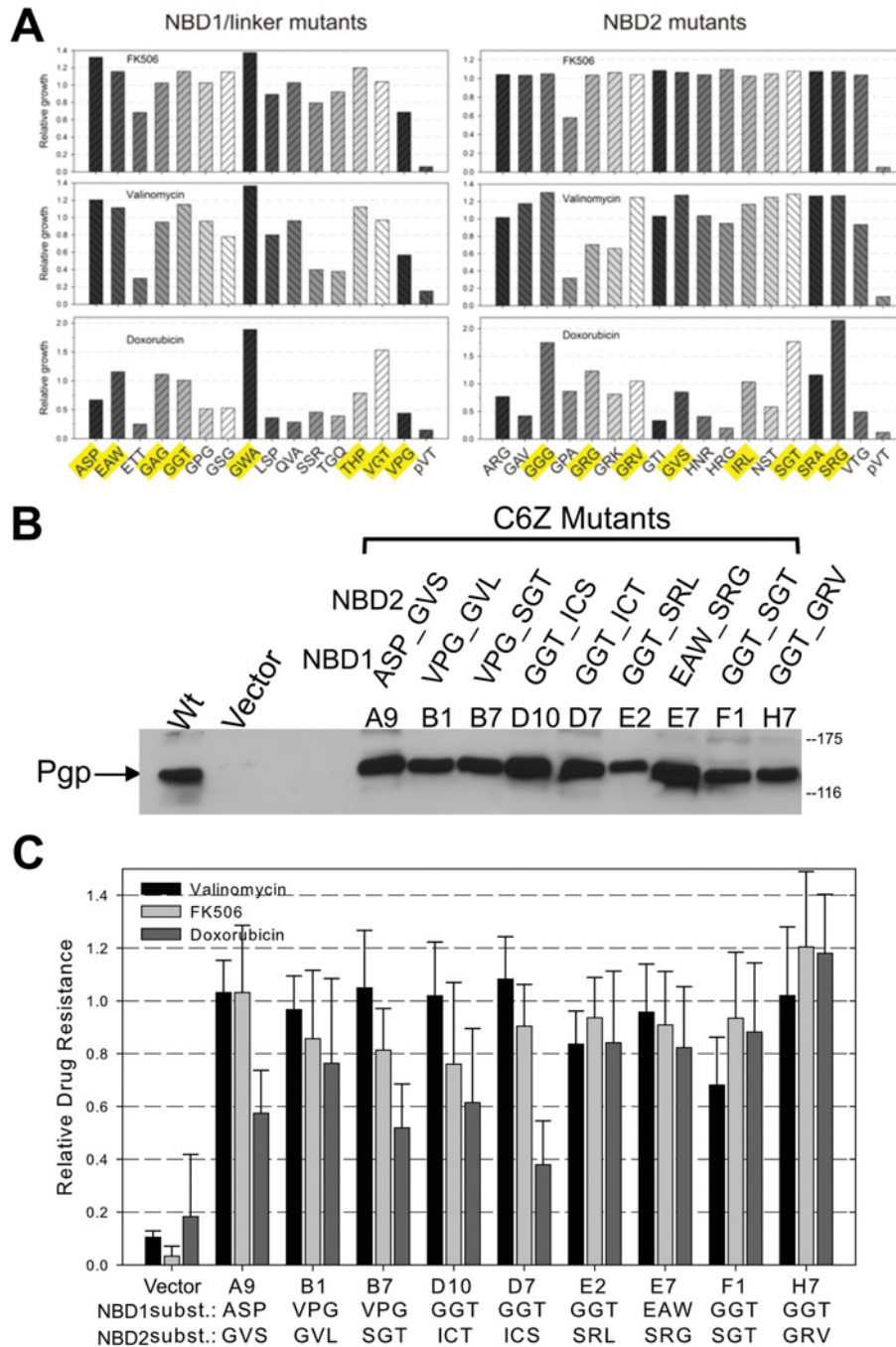
Figure 5 *In vivo function of NBD1/linker and NBD2 Cys mutants*

(**A**) NBD1/linker and NBD2 Cys mutants were screened for resistance to 100 μM valinomycin, 50 μM FK506, and 40 μM doxorubicin. Relative resistance was calculated from the growth of each sample in drug compared with drug-free medium and then normalized to the relative growth of WT Pgp expressing yeast (WT = 1). Amino acid substitutions are given in single letter code, e.g. ASP is C427A/C638S/C669P (left) and ARG is C1070A/C1121R/C1223G (right). A representative of two experiments is shown. Highlighted mutants were used to assemble the two mutant blocks. (**B**) Eight NBD1/linker and nine NBD2 unique mutants (highlighted in **A**) were combined as described in the Materials and Methods section, then plasmids extracted and retransformed into naive yeast cells for analysis of Pgp expression and function. 15 μg of crude microsomal membrane protein from each mutant expressing yeast strain and yeast expressing WT Pgp and empty pVT vector was analysed by Western blot using the Pgp-specific monoclonal antibody Cys[219]. Labels indicate the well number of the mutants from screening (A9 to H7), and the substitutions for the three consecutive Cys positions in NBD1 and NBD2 are given as single letter amino acid abbreviation. (**C**) The same nine mutants were tested for resistance against multiple drugs as in (**A**). Bars represent the mean of multiple replicates from three independent experiments $\pm$ S.D..
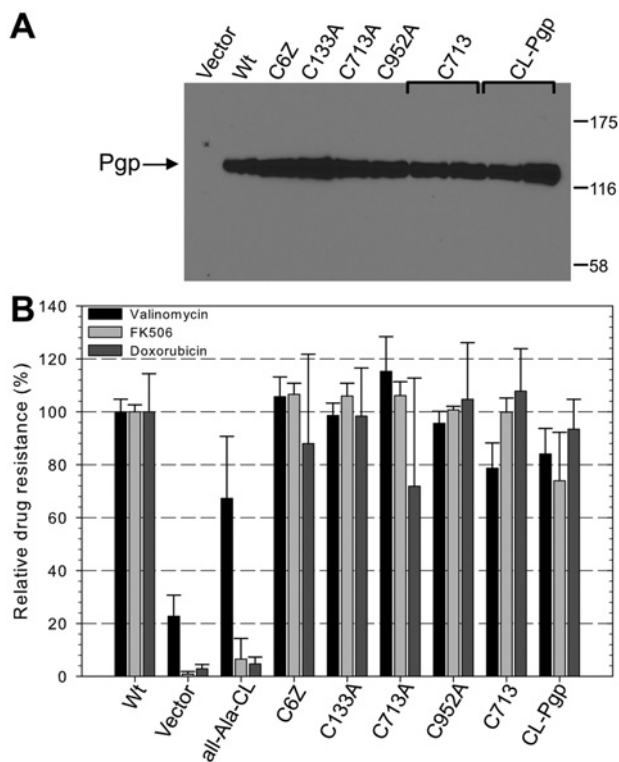
Figure 6 Expression and *in vivo* transport activity of CL and low Cys Pgp mutants

(**A**) Western blot analysis of the diverse mutants was done as in Figure 5 with 15 μg of microsomal membrane protein loaded per lane (**B**). The CL and low-Cys mutant expressing strains were tested for fungicidal resistance as in Figure 5. C6Z- Pgp (H7 from Figure 5C) has all six cytoplasmic Cys removed and retains three native Cys in the TMDs. C133A, C713A and C952A each retain two native Cys in the TMDs, while C713 retains only one native Cys. CL-Pgp is fully CL with all nine Cys removed. All-Ala-CL, with all Cys mutated to Ala in the WT Pgp background, was included for comparison. Bars indicate the mean relative growth of samples normalized to the relative growth of WT Pgp expressing yeast from three independent experiments $\pm$ S.D., nine independent experiments for CL-Pgp.

it is not fully labelled. However, the low CL-Pgp background and the high sensitivity of fluorescence techniques allow for easy detection of the single-Cys. These results demonstrate the usefulness of the functional CL-Pgp as background for site-specific labelling of Cys residues introduced at selected positions in the protein.

ATPase activity of CL-Pgp

Purified proteins were assayed for drug stimulation or inhibition of ATPase activity, which is a benchmark test for comparing purified Pgp proteins [40,44]. Maximum verapamil-stimulated ATPase activity was similar for CL-Pgp, C6Z-Pgp and WT with 3.2$\pm$0.6, 3.5$\pm$0.2 and 3.5$\pm$0.7 μmol/min/mg ATP hydrolysed, respectively (Table 1). This value is somewhat higher than that of *P. pastoris*-purified Pgp (2.1$\pm$0.3 μmol/min/mg) [30], suggesting that increased salt is also beneficial for ATPase activ-

ity. The half-maximal stimulatory concentrations for verapamil were 4.8 μM, 2.0 and 3.8 μM for CL-Pgp, C6Z-Pgp and WT, respectively (Figures 8A and 8C; Table 1), not significantly different in the two tailed *t* test ($P > 0.5$). Inhibition of the verapamil-stimulated ATPase activity by the immunosuppressant cyclosporine A was also comparable for the two proteins, with half-maximal inhibition at 4.8 μM, 3.4 and 2.0 μM for CL-Pgp, C6Z-Pgp and WT, respectively ($P > 0.05$; Figures 8B and 8D; Table 1). Taken together, the enzymatic data indicate unaltered affinities for substrates and inhibitors between WT, C6Z-Pgp and CL-Pgp.

DISCUSSION

This study employed a directed evolutionary approach to determine the preferred amino acid substitutions for Cys in the cytoplasmic domains of Pgp (Figures 3 and 4) to generate several highly-active low-Cys mutants with six Cys residues removed (Figure 5). We initially focused on the cytoplasmic Cys residues because they are located in the highly conserved, complexly folded NBDs where directed evolution could identify beneficial non-conservative substitutions and domain–domain interactions (Figure 1) [2,15]. Then, the three Cys in the transmembrane α helices were replaced with Ala, as the small non-polar side chain is known to support α-helical structure [38]. Substitution of these Cys residues by Ala resulted in a CL protein that was well expressed, displayed high levels of resistance against multiple fungicides, and could be purified in yields suitable for biophysical studies with high ATPase activity. The success of this mutagenesis strategy can be seen by comparing the final CL-Pgp, which conveyed high levels of yeast fungicide resistance, with the previously constructed all-Ala-CL mutant. The latter produced minimal drug resistance when expressed in yeast (Figure 6B). These results show that directed evolution successfully produced a novel CL Pgp that bypassed the limitations of the previously constructed mouse and human CL Pgps [16,17,22,30], as it is both active in yeast functional assays and can be readily purified in milligram quantities.

The site-saturation mutagenesis strategy used in this study replaced each of the six intracellular Cys with a fully degenerate codon that encoded all 20 standard amino acids, including Cys. Although Cys was allowed in this strategy, Cys revertants were only found abundantly at two positions in the linker region (Cys[638] and Cys[669]), suggesting that other amino acids support Pgp function and could directly compete with Cys during selection (Figures 3 and 4). In general, Gly and Ser were the most prominent substitutions at multiple native Cys positions, contrary to the BLOSUM62 (BLOcks of Amino Acid SUbstitution Matrix) matrix where Gly carries one of the highest penalties (-3) for Cys replacement, while Ser is considered a more conservative substitution (penalty $= -1$) [45]. In a similar directed evolution study, Ullman and co-workers found that site saturation mutagenesis of Cys produced substitutions with small non-polar

Table 1 ATPase activity of CL-Pgp and low-Cys mutants

Protein	ATPase activity* (μmole Pi/mg/min)	Ks Verapamil (μM) †	Ki Cyclosporine A (μM)‡
Wt	3.5 ± 0.66	4.8 ± 1.8	2.0 ± 0.8
C6Z	3.5 ± 0.19	2.0 ± 1.1	3.4 ± 0.9
CL-Pgp	3.2 ± 0.56	3.8 ± 3.3	4.8 ± 2.3

*Maximum verapamil-stimulated ATPase activity in the presence of 30 μM verapamil (mean $\pm$ S.E.M.).
†Concentration required for half-maximal ATPase activity derived from plots of V versus [drug] in Figures 8A and 8C (mean $\pm$ S.E.M.); similar plots were obtained for C6Z-Pgp (not shown).
‡Cyclosporine A concentration required for half-maximal inhibition of verapamil-stimulated ATPase in Figures 8B and 8D (mean $\pm$ S.E.M.); similar plots were obtained for C6Z-Pgp (not shown).
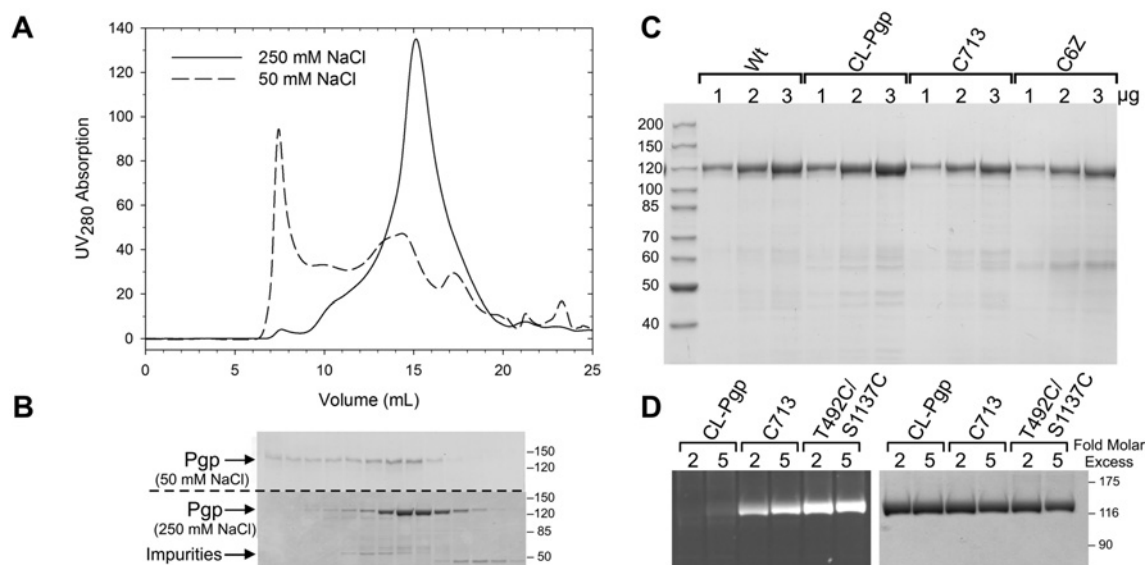
Figure 7   Purification and thiol-selective labelling of Cys mutants
(**A**) Cl-Pgp was purified from *S. cerevisiae* fermentor cells by affinity chromatography on Ni-NTA resin as detailed in the Materials and Methods section , then 2–3 mg (500 μl) loaded on a Superose 6B column and resolved in buffers containing 50 or 250 mM NaCl. (**B**) SEC fractions were resolved on SDS-gels and stained with Coomassie Brilliant Blue. Molecular mass protein markers are given in kDa. Mutants Cys[713] and C6Z-Pgp, or WT Pgp purified and chromatographed in 250 mM NaCl buffers gave very similar elution profiles as Cl-Pgp (not shown). (**C**) Increasing amounts of each purified protein were resolved on 10% SDS/PAGE gels and stained with Coomassie Brilliant Blue. (**D**) CL-Pgp, Cys[713] and a control double Cys mutant T492C/S1137C were labelled for 1 h at room temperature with a 2-fold and 5-fold molar excess of the methanethiosulphonate reactive dye of coumarine (CP-MTS) as described, resolved on an SDS-gel and imaged at 385 nm (left), or stained with Coomassie Brilliant Blue (right). T492C/S1137C was generated by site-directed mutagenesis in the pVT-CL-opti-mdr3 background.

amino acids and also a prevalence of larger non-polar residues such as Leu and Met (penalty $= -1$) [29]. We would argue that the prevalence for a given substitution is likely dependent on the location of the Cys within a given protein. In Pgp for example, Gly and Ser most frequently replaced Cys residues in the glycine-rich Walker A motifs (Cys[427] and Cys[1070]). These two Cys are highly conserved among Pgp orthologues (Supplementary Table S2 at http://www.bioscirep.org/bsr/034/bsr034e116add.htm), but both Gly and Ser are naturally prevalent at these two positions in other ABC transporters [46].

Among the unique mutants, Cys was often substituted by small non-polar or polar amino acids that likely maintain or enhance the flexibility at the Cys position. Cys[1121] of NBD2 was an exception to this trend, as Arg, a large positively charged amino acid, was

highly preferred at this site. The Pgp structure suggests that this Arg would be in close proximity to a negatively charged residue (Glu[269]) in the ICL2 that crosses over from the N-terminal TMD (blue) and docks into NBD2 (green; Figures 1 and 4B). This may explain the preference for Arg, which may form a salt bridge that would stabilize domain interactions. A salt bridge within ICL2 (between Arg[276] and Glu[256]) was recently found to be essential for maturation of human Pgp (ref) indicating that proper folding is necessary for docking of this loop into NBD2 [47]. Maintaining or even strengthening domain interactions can be crucial for assembly of multidomain proteins and their maturation to the cell surface as has been demonstrated for the related ABC transporter CFTR (cystic fibrosis transmembrane conductance regulator). For example, V510D in NBD1 of CFTR can form
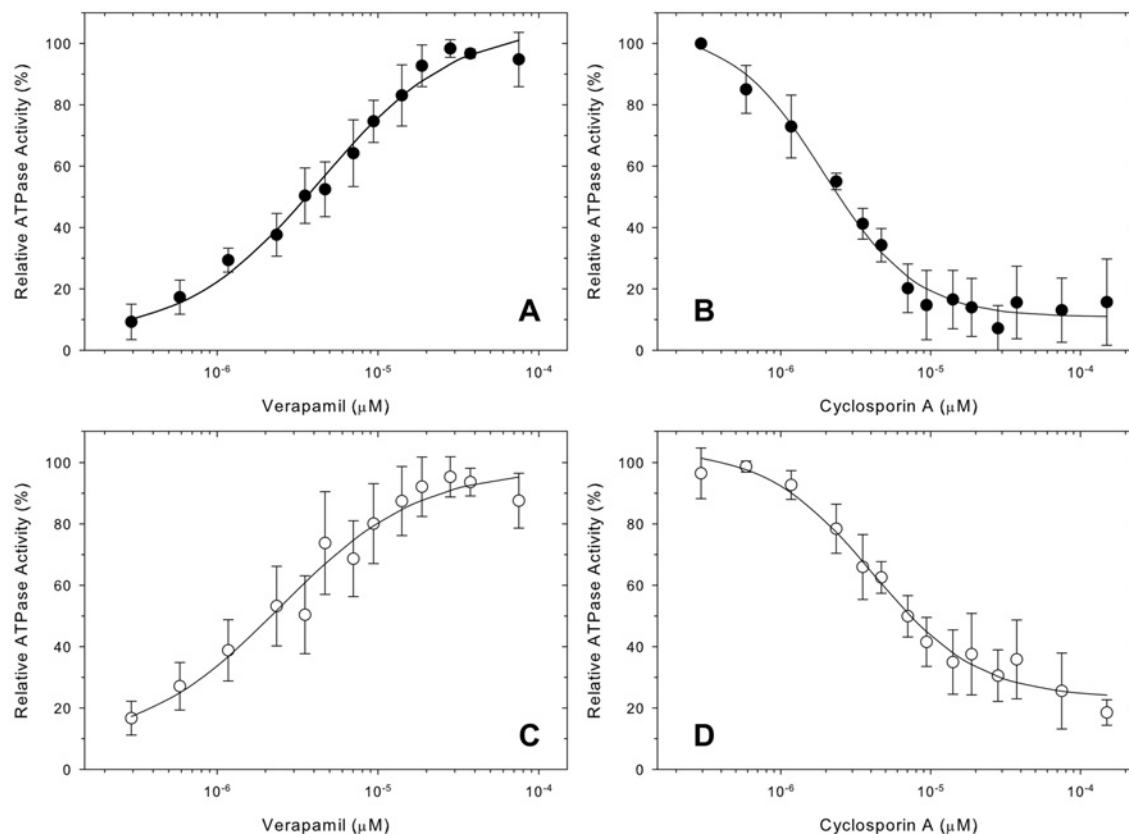
Figure 8   Stimulation and inhibition of ATPase activity of CL-Pgp and WT Pgp
(**A**) WT and (**C**) CL-Pgp ATPase activity was assayed with purified proteins in the presence of increasing concentrations of verapamil. The solid lines represent non-linear regression analyses to the data points (see the Materials and Methods section) No cooperativity was observed with Hill coefficients close to 1.0 (1.04 and 1.00, respectively). Each data point represents the mean from at least three independent experiments $\pm$ S.D.. (**B**) WT and (**D**) CL- Pgp purified proteins were assayed in the presence of 30 μM verapamil to maximally stimulate ATPase activity, but with increasing concentrations of the inhibitor cyclosporine A. The solid lines are non-linear regression fits, Hill coefficients were 1.40 and 1.22, respectively.

a salt bridge with Arg[1070] in ICL4 (crossing-over from the C-terminal TMD), promoting maturation and trafficking to the cell surface, and thereby can rescue the folding defect of the disease causing mutation ΔF508 [48,49].

Surprisingly, the two native Cys positions (Cys[638] and Cys[669]) that showed many revertants are both located in the linker that connects the N-terminal NBD to the C-terminal TMD, which is considered to be highly flexible with little secondary structure (Figure 1) [6]. The amino-acid sequence of the linker, including these two Cys, is not conserved among ABC transporters or Pgp orthologues, for example Asp and His are found in human Pgp (Supplementary Table S2 at http://www.bioscirep.org/bsr/034/bsr034e116add.htm). Previous work has demonstrated that human Pgp maintains some level of drug stimulated ATPase activity when the two halves of Pgp are coexpressed without being connected by the linker, but proper protein expression is severely impaired [50]. Additional work suggests that the linker may also have a regulatory function in Pgp [51]. The tendency to retain Cys at these two positions may indicate some importance for these two polar residues in mouse

Pgp, while other orthologues prefer charged Asp (34%) and Arg (49%) residues. The fact that Cys reversions occurred at these positions to a major extent indicates that site-saturation mutagenesis and stringent mutant selection was able to detect and successfully replace the Cys positions that were more difficult to substitute.

Although mutants generated by directed evolution were active, we did observe some variation in the degree of drug resistance among the C6Z mutants (Figure 5). The differing performance of the mutants between doxorubicin, FK506, and valinomycin may be explained by the mechanisms of action of these compounds. FK506, a calcineurin signalling inhibitor, and valinomycin, an ionophore, likely have more transitory effects on the physiology of the yeast cells than doxorubicin, which intercalates into DNA and interferes with DNA replication even at low intracellular concentrations [52]. Therefore doxorubicin may be more sensitive to assess how efficient Pgp mutants export drugs and prevent their intracellular accumulation.

In summary, this study demonstrated that directed evolution successfully yielded a CL Pgp that is well expressed, completely

active in *in vivo* yeast assays, and can be readily purified. Maintaining activity in yeast functional assays is a critical improvement of our new CL-Pgp that allow us to rapidly screen activity of new mutants where Cys are reinserted at strategic positions for biophysical studies. Furthermore, analysis of the frequency of amino-acid substitutions at a given Cys position, suggests that substitutions are highly dependent on the local environment of the native Cys. Our screen identified novel non-conservative Cys substitutions and revealed important, previously unnoticed domain interactions around Cys[1121]. The information gained from this study may be used to guide the construction of other CL proteins where Ala and Ser have proven unsuccessful.

REFERENCES

1   Eckford, P. D. and Sharom, F. J. (2009) ABC efflux pump-based resistance to chemotherapy drugs. Chem. Rev. **109**, 2989–3011 CrossRef PubMed

2   Dean, M., Rzhetsky, A. and Allikmets, R. (2001) The human ATP-binding cassette (ABC) transporter superfamily. Genome Res. **11**, 1156–1166 CrossRef PubMed

3   Szakacs, G., Paterson, J. K., Ludwig, J. A., Booth-Genthe, C. and Gottesman, M. M. (2006) Targeting multidrug resistance in cancer. Nat. Rev. Drug Discov. **5**, 219–234 CrossRef PubMed

4   Schinkel, A. H. and Jonker, J. W. (2003) Mammalian drug efflux transporters of the ATP binding cassette (ABC) family: an overview. Adv. Drug Deliv. Rev. **55**, 3–29 CrossRef PubMed

5   Gutmann, D. A. P., Ward, A., Urbatsch, I. L., Chang, G. and van Veen, H. W. (2010) Understanding polyspecificity of multidrug ABC transporters: closing in on the gaps in ABCB1. Trends Biochem. Sci. **35**, 36–42 CrossRef PubMed

6   Aller, S. G., Yu, J., Ward, A., Weng, Y., Chittaboina, S., Zhuo, R., Harrell, P. M., Trinh, Y. T., Zhang, Q., Urbatsch, I. L. and Chang, G. (2009) Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. Science **323**, 1718–1722 CrossRef PubMed

7   Ward, A. B., Szewczyk, P., Grimard, V., Lee, C.-W., Martinez, L., Doshi, R., Caya, A., Villaluz, M., Pardon, E., Cregger, C. et al. (2013) Structures of P-glycoprotein reveal its conformational flexibility and an epitope on the nucleotide-binding domain. Proc. Natl. Acad. Sci. U.S.A. **110**, 13386–13391 CrossRef PubMed

8   Dawson, R. J. and Locher, K. P. (2006) Structure of a bacterial multidrug ABC transporter. Nature **443**, 180–185 CrossRef PubMed

9   Ward, A., Reyes, C. L., Yu, J., Roth, C. B. and Chang, G. (2007) Flexibility in the ABC transporter MsbA: alternating access with a twist. Proc. Natl. Acad. Sci. U.S.A. **104**, 19005–19010 CrossRef PubMed

10 Crowley, E., McDevitt, C. A. and Callaghan, R. (2010) Generating inhibitors of P-glycoprotein: where to, now? Methods Mol. Biol. **596**, 405–432

11 Cha, A., Snyder, G. E., Selvin, P. R. and Bezanilla, F. (1999) Atomic scale movement of the voltage-sensing region in a potassium channel measured via spectroscopy. Nature **402**, 809–813 CrossRef PubMed

12 Pohl, T., Spatzal, T., Aksoyoglu, M., Schleicher, E., Rostas, A. M., Lay, H., Glessner, U., Boudon, C., Hellwig, P., Weber, S. and Friedrich, T. (2010) Spin labeling of the *Escherichia coli* NADH ubiquinone oxidoreductase (complex I). Biochim. Biophys. Acta **1797**, 1894–1900 CrossRef PubMed

13 Wen, P. C., Verhalen, B., Wilkens, S., McHaourab, H. S. and Tajkhorshid, E. (2013) On the origin of large flexibility of P-glycoprotein in the inward-facing state. J. Biol. Chem. **288**, 19211–19220 CrossRef PubMed

14 Cooper, R. S. and Altenberg, G. A. (2013) Association/dissociation of the nucleotide-binding domains of the ATP-binding cassette protein MsbA measured during continuous hydrolysis. J. Biol. Chem. **288**, 20785–20796 CrossRef PubMed

15 Liu, R. and Sharom, F. J. (1996) Site-directed fluorescence labeling of P-glycoprotein on cysteine residues in the nucleotide binding domains. Biochemistry **35**, 11865–11873 CrossRef PubMed

16 Loo, T. W. and Clarke, D. M. (1995) Membrane topology of a cysteine-less mutant of human P-glycoprotein. J. Biol. Chem. **270**, 843–848 CrossRef PubMed

17 Taylor, A. M., Storm, J., Soceneantu, L., Linton, K. J., Gabriel, M., Martin, C., Woodhouse, J., Blott, E., Higgins, C. F. and Callaghan, R. (2001) Detailed characterization of cysteine-less P-glycoprotein reveals subtle pharmacological differences in function from wild-type protein. Br. J. Pharmacol. **134**, 1609–1618 CrossRef PubMed

18 Crowley, E., O'Mara, M. L., Kerr, I. D. and Callaghan, R. (2010) Transmembrane helix 12 plays a pivotal role in coupling energy provision and drug binding in ABCB1. FEBS J. **277**, 3974–3985 CrossRef PubMed

19 Loo, T. W., Bartlett, M. C., Detty, M. R. and Clarke, D. M. (2012) The ATPase activity of the P-glycoprotein drug pump is highly activated when the N-terminal and central regions of the nucleotide-binding domains are linked closely together. J. Biol. Chem. **287**, 26806–26816 CrossRef PubMed

20 Lerner-Marmarosh, N., Gimi, K., Urbatsch, I. L., Gros, P. and Senior, A. E. (1999) Large scale purification of detergent-soluble P-glycoprotein from *Pichia pastoris* cells and characterization of nucleotide binding properties of wild-type, Walker A, and Walker B mutant proteins. J. Biol. Chem. **274**, 34711–34718 CrossRef PubMed

21 Urbatsch, I. L., Gimi, K., Wilke-Mounts, S., Lerner-Marmarosh, N., Rousseau, M. E., Gros, P. and Senior, A. E. (2001) Cysteines 431 and 1074 are responsible for inhibitory disulfide cross-linking between the two nucleotide-binding sites in human P-glycoprotein. J. Biol. Chem. **276**, 26980–26987 CrossRef PubMed

22 Tombline, G., Urbatsch, I. L., Virk, N., Muharemagic, A., White, L. B. and Senior, A. E. (2006) Expression, purification, and characterization of cysteine-free mouse P-glycoprotein. Arch. Biochem. Biophys. **445**, 124–128 CrossRef PubMed

23 Lee, S. H. and Altenberg, G. A. (2003) Transport of leukotriene C-4 by a cysteine-less multidrug resistance protein 1 (MRP1). Biochem. J. **370**, 357–360 CrossRef PubMed

24 van Iwaarden, P. R., Pastore, J. C., Konings, W. N. and Kaback, H. R. (1991) Construction of a functional lactose permease devoid of cysteine residues. Biochemistry (Mosc) **30**, 9595–9600 CrossRef

25 Holstead, R. G., Li, M. S. and Linsdell, P. (2011) Functional differences in pore properties between wild-type and cysteine-less forms of the CFTR chloride channel. J. Membr. Biol. **243**, 15–23 CrossRef PubMed

26 Hu, Y. K., Eisses, J. F. and Kaplan, J. H. (2000) Expression of an active Na,K-ATPase with an a-subunit lacking all twenty-three native cysteine residues. J. Biol. Chem. **275**, 30734–30739 CrossRef PubMed

27 Kuo, P. H., Ketchum, C. J. and Nakamoto, R. K. (1998) Stability and functionality of cysteine-less FOF1 ATP synthase from *Escherichia coli*. FEBS Lett. **426**, 217–220 CrossRef PubMed

28 Mense, M., Vergani, P., White, D. M., Altberg, G., Nairn, A. C. and Gadsby, D. C. (2006) *In vivo* phosphorylation of CFTR promotes formation of a nucleotide-binding domain heterodimer. EMBO J. **25**, 4728–4739 CrossRef PubMed

29 Arendt, C. S., Ri, K., Yates, P. A. and Ullman, B. (2007) Genetic selection for a highly functional cysteine-less membrane protein using site saturation mutagenesis. Anal. Biochem. **365**, 185–193 CrossRef PubMed

30 Bai, J. P., Swartz, D. J., Protasevich, II, Brouillette, C. G., Harrell, P. M., Hildebrandt, E., Gasser, B., Mattanovich, D., Ward, A., Chang, G. and Urbatsch, I. L. (2011) A gene optimization strategy that enhances production of fully functional P-glycoprotein in *Pichia pastoris*. PLoS ONE **6**, e22577 CrossRef PubMed

31 Green, M. R., Sambrook, J. and Sambrook, J. (2012) Molecular cloning: a laboratory manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y

32 Kwan, T. and Gros, P. (1998) Mutational analysis of the P-glycoprotein first intracellular loop and flanking transmembrane domains. Biochemistry **37**, 3337–3350 CrossRef PubMed

33 Beaudet, L. and Gros, P. (1995) Functional dissection of P-glycoprotein nucleotide-binding domains in chimeric and mutant proteins. Modulation of drug resistance profiles. J. Biol. Chem. **270**, 17159–17170 CrossRef PubMed

34 Raymond, M., Gros, P., Whiteway, M. and Thomas, D. Y. (1992) Functional complementation of yeast ste6 by a mammalian multidrug resistance mdr gene. Science **256**, 232–234 CrossRef PubMed

35 Vernet, T., Dignard, D. and Thomas, D. Y. (1987) A family of yeast expression vectors containing the phage f1 intergenic region. Gene **52**, 225–233 CrossRef PubMed

36 Urbatsch, I. L., Beaudet, L., Carrier, I. and Gros, P. (1998) Mutations in either nucleotide-binding site of P-glycoprotein (Mdr3) prevent vanadate trapping of nucleotide at both sites. Biochemistry **37**, 4592–4602 CrossRef PubMed

37 Urbatsch, I. L., Sankaran, B., Weber, J. and 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 CrossRef PubMed

38 Lopez-Llano, J., Campos, L. A. and Sancho, J. (2006) a-helix stabilization by alanine relative to glycine: roles of polar and apolar solvent exposures and of backbone entropy. Proteins–Struct. Funct. Bioinform. **64**, 769–778 CrossRef

39 Raymond, M., Ruetz, S., Thomas, D. Y. and Gros, P. (1994) Functional expression of P-glycoprotein in *Saccharomyces cerevisiae* confers cellular resistance to the immunosuppressive and antifungal agent FK520. Mol. Cell. Biol. **14**, 277–286 PubMed

40 Swartz, D. J., Weber, J. and Urbatsch, I. L. (2013) P-glycoprotein is fully active after multiple tryptophan substitutions. Biochim. Biophys. Acta – Biomembr. **1828**, 1159–1168 CrossRef

41 Hani, J. and Feldmann, H. (1998) tRNA genes and retroelements in the yeast genome. Nucleic Acids Res. **26**, 689–696 CrossRef PubMed

42 Kodan, A., Shibata, H., Matsumoto, T., Terakado, K., Sakiyama, K., Matsuo, M., Ueda, K. and Kato, H. (2009) Improved expression and purification of human multidrug resistance protein MDR1 from baculovirus-infected insect cells. Protein Expr. Purif. **66**, 7–14 CrossRef PubMed

43 Verhalen, B., Ernst, S., Borsch, M. and Wilkens, S. (2012) Dynamic ligand-induced conformational rearrangements in P-glycoprotein as probed by fluorescence resonance energy transfer spectroscopy. J. Biol. Chem. **287**, 1112–1127 CrossRef PubMed

44 Urbatsch, I. L., Julien, M., Carrier, I., Rousseau, M. E., Cayrol, R. and Gros, P. (2000) Mutational analysis of conserved carboxylate residues in the nucleotide binding sites of P-glycoprotein. Biochemistry **39**, 14138–14149 CrossRef PubMed

45 Henikoff, S. and Henikoff, J. G. (1992) Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. U.S.A. **89**, 10915–10919 CrossRef PubMed

46 Smith, P. C., Karpowich, N., Millen, L., Moody, J. E., Rosen, J., Thomas, P. J. and Hunt, J. F. (2002) ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimer. Mol. Cell **10**, 139–149 CrossRef PubMed

47 Loo, T. W. and Clarke, D. M. (2013) A salt bridge in intracellular loop 2 is essential for folding of human P-glycoprotein. Biochemistry **52**, 3194–3196 CrossRef PubMed

48 Loo, T. W., Bartlett, M. C. and Clarke, D. M. (2010) The V510D suppressor mutation stabilizes delta F508-CFTR at the cell surface. Biochemistry (Mosc). **49**, 6352–6357 CrossRef

49 Dawson, J. E., Farber, P. J. and Forman-Kay, J. D. (2013) Allosteric coupling between the intracellular coupling Helix 4 and regulatory sites of the first nucleotide-binding domain of CFTR. PLoS ONE **8**, e74347 CrossRef PubMed

50 Loo, T. W. and Clarke, D. M. (1994) Reconstitution of drug-stimulated ATPase activity following coexpression of each half of human P-glycoprotein as separate polypeptides. J. Biol. Chem. **269**, 7750–7755 PubMed

51 Sato, T., Kodan, A., Kimura, Y., Ueda, K., Nakatsu, T. and Kato, H. (2009) Functional role of the linker region in purified human P-glycoprotein. FEBS J. **276**, 3504–3516 CrossRef PubMed

52 Foor, F., Parent, S. A., Morin, N., Dahl, A. M., Ramadan, N., Chrebet, G., Bostian, K. A. and Nielsen, J. B. (1992) Calcineurin mediates inhibition by FK506 and cyclosporin of recovery from a-factor arrest in yeast. Nature **360**, 682–684 CrossRef PubMed

OPEN ACCESS

SUPPLEMENTARY DATA

Directed evolution of P-glycoprotein cysteines reveals site-specific, non-conservative substitutions that preserve multidrug resistance

Douglas J. SWARTZ*†[1], Leo MOK*†[1], Sri K. BOTTA*†, Anukriti SINGH*†, Guillermo A. ALTENBERG†‡ and Ina L. URBATSCH*†[2]

*Department of Cell Biology and Biochemistry, Texas Tech University Health Sciences Center, Lubbock, Texas, U.S.A.
†Center for Membrane Protein Research, Texas Tech University Health Sciences Center, Lubbock, Texas, U.S.A.
‡Department of Cell Physiology and Molecular Biophysics, Texas Tech University Health Sciences Center, Lubbock, Texas, U.S.A.

Table S1 Oligonucleotide sequences used for site-saturation mutagenesis and site-directed mutagenesis of single TM Cys mutants
Only the forward mutagenic primer sequences are shown. Reverse primer sequences are identical but complimentary to the forward primer.

Primer name	Sequence (5′ to 3′)
Mutagenic primers	
C427-NNN	CGGAAATTCTGGTNNNGGCAAGTCCAC
C638-NNN	GGAAATGAGGCTNNNAAGTCTAAGGACGAG
C669-NNN	CACTAGAAAGTCCATTNNNGGTCCACACGATC
C1070-NNN	GGTTCTTCCGGTNNNGGCAAGTCCACTG
C1121-NNN	CTATTCTTTTCGACNNNAGCATTGCAGAG
C1223-NNN	CCAGAGAAGGAAGAACTNNNATCGTCATTGCTC
C133A	TCCTTCTGGGCCCTGGCCGCTGGAAGACAAATCCACAAGATTAGACAGAAG
C713A	GGTATCTTCGCTGCCATCATTAACGGTGGTTTGCAGCCAGCTTTTTCTGTG
C952A	TACGCTGCCGCATTCAGATTCGGTGCTTATTTGGTGACTCAGCAATTGATG
Flanking primers	
NBD1/linker	CCTTGTTTCTTTTTCTGCACAA (Forward)
	CTGTCTTTGGGTTTCTGGTGGACCACC (Reverse)
NBD2	CCATACCGTAACGCTATGAAAAAGGCTC (Forward)
	GAAGTGTCAACAACGTATCTACCAACG (Reverse)

[1] These authors contributed equally to this work.
[2] To whom correspondence should be addressed (email ina.urbatsch@ttuhsc.edu).

Table S2 Relative frequency of each amino acid found at the nine native Cys positions in closely related Pgp orthologs

Position	Cys[133]*	Cys[427]	Cys[638]	Cys[669]	Cys[713]	Cys[952]	Cys[1070]	Cys[1121]	Cys[1223]
A									
C	84	100	11	5	98	98	100	96	98
D			34						
E			5						
F								4	
G	2		23						
H				23					
I									
K				16					
L	2								
M									
N			7						
P									
Q				2					
R				49					2
S			14	2	2	2			
T	11			2					
V									
W									
Y			7						
Total%	100	100	100	100	100	100	100	100	100

*Human and mouse Pgp were aligned with 46 closest related orthologs (⩾70% sequence identity) from the Ensembl database. The relative frequency (%) of amino acids found at each Cys position is shown.