Overcoming Multidrug Resistance: search for human P-glycoprotein modulators and insights on drug efflux mechanism

Cátia Alexandra Marques Bonito Ferreira

Programa Doutoral em Química
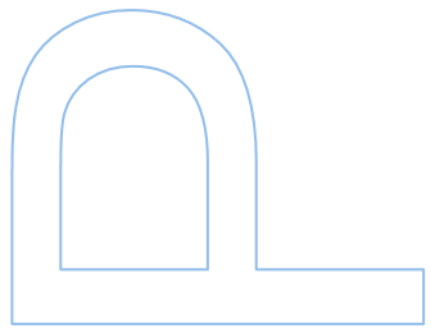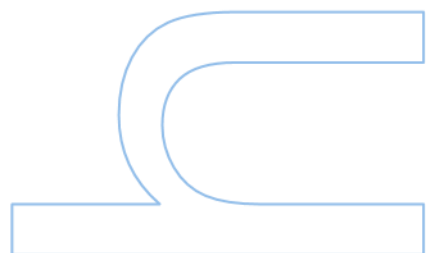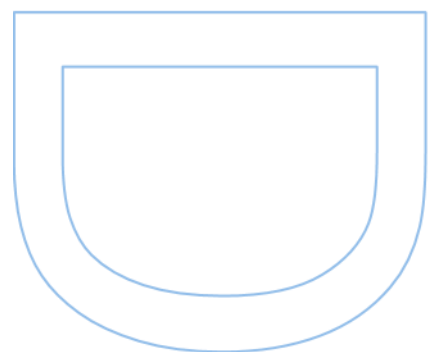Departamento de Química e Bioquímica
2023

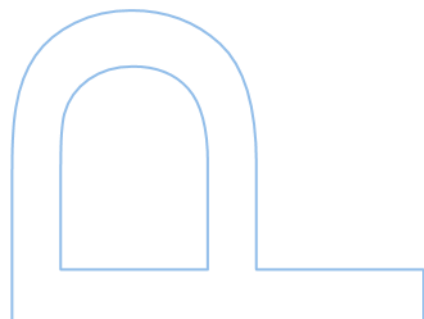**Orientador**
Daniel José Viegas Antunes dos Santos, Investigador Auxiliar,
Universidade Lusófona.

Co-orientadores
Maria Natália Dias Soeiro Cordeiro, Professora Associada,
Faculdade de Ciências da Universidade do Porto

Jean-Pierre Gillet, Professor,
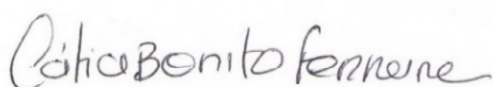Faculdade de Medicina, Universidade de Namur, Bélgica

Declaração de Honra

Eu, Cátia Alexandra Marques Bonito Ferreira, inscrito(a) no Programa Doutoral em Química da Faculdade de Ciências da Universidade do Porto declaro, nos termos do disposto na alínea a) do artigo 14.º do Código Ético de Conduta Académica da U.Porto, que o conteúdo da presente tese reflete as perspetivas, o trabalho de investigação e as minhas interpretações no momento da sua entrega.

Ao entregar esta tese, declaro, ainda, que a mesma é resultado do meu próprio trabalho de investigação e contém contributos que não foram utilizados previamente noutros trabalhos apresentados a esta ou outra instituição.

Mais declaro que todas as referências a outros autores respeitam escrupulosamente as regras da atribuição, encontrando-se devidamente citadas no corpo do texto e identificadas na secção de referências bibliográficas. Não são divulgados na presente tese quaisquer conteúdos cuja reprodução esteja vedada por direitos de autor.

Tenho consciência de que a prática de plágio e auto-plágio constitui um ilícito académico.

Eslöv, 15 de Setembro de 2022

Agradecimentos/ Acknowledgments

Quero expressar os meus sinceros agradecimentos ao meu orientador Doutor Daniel dos Santos, Investigador Auxiliar, CBIOS- Research Center for Biosciences & Health Technologies, Universidade Lusófona de Humanidades e Tecnologias, e à minha co-orientadora Professora Doutora Maria Natália Cordeiro, professora Associada do Departamento de Química e Bioquímica da Faculdade de Ciências da Universidade do Porto (FCUP) pela dedicação que demonstraram ao longo destes 4 anos de tese, e pela confiança que depositaram em mim durante a realização do projeto.

I also express my gratitude to my co-supervisor Professor Doctor Jean-Pierre Gillet from the Laboratory of Molecular Cancer Biology, Faculty of Medicine, University of Namur, Belgium, for accepting me as his Ph.D. student. Professor Jean-Pierre demonstrated in several ways a great dedication to this work during these years. I also want to recognize the commitment expressed by Professor Jean-Pierre in the preparation for my internship in Belgium, and I want to acknowledge the warm reception that I had there. I must also recognize the willingness of Professor Jean-Pierre in doing the biological assays that, unfortunately, I could not perform, due to the beginning of the pandemic. I want to also give a special acknowledgment to Simon Lefevre (LBMC, UNamur) for his technical support in gathering the ATPase assays data, and all the other members of the team involved in the logistics of the ATPase assay Systems.

Quero agradecer à Doutora Goreti Carneiro, técnica superior do Departamento de Química e Bioquímica da Faculdade da FCUP, pela sua incansável dedicação e rapidez na logística de aquisição e envio tanto de compostos como de material necessário para efetuar os ensaios biológicos na Bélgica.

Agradeço à Professora Doutora Maria José U. Ferreira, professora Associada com Agregação no Grupo de Produtos Naturais da Faculdade de Farmácia da Universidade de Lisboa, e à Professora Doutora Maria Emília Sousa, professora Associada no Departamento de Ciências Químicas da Faculdade de Farmácia da Universidade do Porto, pela amabilidade que ambas demonstraram em disponibilizar compostos sintetizados in-house para poderem ser testados nos ensaios biológicos na Bélgica. Desta forma, agradeço também aos meus colegas Ricardo J. Ferreira e Fernando Durães pela síntese dos mesmos. Ao meu colega Ricardo J. Ferreira também quero agradecer todo o apoio técnico de âmbito computacional que ofereceu durante o desenvolvimento desta tese.

Quero agradecer aos meus filhos, à minha família e aos meus amigos pelo apoio, pelo carinho e pela motivação durante a realização e a escrita da tese.

Por último, quero tecer um profundo agradecimento ao meu marido Ricardo Ferreira, pelo incansável apoio a vários níveis para que eu pudesse concluir esta tese. Perante um contexto de pandemia após o segundo ano de doutoramento e com dois filhos, presentemente com 4 e 11 anos, o plano de trabalhos da tese foi em algumas ocasiões difícil de realizar! Contudo, devo reconhecer todo o carinho e paciência que o meu marido demonstrou, especialmente em aturar o meu mau humor, tenho de confessar! O meu marido foi em termos pessoais o grande pilar emocional para a realização da minha tese... muito obrigada Ricardo!

ABSTRACT

Multidrug resistance (MDR) to anticancer drugs is one of the major contributor to chemotherapy failure. Among others, the over-expression of the ATP-binding cassette (ABC) transporters such as P-glycoprotein (P-gp, ABCB1) is one of the most significant mechanisms in MDR cancer cells, thus contributing to decrease the intracellular concentration of anticancer drugs and leading to unsuccessful chemotherapeutic regimens. Hence, modulating drug efflux by P-gp is still considered one of the most promising strategies to overcome MDR in cancer.

Nevertheless, and despite decades of research, the mechanisms of drug recognition/specificity and efflux-related signal-transmission between the transmembrane domains (TMDs) and nucleotide-binding domains (NBDs) are still unsolved, thus contributing to impair the development of clinical modulators. Additionally, the poly-specificity of the drug-binding pocket (DBP) is considered by many the biggest drawback for the development of more selective and effective inhibitors. This way, other strategies that avoid competition within the DBP are required for the development of a novel generation of efflux modulators. Thus, the main goals of this work were to contribute with additional knowledge concerning the molecular mechanisms underlying drug promiscuity and TMD-NBD communication, and to explore a new strategy for P-gp modulation by a non-competitive mechanism.

To achieve such goals the impact of four P-gp mutations (G185V, G830V, F978A and ∆F335), located at the TMDs, and experimentally related to changes in drug binding and efflux, in the P-gp architecture was thoroughly characterized by means of Molecular Dynamics simulations. The studies were performed in the absence and presence of molecules, that experimentally showed altered efflux patterns in the respective variants. From these studies it is possible to conclude that perturbations at the transmembrane region as the presence of mutations and/or drugs, may induce a repacking of the transmembrane helices (TMHs), that in turn, lead to changes in the drug-binding sites (DBSs) physical properties. This changes the way how molecules bind to P-gp, which suggests that the TMHs repacking may be one of the molecular mechanisms underlying P-gp poly-specificity. Additionally, long-range effects at the TMD-NBD interfaces were observed as a result of the TMHs repacking, in particular changes in the residues interactions between the intracellular coupling helices (ICHs) and the NBDs.

Considering the notable involvement of the ICHs in TMD-NBD signal-transmission mechanism, targeting such motifs with small molecules, is considered to be an

alternative approach for overcoming MDR. Therefore, by means of computational Fragment-Based Drug Discovery two allosteric DBSs (aDBSs) were identified, located in-between the ICHs and the A-loop, Q-loop and Walker A motifs of the respective NBD. Following, the putative aDBSs were characterized in terms of size, polarity, and lining residues. The results indicate that both aDBSs have similar volumes but, nevertheless, are smaller and more polar than the modulator site (M-site), located at the top of the canonical DBP (outer leaflet). Furthermore, Molecular Docking studies indicate that in these hypothetical aDBSs, small molecules are allowed to bind as thioxanthone and flavanone derivatives, with favorable binding energies. Interestingly, most of the compounds evaluated in the ATPase assays proved to successfully decrease the verapamil-stimulated P-gp ATPase activity, thus inferring that these aDBSs are indeed druggable and that molecules that target these motifs are able to modulate drug efflux.

Finally, this work demonstrates the "proof-of-concept" that P-gp can be modulated by an allosteric mechanism, having the great advantage to avoid competition with substrates at the DBP. One flavanone derivative was further identified as possible building block for the design of allosteric P-gp modulators. Thereby, this work opens new perspectives for the development of small molecules able to inhibit P-gp by a non-competitive mechanism. The elaboration of a forth-generation of MDR modulators is, therefore, possible, in this case highly specific for P-gp but nonetheless transferable to other efflux pumps of the ABC transporter superfamily, while designed to improve the pharmacokinetic limitations shown by the previous generations of MDR reversal agents.

KEYWORDS

RESUMO

A resistência a múltiplos fármacos anti-tumorais é um dos principais fatores que contribuem para o fracasso da quimioterapia. Entre outros, a sobre-expressão de transportadores da família ABC tais como a glycoproteína-P (P-gp, ABCB1), é um dos mecanismos mais significativos presente nas células tumorais multi-resistentes, contribuindo desta forma para diminuir a concentração intracelular de fármacos anti-cancerígenos e inviabilizando o sucesso dos esquemas de quimioterapia. Assim, a modulação do efluxo de fármacos pela P-gp é uma das estratégias mais promissoras para vencer a multi-resistência no cancro.

No entanto, e apesar de décadas de investigação, os mecanismos de reconhecimento/especificidade de fármacos e de transmissão de sinal entre os domínios transmembranares (TMDs) e os domínios de ligação a nucleótidos (NBDs), intimamente relacionado com o mecanismo de efluxo, ainda não foram solucionados, dificultando desta forma forma o desenvolvimento de moduladores com grande afinidade para a P-gp. Além disso, a poli-especificidade do local de ligação de fármacos/xenobióticos (DBP) à P-gp, é considerada por muitos a maior desvantagem para o desenvolvimento de moduladores mais seletivos e eficazes para esta bomba de efluxo. Sendo assim, para o desenvolvimento de uma nova geração de moduladores do efluxo, são necessárias outras estratégias que evitem a competição de fármacos dentro do DBP. Neste contexto, o principais objetivos deste trabalho foram contribuir com conhecimento adicional sobre os mecanismos moleculares subjacentes à promiscuidade de fármacos na P-gp e à comunicação entre os TMDs e NBDs, e explorar uma nova estratégia para a modulação da P-gp, idealmente através de um mecanismo não competitivo.

Para atingir tais objetivos, o impacto induzido por quatro mutações (G185V, G830V, F978A e ∆F335), localizadas nos TMDs, e experimentalmente relacionadas com alterações na ligação e efluxo de fármacos, na arquitetura da P-gp foi minuciosamente caracterizado por meio de simulações de Dinâmica Molecular. Os estudos foram realizados na ausência e presença de moléculas, que experimentalmente mostraram padrões de efluxo alterados nas respetivas variantes. A partir desses estudos é possível concluir que perturbações nos TMDs tais como a presença de mutações e/ou ligandos induzem um re-empacotamento das hélices transmembranares (TMHs), que por sua vez, levam a alterações nas propriedades físicas dos locais de ligação de fármacos (DBSs) à P-gp. Estes eventos alteram a forma como os fármacos se ligam à P-gp, o que indica que o re-empacotamento das TMHs pode ser um dos mecanismos moleculares

subjacentes à poli-especificidade da P-gp. Adicionalmente, foram observados efeitos de longo alcance nas interfaces TMD-NBD como resultado do re-empacotamento das TMHs, em particular, alterações nas interações entre os resíduos das hélices intra-celulares de acoplamento (ICHs) e os resíduos dos NBDs.

Considerando o notável envolvimento das ICHs no mecanismo de transmissão de sinal entre os TMDs e os NBDs, o desenho de moléculas capazes de interagir especificamente com tais motivos, é considerado por muitos uma abordagem alternativa para reverter a multi-resistência. Portanto, por meio da Descoberta Computacional de Fármacos Baseada em Fragmentos foram identificados dois DBSs alostéricos (aDBSs) na P-gp, entre as ICHs e os motivos A-loop, Q-loop e Walker A do respetivo NBD. Além disso, os hipotéticos aDBSs foram caracterizados em termos de tamanho, polaridade e o tipo de resíduos que delimitam o possível local de ligação, sendo os resultados indicadores de que ambos os aDBSs possuem volumes semelhantes, mas porém, são menores e mais polares que o sítio modulador (sítio M) localizado no topo do DBP. Além disso, as energias de ligação obtidas através de estudos de Docking Molecular corroboram que pequenas moléculas, tais como derivados das tioxantonas e dos flavonóides, são capazes de se ligar aos possíveis aDBSs com energias de ligação favoráveis. Curiosamente, a maioria dos compostos avaliados nos ensaios experimentais de ATPase demonstraram reduzir com sucesso a atividade ATPase estimulada pelo verapamil, inferindo desta forma, que esses aDBSs são considerados potenciais alvos para fármacos, e que moléculas que visam esses motivos são capazes de modular o efluxo de drogas.

Por fim, este trabalho demonstra uma "prova de conceito" de que a P-gp pode ser modulada por um mecanismo alostérico, apresentando a grande vantagem de evitar a competição com substratos no DBP. Foi ainda identificado um derivado da narigenina (IC$_{50}$ 81 ± 6.6 μM) como um possível composto de base para o desenvolvimento de moduladores alostéricos. Desta forma, este trabalho abre novas perspetivas para a procura de moléculas pequenas, capazes de inibir a P-gp por um mecanismo não competitivo. O desenvolvimento da próxima geração de moduladores da MDR é, como tal, possível, neste caso altamente específico para P-gp, mas ainda assim transferível para outras bombas de efluxo da família dos transportadores ABC, enquanto projetado para melhorar as limitações farmacocinéticas mostradas pelo anterior gerações de agentes de reversão MDR.

PALAVRAS-CHAVE

Multi-resistência, glicoproteína-P, inibidores específicos do efluxo, transmissão de sinal, hélices intra-celulares, locais alostéricos, tioxantonas, flavanonas, mecanismo alostérico, modelação por homologia, Dinâmica Molecular, Docking Molecular, Descoberta de Fármacos Baseada em Fragmentos, Ensaios de ATPase.

PREAMBLE

The work presented in this thesis is divided in 6 chapters.

Chapter 1 - Introduction

This chapter starts with a brief overview about the multidrug-resistance in cancer, and the important role that over-expression of membrane efflux pumps, such as P-glycoprotein (P-gp), has in cancer cells survival. A review about the most important research done in P-gp in the recent decades is also presented, emphasizing the discovery of a new target region within the P-gp architecture and potential scaffolds that would be suitable for modulating P-gp, when bound at this new location and hipotetically through an allosteric mechanism. Finally, this chapter ends with the goals of this PhD project.

The information contained in this chapter was adapted from the paper

Ferreira RJ, Bonito CA, Ferreira MJ-U, dos Santos DJVA. About P-glycoprotein: a new drugable domain is emerging from structural data. *WIREs Comput. Mol. Sci.* **2017,** *7, e1316* (doi: https://doi.org/10.1002/wcms.1316).

and from the preprint

Bonito CA, Ferreira MJU, Ferreira RJ, dos Santos DJVA. Does human P-glycoprotein efflux involve transmembrane alpha helix breakage? *bioRxiv.* **2019,** 692988 (doi: https://doi.org/10.1101/692988)

Chapter 2 – Methodology

This chapter summarizes all of the computational and experimental methods applied in this work, and the rationale underlying their usage.

Chapters 3, 4 and 5

These chapters describe and discuss the main findings obtained during this work. Each chapter corresponds to publications published in journals with peer review (Chapter 3), or in preprint servers (Chapters 4 and 5):

Chapter 3 – Paper I: "Theoretical insights on helix repacking as the origin of P-glycoprotein promiscuity."

Bonito CA, Ferreira RJ, Ferreira M-JU, Gillet J-P, Cordeiro MNDS, dos Santos DJVA. Theoretical insights on helix repacking as the origin of P-glycoprotein promiscuity. *Sci. Rep.* **2020,** *10, 9823.* (doi: https://doi.org/10.1038/s41598-020-66587-5)

Chapter 4 - Paper II: "Long-range communication between transmembrane- and nucleotide-binding domains does not depend on drug binding to mutant P-glycoprotein."

Bonito CA, Ferreira RJ, Ferreira M-JU, Gillet J-P, Cordeiro MNDS, dos Santos DJVA *bioRxiv 2022.06.30.498271*. (doi: https://doi.org/10.1101/2022.06.30.498271; *submitted a peer-reviewed journal*)

Chapter 5 - Letter: "Probing the allosteric modulation of P-glycoprotein: A medicinal chemistry approach towards the identification of non-competitive P-gp inhibitors."

Bonito CA, Ferreira RJ, Ferreira M-JU, Durães F, Sousa E, Gillet J-P, Cordeiro MNDS, dos Santos DJVA. *ResearchSquare* **2022,** *12,* September 22 PREPRINT (doi: https://doi.org/10.21203/rs.3.rs-2038969/v1; *submitted a peer-reviewed journal*)

Chapter 6 - Conclusions

This chapter summarizes the main knowledge gathered during this work concerning the mechanisms of drug specificity and efflux-related signal-transmission. Moreover, a "proof-of-concept" that P-gp can be modulated by an allosteric mechanism is also presented.

References

This section consists in a list of all bibliographic references used along the thesis and sorted by alphabetic order.

Appendix

This section includes, the final PDF format of the published paper in a journal with peer review (*Scientific Reports*, *Annex S1*), and the *Supporting Information* files (Annexes S2-S4) regarding the three publications produced during the PhD project.

Table of Contents

Index of Tables

Index of Figures

FCUP

Overcoming Multidrug Resistance: search for human P-glycoprotein modulators and insights on drug efflux mechanism | xviii

LIST OF ABBREVIATIONS

ABC	ABC transporter superfamily
aCAP	Cyclic peptide P-gp inhibitor
ACT	Actinomycin D
aDBS(s)	Allosteric drug-binding site(s)
AM1-BCC	Austin Model 1 with bond charge correction
ATP	Adenosine Triphosphate
BCRP	Breast Cancer Resistance Protein
BDE-100	Polybrominated diphenyl ether 100
BUM	2-[4-(diethylamino)-2-hydroxybenzoyl]benzoic acid
C. elegans	*Caenorhabditis elegans*
CE	Contact efficiency ratio
CFTR	Cystic Fibrosis Transmembrane Regulator
cFBDD	computational Fragment-Based Drug Design
COL	Colchicine
Cryo-EM	Criogenic electron microscopy
Cα	Alpha carbon(s)
CYC	Cyclosporin A
DBP	Drug-binding pocket
DBS(s)	Drug-binding site(s)
DHP	Dihydropyridine
DMSO	Dimethylsulfoxide
DOX	Doxorrubicin
ECL	Extracellular loop
ECS	Extracellular segment
GROMACS	GROningen MAchine for Chemical Simulation
H-site	H33342-binding site
H33342	Hoescht 33342
IC$_{50}$	Half maximal inhibitory concentration
ICH(s)	Intracellular coupling helix(ces)
ICS	Intracellular segment
IF	Inward-facing
K	Kelvin
LRET	Luminescense resonance energy transfer
M-site	Modulator-binding site
MD	Molecular Dynamics

MDR	Multidrug-resistance
mL	mililiter
mM	milimolar
MMFF94x	Merck molecular force field 94x
MOE	Molecular Operating Environment
MRP1	Multidrug resistance protein 1
Na_3VO_4	Sodium orthovanadate
NBD(s)	Nucleotide-binding domain(s)
NBD-CsA	[N-ε-(4-nitrobenzofurazan)-7-yl]-dLys8-cyclosporin A
NBS(s)	Nucleotide-binding site(s)
NPA	1-*N*-naphtylphtalamic acid
nm	nanometer
ns	nanosecond
OF	Outward-facing
OPM	Orientation of Proteins in Membranes
PBC	Periodic Boundary Conditions
PDB	Protein Data Bank
P-gp	P-glycoprotein
hP-gp	human P-glycoprotein (*Homo sapiens*)
mP-gp	murine P-glycoprotein (*Mus musculus)*
PME	Particle Mesh Ewald
PMF	Potential of Mean Force
POPC	1-palmitoyl-2-oleoyl-phosphatidylcholine
ps	Picosecond
P(V)	Probability distribution function (Volume)
R-site	R123-binding site
R123	Rhodamine-123
RLU	Relative light units
RMSD	Root mean square deviation
SaMD	Simulated-annealing Molecular Dynamics
Sav1866	*Staphylococcus aureus* multidrug ABC transporter
SBS	Substrate-binding sites(s)
SPC	Single-point charge water model
TC	Test compound
TMD(s)	Transmembrane domain(s)
TMH(s)	Transmembrane helix(ces)
VER	Verapamil

VIN	Vinblastine
VLS	Valspodar (PSC833)
VMD	Visual Molecular Dynamics
WA	Walker A
WT	wild-type
Å	Ångstrom
Δ*G*	free-energy of binding
µg	Microgram
µM	Micromolar

CHAPTER 1 - INTRODUCTION

1.1 Multidrug resistance in cancer

Multidrug resistance (MDR) to anticancer drugs remains as one of the largest threats for the success of chemotherapy regimens worldwide. As this phenomenon often extends to structurally distinct chemotherapeutic agents, MDR is one of the major contributors for cancer treatment failures.[1] Although a multifactorial process (**Figure 1.1**), the over-expression of the ATP-binding cassette (ABC) transporters such as P-glycoprotein (P-gp, ABCB1),[2] Multidrug Resistance Protein 1 (MRP1, ABCC1),[3] the Breast Cancer Resistance Protein (BCRP, ABCG2),[4-6] and ABCB5,[7,8] represent one of the most relevant mechanisms in cancer cells survival. Currently, 49 ABC transporter members were identified in the human genome and classified in seven different subfamilies, ABCA to ABCG.[9] Among them, 13 ABC transporter members seem to be involved in MDR in cancer namely, ABCA1/2, ABCB1/2/5, ABCC1/2/3/4/5/6/10 and ABCG2, mediating the efflux of anticancer drugs to the extracellular space, and thus lowering their concentration inside the cell.[10] Facing the notable relevance that ABC transporters have in cancer cells, the modulation of drug efflux by small molecules is of the utmost importance to overcome MDR and to improve the success of the chemotherapy treatments.[11]
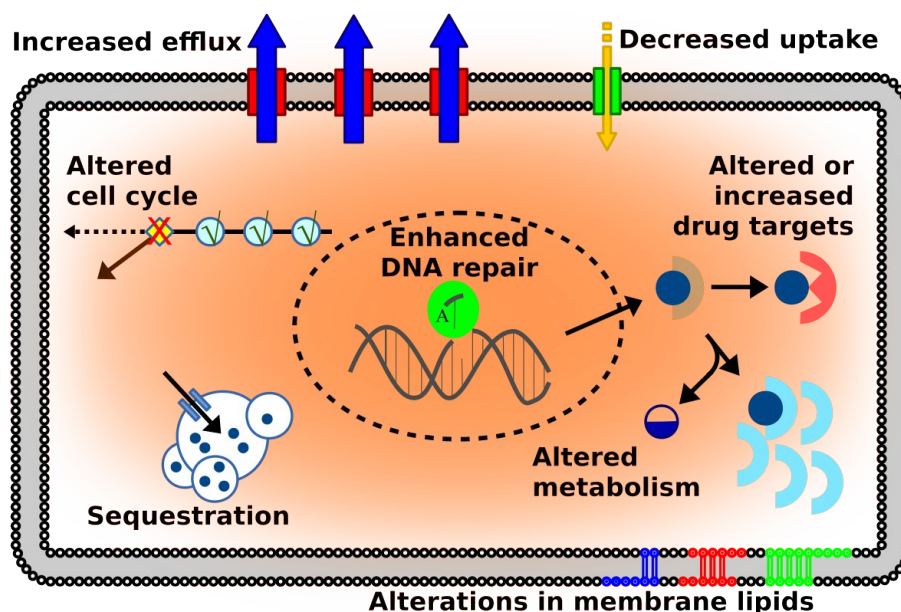
Figure 1.1. Mechanisms of drug resistance (adapted from Gottesmann et al., 2002).[12]

1.2 P-gp architecture and drug efflux: Insights from structural data

1.2.1 Crystallographic structures

P-glycoprotein is one of the most studied efflux pumps to date. Firstly identified in 1976 in Chinese ovary cell mutants, it was promptly associated with altered drug permeation

in cancer cells.[2] Despite having a higher relevance in MDR, P-gp also plays a key role in cell detoxification, and it is constitutively expressed in many organs as liver, kidney, gut, and pancreas as well as in the blood–brain and placental barriers.

Only after the publication of the bacterial ABC transporters MsbA in 2001 (retracted but later corrected)[13,14] and Sav1866[15] X-ray structures, new insights on the ABC architecture and how the efflux mechanism could occur were acquired. Yet, the publication of the first murine P-gp crystallographic structure in 2009[16] was considered an important hallmark, contributing with new information about the P-gp structure, efflux mechanism and drug specificity.

From its crystallographic structural data (PDB ID: 3G5U), a pseudo two-fold symmetry was noticeable, much similar to that described for Sav1866. Two homologous functional units (N- and C- terminal halves) comprising six α-helical bundles spawning the membrane (transmembrane domains, TMDs), while the cytoplasmic nucleotide-binding domains (NBDs), where the adenosine triphosphate (ATP) binds and hydrolysis occurs, revealed a head-to-tail organization. Connecting the transmembrane and cytoplasmic domains, four short intracellular coupling helices (ICHs) between the transmembrane helices (TMHs) 2/3 (ICH1-NBD1), 4/5 (ICH2-NBD2), 8/9 (ICH3-NBD2) and 10/11 (ICH4-NBD1) were additionally identified. The TMDs harbor the drug-binding pocket (DBP) – a large internal cavity (~ 6000 Å^3), opened to the cytoplasm and to the lipid bilayer through two portals formed by the TMHs 4/6 and 10/12 (**Figure 1.2**).[16] The co-crystallization of the murine P-gp structure with two cyclic peptide inhibitors (isomers QZ59-SSS and QZ59-RRR) (PDB IDs: 3G60 and 3G61) confirmed the existence of a specific drug-binding site (DBS), mostly containing hydrophobic and aromatic residues, located at an upper position within the large cavity, in agreement with previous biochemical studies.[17,18] Interestingly, many of the residues described by the authors as interacting with QZ59 ligands had already been previously reported by several biochemical studies using verapamil, the gold-standard P-gp efflux modulator.[19-21] Although obtained in a nucleotide-free inward-facing (IF) conformation, where the DBP is opened to the cytoplasm and the NBDs are separated, the transporter was still thought to be competent for substrate binding. This lead the authors to propose that the IF conformation may represent only one of the initial steps in the catalytic cycle.[16] Later, four new murine P-gp crystallographic structures, co-crystallized with several QZ-ligands (PDB IDs: 4Q9I, 4Q9J, 4Q9K and 4Q9L),[22] were obtained and made available, again corroborating the existence of the previously identified hydrophobic DBS at the top of DBP.[16] Additionally, large conformational changes in TMH4 were also observed upon ligand binding.[22]
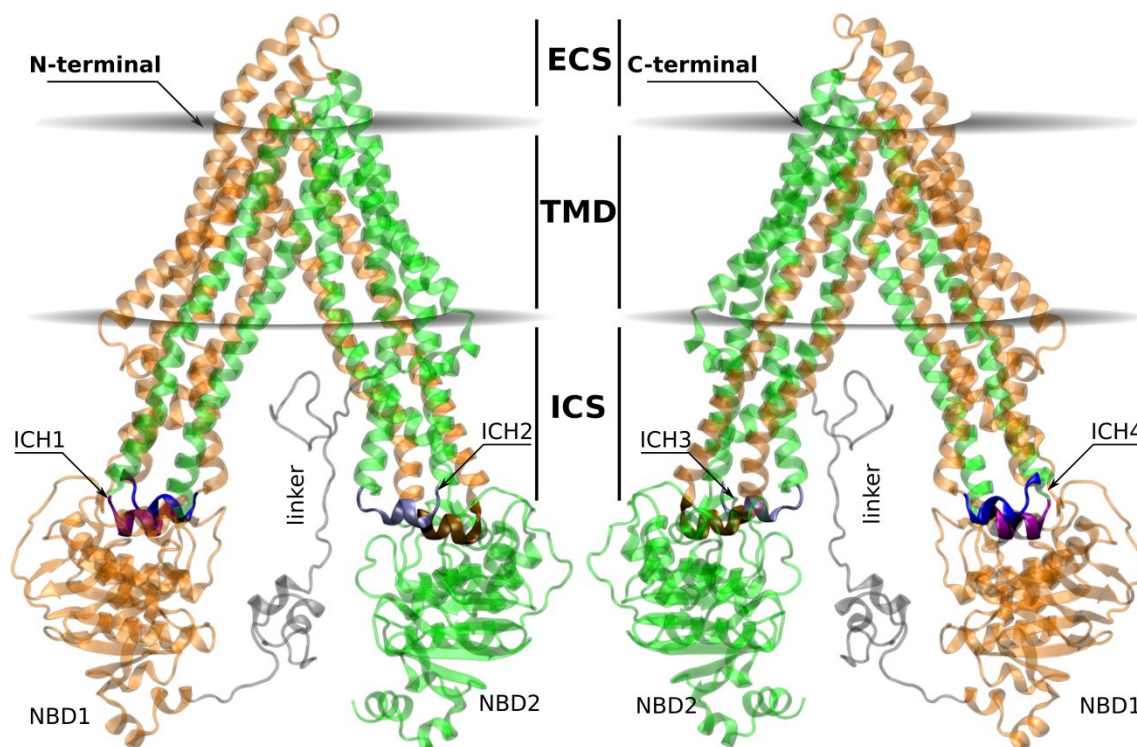
Figure 1.2. Structural representation of the human P-glycoprotein in an IF conformation. The 12 TMHs are distributed within two halves (N-terminal, orange; and C-terminal, green), being physically linked to the respective NBD by coils bridging TMH6/NBD1 and TMH12/NBD2, and also by non-covalent interactions involving short intracellular coupling helices: ICH1 (purple) /ICH4 (blue) with NBD1 and ICH2 (silver) /ICH3 (brown) with NBD2. The DBP is a large cavity between both TMDs. Both P-gp halves are connected by a small peptide sequence — the linker (gray). Each segment within both halves can be subdivided according to its location, namely extracellular (ECS), transmembrane (TMD) or intracellular (ICS).

Interestingly, the TMH4 kink described by Szewczyk and co-workers only seemed to occur upon binding of ligands acting as activators of ATPase activity (QZ-Ala and QZ-Val). Oppositely, ATPase inhibitors (QZ-Ile and QZ-Phe) induced a more straight helical conformation, similar to that observed in the *apo* structure (PDB ID: 4Q9H).[22] More recently, asymmetries in the ATP-binding between the NBDs were observed in four X-ray crystallographic structures of mouse P-gp in an ATP-bound IF conformation (PDB IDs: 5KO2, 5KOY, 5KPD, 5KPJ).[23] Although this configuration may not represent a real conformation during the catalytic cycle, it provided insights that ATP binds preferentially to NBD1 in the absence of substrates. It also suggested that differences in the ATP-binding affinity between the two NBDs may exist, in agreement with at least another experimental study from the same year.[24]

Lastly, the recently published X-ray crystallographic structures of mouse P-gp (PDB IDs: 6UJW, 6UJP, 6UJR, 6UJS, 6UJT, 6UJN),[25] in complex with the environmental pollutant BDE-100, further demonstrated that the content of aromatic residues within the DBP is important for drug recognition and poly-specificity. Aiming for a more thorough

identification of possible binding sites, Le et al. proposed three DBSs at the top of the DBP: i) one located at C-terminal P-gp halve and surrounded by multiple TMHs, in agreement to previous structural data,[26] ii) a second one located at the opposite N-terminal P-gp halve, and iii) a third DBS lying on the axis of pseudo-symmetry between both P-gp halves. According to the authors, the first site is the most hydrophobic one and containing 73% of aromatic residues, while the latter is the most polar, with only 56% of aromatic residues. Quite interestingly, although BDE-100 preferentially interact with the former, the ligand shifted to the other binding sites when some aromatic residues of the former were mutated to alanine, but without changes of the inhibitory effect on P-gp ATPase activity. These findings led the authors to conclude that the presence of aromatic residues at the DBP was, therefore, important for drug poly-specificity.

Despite the valuable information provided by these crystallographic structures, there are some issues concerning the crystallization process of membrane proteins that may lead to unreliable results. In particular, as the lipid membrane plays an important role on P-gp architecture and function,[27-31] the removal of the membrane during crystallization may also compromise the stability and function of the transporter. An example is the NBD-NBD distances reported in the literature for some murine P-gp structures in the IF conformation. Ranging from 30 Å (PDB ID: 3G5U)[16] and 31 Å (PDB ID: 4KSB)[32] up to 45.5 Å (PDB ID: 5KOY)[23] or 59.6 Å (PDB ID: 5KPJ),[23] all these structures reflect the high conformational flexibility of the P-gp transporter and highlight many conformation possibilities with similar energies upon lipid removal.[31] Additionally, and likewise relevant is the absence of highly flexible regions such as the linker (residues 627-688), unable to be identified in the electron densities of all the previously published P-gp crystallographic structures to date. As the linker motif is described to be crucial for P-gp stability, folding, substrate recognition and ATP hydrolysis,[33,34] its importance on P-gp structure cannot be neglected. Another issue concerns the fact that all P-gp crystallographic structures described above are incomplete, lacking the first N-terminal residues: residues 1-32 (PDB ID: 3G5U), residues 1-30 (PDB ID: 4Q9H) and residues 1-31 (PDB ID: 5KOY). Finally, the introduction of mutations to reduce P-gp flexibility or to trap the protein in a specific conformation during the crystallization process are other examples that may introduce artifacts in the structure and function of the transporter.[23,25,35]

1.2.2 Cryogenic electron microscopy (cryo-EM) structures.

In 2018, the first human P-gp in the ATP-bound outward-facing conformation (OF) (PDB ID: 6C0V)[36] was obtained using cryogenic electron microscopy (cryo-EM). While in the IF the NBDs are completely separated, the OF conformation showed a closed NBD

dimer trapping two ATP molecules. Additionally, the structural data also showed that the DBP formed by the TMDs had opened to the extracellular milieu. This was another important hallmark that provided additional insights in the P-gp efflux mechanism: i) the IF→OF transition implied large rearrangements in the TMHs, most particularly in the "crossing helices" 4/5 and 10/11, which were already known to play an important role in the formation of the NBD dimer upon ATP-binding,[37] ii) during the IF→OF transition, all TMD-NBD interfaces moved as one "concerted rigid body" that allowed the signal to propagate from the NBDs to TMDs, and iii) the TMHs rearrangement and subsequent substrate release occurred after ATP-binding and before ATP hydrolysis, as no substrate was found inside the DBP. In the same year, a mouse cryo-EM structure was obtained in the post-hydrolytic state (PDB ID: 6GDI).[38] Although the existence of populations in an OF was not clear, most of those obtained were indubitably found in the IF conformation, even in the presence of adenosine diphosphate (ADP). Herein, the structural data showed two densities corresponding to regions near the inner leaflet of the lipid bilayer (one in each P-gp halve) as possible DBSs, in agreement with previous studies.[39-41] Additionally, a 25-Å long curve density connecting the end of NBD1 with the cytoplasmic portion of TMH7 was identified, thus suggesting a possible location for the linker region, also in agreement with a previously published refined murine P-gp structure.[30]

Recently, several human and human-mouse chimeric P-gp cryo-EM structures in the IF conformation were published, providing even more insights on drug recognition and specificity. The first one is the human-mouse chimeric P-gp, in complex with the antigen-binding fragment of the human-specific inhibitory antibody (UIC2 Fab), with two molecules of the P-gp inhibitor (modulator) zosuquidar bound inside the DBP (PDB ID: 6FN1).[42] The structure showed conformational changes in the TMHs, most particularly a kinking and bending of the TMHs 4 and 10 upon zosuquidar binding. This feature was found to trap both molecules in a large central region within the DBP, described as "a fully occluded ligand-binding cavity". Later on, two other cryo-EM structures, also bound to the UIC2 Fab, were made available: i) a human-mouse chimeric P-gp (PDB ID: 6QEE)[43] with two zosuquidar molecules bound inside the DBP, and ii) a human P-gp in the presence of taxol, a known P-gp substrate (PDB ID: 6QEX).[43] Herein, the superimposition of both P-gp structures revealed that taxol and zosuquidar bind at the previously described "fully occluded ligand-binding cavity".[42] However, in the presence of taxol, the NBDs are closer than in the presence of zosuquidar, leading the authors to suggest that the differences between substrates and modulators could be associated to NBD movements and ATPase activity.[43] More recently, another set of human P-gp cryo-

EM stuctures, in complex with UIC2 Fab fragment of the inhibitory monoclonal antibody MRK16, were obtained in the presence of substrates and modulators (PDB IDs: 7A65, 7A69, 7A6C, 7A6E, 7A6F).[44] Interestingly, two DBSs within the DBP were characterized in the novel P-gp structures. The first one is the "occluded cavity" in a central region of the DBP, formed by the conformational changes in the TMHs 4 and 10, upon drug-binding, and in agreement with previous cryo-EM structures. [42,43] The other one, found at the C-terminal P-gp halve, was surrounded by multiple TM helices, in agreement with a previous study by Le and co-workers.[25] The authors described this second binding site as an "access tunnel", rich in phenylalanine residues (F239, F303, F343, F994) interacting with the modulators tariquidar and elacridar and spawning between the L-shaped and U-shaped regions where these molecules bind.[44]

Despite the importance of these recent P-gp structures for the efflux mechanism and drug specificity, the most striking difference from the murine X-ray structures is the "kink" of TM helices 4 and 10, when ligands are present, to form the "occluded" conformation.[42-44] Although a structural kink had already been reported for a single P-gp crystal structure, this specific feature was only observed in the TMH4 and in the presence of ligands increasing its ATPase activity (QZ-Ala and QZ-Val).[22] Even so, it is not totally clear what happened to the conformation of the TMHs 4 and 10 in a drug-free IF conformation. While a straight conformation of the TMHs 4 and 10 was still reported for the drug-free human–mouse chimeric ABCB1-UIC2-Fab (only depicted in the electron-density maps), more poorly ordered TM helices 4 and 10 were instead reported for the human ABCB1-UIC2-Fab (PDB ID: 6QEX) and the ABCB1- MRK16-Fab structures (PDB ID: 7A65). Intriguingly, the only other example where kinks were reported for TM4 (but not TM10) refers to a crystallographic structure of the *Cyanidioschyzon merolae* P-glycoprotein homolog, *Cm*ABCB1 (PDB ID: 3WMG),[35] in complex with the cyclic peptide inhibitor aCAP bound to the P-gp's extracellular domains.

On the other hand, recent studies using targeted Molecular Dynamics (MD) simulations, aimed to evaluate the P-gp conformational changes during the IF→OF transition, reported different results. According to Zhang et al.,[45] extensive TMHs reorientation upon ligand binding occurred, similar to those reported in the recent cryo-EM P-gp structures, but without observing any kink or unfolding events in TMHs 4 and 10. Following, and aiming to assess the effects of modulator and substrate binding on the P-gp conformation, Xing and co-workers[46] also uphold the importance of changes in the TMH rearrangement, but once again no kinked or distorted conformations were observed for TMH4 and 10.

Similar to the X-ray structures, any structural modifications inserted on these cryo-EM P-gp structures to promote protein stabilization and trapping in a specific conformation may induce severe changes in the P-gp structure and function, likely leading to erratic conclusions.[47] Such modifications include i) mutations at the NBD catalytic residues to trap P-gp in an OF,[36] ii) disulfide cross-linking of the NBDs to maintain the NBD closure in absence of nucleotide,[42] iii) ADP/vanadate inclusion to trap P-gp in a post-hydrolytic state[38] and/or iv) antibodies bound to the external side of the transporter.[43,44]

Very recently, a cryo-EM structure of a murine P-gp bound to ivacaftor (PDB ID: 7OTG)[47] was obtained and made available. Ivacaftor is a novel cystic fibrosis therapeutic, and is also a transported substrate by P-gp.[47] The P-gp structure obtained without any structural modifications and/or antibodies, again revealing straight helical conformational of the TMHs 4 and 10, strikingly similar to that observed in most of the X-ray structures of murine P-gp. These results reinforce the idea that the experimental conditions must be taken into account to avoid misinterpretations.

1.3 P-glycoprotein catalytic cycle

From the structural data it is possible to identify the IF and OF as the two main distinct P-gp conformations during the catalytic cycle. The first one is thought to represent the initial step of the efflux cycle,[16] wherein the DBP is fully opened to the cytoplasm and the periplasmic pore is closed. In this conformation, drugs bind to high-affinity DBSs, and both NBDs are dissociated. Upon ATP/Mg^{2+} binding, extensive conformational changes occur, driven by NBD dimerization, and the transporter is prompted for ATP hydrolysis.[48] The NBD dimerization involves the formation of two nucleotide-binding sites (NBSs), each one formed by the Walker A and Walker B motifs of one NBD and the ABC-signature (LSGGQ) motif of the other (**Figure 1.3**).[36,49]

Cross-linking and mutagenesis studies showed that the catalytic residues E556 (NBD1) and E1201 (NBD2), together with the Q-loop residues Q475 (NBD1) and Q1118 (NBD2), located at the NBD-NBD interface, are part of a "spring-like mechanism" that is required for the NBD movement and dimer formation.[51,52] Following, large conformational changes in the TMHs spatial position takes place, reorganizing the TM helical bundle towards the opening of the periplasmic pore, and thus reaching the OF state.[48] In this conformation, the DBSs show low-affinity for substrates, which leads to drug release. Finally, after ATP hydrolysis P-gp returns to the initial conformation, resetting itself and allowing a new catalytic cycle to initiate.[53]
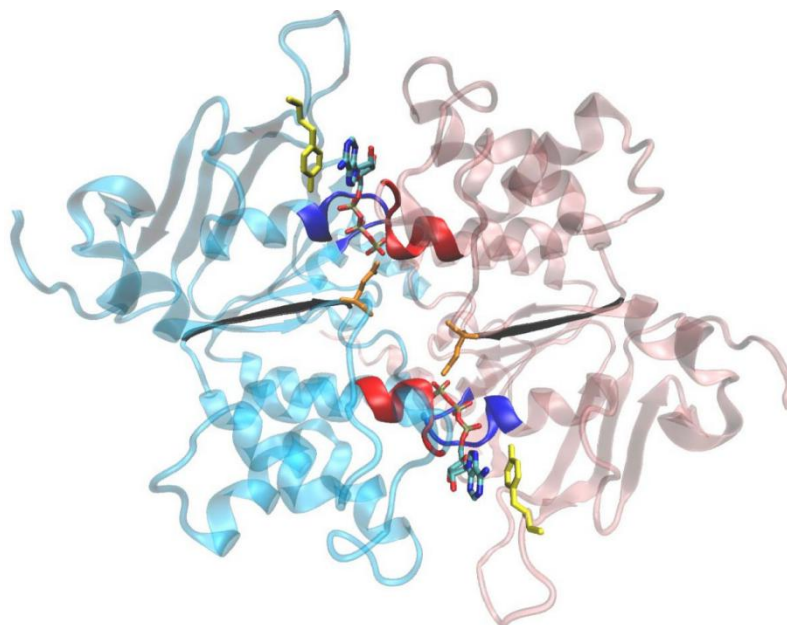
Figure 1.3. Ribbon representation of the NBD closed dimer (top view from the membrane) in the presence of two ATP molecules (depicted as sticks). The N-terminal NBD (light blue) and the C-terminal NBD (light pink) are associated in a head-to-tail fashion. Each ATP-binding pocket is formed by the A-loop (yellow), the Walker A motif (dark blue), the Walker B motif (black) and the glutamine residue of the Q-loop (depicted in orange) from one NBD and closed by the Signature motif colored in red from the other NBD (image from Becker, 2009).[50]

Oppositely, other studies proposed an alternate cycle for ATP hydrolysis,[54,55] where only one ATP molecule is tightly bound to the NBS ("occluded") and hydrolyzed at a time (although a second ATP molecule is necessary to be bound to the opposite NBD). According to these studies, the asymmetry between the NBSs may represent an intermediate conformation after ATP-binding but prior to ATP hydrolysis. More recently, Verhalen and co-workers[24] suggested a model of the P-gp conformation cycle, corroborating the previously proposed alternate cycle concerning ATP hydrolysis.[54,55] According to the authors, the proposed model comprises four main stages, namely i) binding of two ATP molecules at each NBD, thus promoting NBD dimerization, followed by the substrate binding to DBP, ii) induction of a high-affinity "occluded state" in one ATP molecule, while the second one remains in a low-affinity "loose state", iii) hydrolysis of the first ATP molecule and allowing ADP and Pi to dissociate by shifting the "occluded state" to the ATP molecule located at the opposite NBD, and iv) hydrolysis of the second ATP molecule, inducing conformational changes and driving P-gp from IF→OF to allow substrate efflux. This model is also in perfect agreement with the constant contact model earlier proposed by Jones and George.[56,57]

Nonetheless, while several studies shed more light on how the P-gp catalytic cycle occurs, others provided somewhat contradictory information. A recent study using luminescence resonance energy transfer (LRET), showed that the conformational

changes in the NBDs occur immediately after substrate binding,[58] in contrast with earlier findings.[24] Additionally, while the human P-gp cryo-EM structure in the ATP-bound OF suggested that drug efflux likely occur before ATP hydrolysis,[36] Dastvan et al.,[59] demonstrated that substrate efflux must occur simultaneously with, or immediately after, ATP hydrolysis, thus corroborating the alternate cycle for ATP hydrolysis described by Verhalen and co-workers.[24]

1.4 Insights on P-gp drug recognition and specificity

Earlier experimental studies,[39,40,60,61] described the existence of two DBSs, deeply buried in the inner leaflet of the lipid bilayer, responsible for binding and transport of substrates. These DBSs were classified as the H-site (N-terminal P-gp halve) and R-site (C-terminal P-gp halve), due to their ability to specifically transport Hoechst-33342 (H33342) and Rhodamine-123 (R123), respectively. According to the authors, both sites interact in a positively cooperative manner, with drug-binding at the H-site stimulating the efflux of substrates by the R-site and vice-versa. The evidence of a third DBS was postulated from additional studies with prazosin and progesterone, as the binding of these ligands also stimulated the transport of H33342 and R123.[62] According to Shapiro and co-workers, prazosin and progesterone must bind to a different DBS, likely located in-between the P-gp halves, in order to have a positive allosteric effect on the transport of substrates by both sites. However, the existence of a third DBS located in an upper position of the DBP was only confirmed many years later, with the publication of the first murine P-gp crystallographic structure.[16] In 2013, by means of molecular docking, Ferreira et al.,[41] defined the location and lining residues for the three known DBSs within the DBP. Accordingly, the modulator site (M-site), which preferentially bind P-gp modulators, the most hydrophobic, is located at the top of the DBP. Oppositely, the other binding sites, H and R, have an increased percentage of polar residues and mostly bind substrates (SBSs). Later, a mouse cryo-EM structure showed two densities nearby the inner leaflet of the lipid bilayer, in the same regions of the predicted SBSs, inferring that the densities found may correspond to H (N-terminal P-gp halve) and R (C-terminal P-gp halve) sites.[38] More recently, Barbieri and co-workers[47] described the binding site for ivacaftor in P-gp, a potentiator of the Cystic Fibrosis Transmembrane Conductance regulator (CFTR/ABCC7) activity,[63] also transported by the P-gp efflux pump. The cryo-EM structure of the IF murine P-gp (PDB ID: 7OTG) showed an ivacaftor molecule in a "channel gate" at the N-terminal P-gp halve, interacting with residues from TMHs 6 and 10, and two residues from TMH12 in the opposite P-gp halve, in a location compatible with the H-site. On the other hand, the existence of multiple flexible DBSs within the DBP had already been suggested. Earlier studies by Loo and Clarke, using cysteine-scanning

mutagenesis and thiol-reactive P-gp substrate analogs, showed that the DBP is lined by residues from different TMHs.[19-21,64,65] The authors also hypothesized that drug-binding should occur by an induced-fit mechanism, where the size and shape of the substrate induce rearrangements in the TMHs and, thus several flexible DBSs may coexist within a common DBP.[66] Indeed, some recent structural data points to the existence of flexible DBSs, mainly located in a central region and at the C-terminal P-gp halve.[42,43,44] Accordingly, these DBSs features seem to be dependent on the ligand, and may bind either substrates or modulators. Another evidence that drugs may bind in flexible DBSs with different orientations was provided by Loo and Clarke[67] by mapping a potential tariquidar binding site. According to the authors tariquidar may bind with high affinity in a central region within the DBP between both P-gp halves, thus trapping P-gp in a conformation that allows stimulation of ATPase activity but without drug transport. Moreover, a previous mutagenesis study additionally identified three structural motifs, in a similar region at the C-terminal P-gp halve, that could be related to drug-binding.[68] According to the latter, each structural motif is formed by several amino acid pairs, being most of them aromatic residues: i) Y953 (TMH11)/F978 (TMH12), ii) Y310 (TMH5)/F728 (TMH7), and iii) Y307 (TMH5)/Q725 (TMH7). Additionally, the authors also showed that while the structural motif Y953/F978 is involved in binding of zosuquidar, the binding of other P-gp modulators may require a combination of several motifs: i) Y953/F978-Y310/F728 (elacridar), and ii) Y953/F978-Y310/F728-Y307-Q725 (tariquidar). Yet, structural data showed that residue pairs Y307/Q725 and Y310/F278 are also involved in the binding of the substrate taxol.[43] Furthermore, evidence that P-gp modulators and substrates could bind at the same binding site(s) was also reported by Subramanian and co-workers.[69] Through the analysis of a potential of mean force (PMF) profiles, these authors showed that H33342/R123 have distinct binding locations within the DBP and acted as non-competitive substrates. Still, H33342 and tariquidar seemed to compete to the same DBS. Lastly, the important role of the aromatic residues in substrate recognition and poly-specificity was also reported.[25] Several X-ray structures of murine P-gp identified three DBSs at the top of the DBP, able to bind the pollutant BDE-100. The authors showed that the ligand seemed to preferentially interact with the most hydrophobic DBS, rich in aromatic residues, changing its binding location according to the percentage of aromatic residues of each DBS.

The table below (**Table 1.1**) lists the residues within the DBP involved in drug-binding found in the available literature.[16,19-22,25,42-44,47,64,67,70-72]

Table 1.1. Comparison between residues identified in biochemical assays and derived from crystallographic and/or cryo-EM structures.

TMH	Resnr	Biochem assays	Aller et al., 2009	Szewczyk et al., 2015	Le et al., 2020	Alam et al., 2018	Alam et al., 2019	Nosol et al., 2020	Barbieri et al., 2021
1	H61	tariquidar verapamil							
	G64	tariquidar verapamil							
	L65	tariquidar verapamil				zosuquidar	taxol zosuquidar	elacridar tariquidar	
	M68						taxol	vincristine	
	M69	tariquidar		QZ-Val	BDE-100		taxol		
	F72	tariquidar		QZ-Val	BDE-100			tariquidar	
2	A129	tariquidar							
4	S222	verapamil vinblastine cyclosporin A colchicine							
	L225			QZ-Ala QZ-Val					
	A229			QZ-Ala QZ-Val					
	W232			QZ-Ala QZ-Val		zosuquidar	taxol	elacridar tariquidar	
	A233					zosuquidar			
	F239							elacridar	
5	I299	vinblastine							
	F303	tariquidar				zosuquidar	taxol zosuquidar	zosuquidar	
	I306	verapamil vinblastine cyclosporin A tariquidar				zosuquidar	taxol	vincristine elacridar tariquidar	
	Y307	tariquidar					taxol	tariquidar	
	S309	tariquidar							
	Y310	elacridar tariquidar					taxol	vincristine elacridar tariquidar	
	F314				BDE-100				
6	F335				BDE-100				
	F336	tariquidar	QZ59-SSS QZ59-RRR	QZ-Ala QZ-Val QZ-Leu QZ-Phe	BDE-100	zosuquidar	taxol	elacridar tariquidar zosuquidar	

Table 1.1. (continued)

6	L339	vinblastine cyclosporin A verapamil	QZ59-RRR	QZ-Leu		zosuquidar	taxol zosuquidar	zosuquidar	
	I340	rhodamine B	QZ59-RRR QZ59-SSS			zosuquidar			
	A342	verapamil vinblastine cyclosporin A	QZ59-RRR QZ59-SSS						
	F343	colchicine cyclosporin A tariquidar verapamil vinblastine		QZ-Ala QZ-Val QZ-Phe		zosuquidar	taxol zosuquidar	vincristine elacridar	ivacaftor
	S344						taxol		
	G346	colchicine verapamil vinblastine							
	Q347						taxol zosuquidar	vincristine elacridar	ivacaftor
	P350	verapamil cyclosporin A vinblastine colchicine							
7	Q725	tariquidar	QZ59-RRR QZ59-SSS	QZ-Phe	BDE-100	zosuquidar zosuquidar	taxol zosuquidar	elacridar tariquidar zosuquidar	
	F728	elacridar tariquidar verapamil	QZ59-RRR QZ59-SSS		BDE-100		taxol	elacridar tariquidar	
	A729			QZ-Leu QZ-Phe	BDE-100				
	F732	tariquidar		QZ-Ala	BDE-100				
	I735				BDE-100				
8	F759				BDE-100				
	L762		QZ59-SSS						
	F770					zosuquidar			
9	T837		QZ59-SSS						
	A841	rhodamine B							
10	V865	tariquidar							
	I868	colchicine cyclosporin A tariquidar verapamil vinblastine					taxol		ivacaftor

Table 1.1. (continued)

10	G872	cyclosporin A tariquidar verapamil vinblastine							ivacaftor
	E875					zosuquidar	taxol	vincristine tariquidar zosuquidar	ivacaftor
	M876								ivacaftor
	L879					zosuquidar			ivacaftor
11	F942	colchicine cyclosporin A tariquidar verapamil vinblastine							
	T945	colchicine cyclosporin A tariquidar verapamil vinblastine							
	Q946	tariquidar				zosuquidar	zosuquidar	vincristine	
	M949	tariquidar		QZ-Phe		zosuquidar	taxol zosuquidar	vincristine elacridar tariquidar	
	Y950	tariquidar				zosuquidar	zosuquidar	elacridar	
	F951					zosuquidar			
	S952	tariquidar							
	Y953	zosuquidar elacridar tariquidar	QZ59-RRR QZ59-SSS	QZ-Ala	BDE-100	zosuquidar	taxol zosuquidar	vincristine elacridar zosuquidar	
	A954	colchicine vinblastine							
12	L975	Rhodamine B tariquidar verapamil			BDE-100				
	F978	elacridar tariquidar zosuquidar	QZ59-RRR QZ59-SSS		BDE-100			tariquidar	
	S979		QZ59-RRR					elacridar tariquidar	
	V981	daunomycin R123 Rhodamine B							
	V982	cyclosporin A Rhodamine B tariquidar verapamil	QZ59-RRR QZ59-SSS	QZ-Val QZ-Leu	BDE-100				

Table 1.1. (continued)

12	F983	daunomycin R123		QZ-Ala QZ-Val QZ-Leu QZ-Phe	BDE-100	zosuquidar	taxol zosuquidar	vincristine tariquidar zosuquidar	
	G984	cyclosporin A verapamil							
	A985	cyclosporin A verapamil	QZ59-SSS						ivacaftor
	M986	colchicine cyclosporin A verapamil vinblastine	QZ59-RRR QZ59-SSS	QZ-Val QZ-Leu QZ-Phe		zosuquidar	zosuquidar	elacridar tariquidar zosuquidar	
	A987					zosuquidar	zosuquidar		
	G989	colchicine verapamil vinblastine							
	Q990					zosuquidar	taxol zosuquidar	vincristine tariquidar zosuquidar	
	V991					zosuquidar	zosuquidar	zosuquidar	
	S993	colchicine cyclosporin A verapamil vinblastine							
	F994							elacridar zosuquidar	

From the ensemble of experimental data, it is noticeable the existence of a common DBS, able to bind both substrates and modulators (**Figure 1.4**). Most of the drugs described interact with residues of the TMHs 6, 10, 11, and 12. Concerning the location of other DBSs, such as the H and R sites, no additional information could be retrieved from any of the crystallographic structures. Therefore, the precise location of these SBSs, or even their existence, is still to be unequivocally proven. However, while biochemical and structural data agree regarding the binding location of the P-gp modulators verapamil, QZ isomers, tariquidar and elacridar (U-shaped forms), only those in which more than one molecule was found within the DBP provided a glimpse on the hypothetical locations of the postulated substrate-binding sites.

Nevertheless, although decades of research to unveil the mechanisms of drug recognition and poly-specificity, the number of DBSs and their precise location into DBP are still not completely clear. Moreover, how P-gp distinguishes substrates from modulators is another unsolved question. Structural data[43] suggests that the differences in the NBDs distance induced by substrates and modulators may be one of the mechanisms by which substrates stimulate the ATPase activity, while modulators induce a reduction of the ATPase activity. A later mechanistic study by Xing et al.,[46] also

provided evidence that the different NBD conformations between the binding of substrates and modulators may be associated with their effects on the P-gp ATPase activity. While in the presence of substrates the NBDs are more aligned and closer, and thus prompted to NBD dimer formation, the binding of P-gp modulators resulting in more separated and misalignment NBDs. In the latter, larger NBD-NBD distances and/or misalignments are thought to impair dimer formation and, consequently, to reduce ATPase activity. However, these mechanisms need to be deeply explored to develop new and high specific P-g modulators.
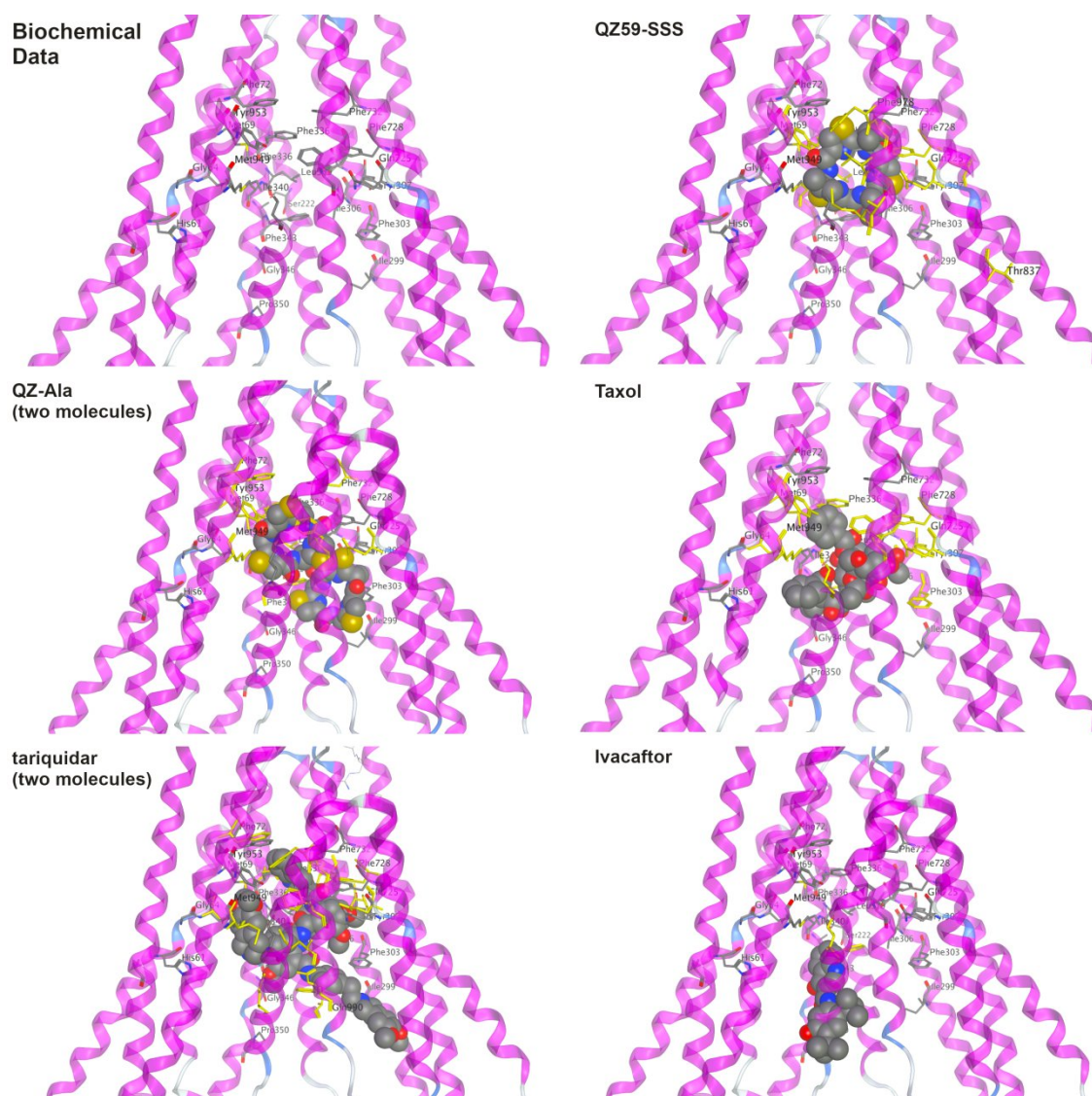
Figure 1.4. Comparison between residues identified by biochemical experiments and those derived from structural data. Ligands are represented in space-fill while matching residues (biochemical and structural studies) are depicted in yellow.

1.5 Insights on P-gp signal-transmission mechanism

To date, three distinct generations of P-gp modulators were categorized, but no clinical candidate has ever reached the market. The *first-generation* of modulators comprised

molecules approved for other therapeutic usages that were also able to modulate the efflux of cytotoxic drugs, e.g. verapamil,[73] quinidine[74] and cyclosporin A.[75] Their low affinities toward P-gp, however, required high *in vivo* dosages for an effective MDR reversion, which impaired any clinical usage of these compounds. Thus, the development of a *second-generation* of modulators, e.g. dexverapamil,[76] dexniguldipine,[77] biricodar[78] or valspodar (PSC 833),[79] followed. Although more specific and deprived of pharmacological activity, other toxic effects and major alterations in the pharmacokinetics of anticancer agents were observed in the clinical trials that did not allow further developments. After the *second-generation* modulators failure, a *third-generation* of efflux modulators e.g. tariquidar (XR9576),[80] elacridar (GF120918),[81] zosuquidar (LY335979),[82] laniquidar (R101933),[83] with increase potency and specificity toward the transporter, was defined from compounds obtained through high-throughput screening approaches. Nonetheless, and although more effective in modulating drug efflux, phase III clinical trials showed a dramatic increase in cellular toxicity (tariquidar) or reduced in *vivo* effectiveness (zosuquidar, laniquidar). In the following years, many potent efflux modulators were depicted in the literature,[84,85] but even those that inhibited P-gp at the nanomolar range, like tariquidar, failed to demonstrate its efficacy in the clinical environment.[86]

The poly-specificity of the DBP, able to bind many structurally unrelated compounds, is pointed out as the biggest drawback for the development of more selective and effective P-gp modulators.[16,68,87,88] Although most of them act as competitive inhibitors,[89] other compounds as tariquidar,[90] verapamil,[91] methadone[92] or HG-829[93] are also identified as non-competitive inhibitors, thus suggesting the existence of allosteric binding site(s) within the P-gp structure. Thus, it is of the utmost importance to identify the precise mechanism of action for additional allosteric modulatory sites within the P-gp structure.

Earlier studies proposed the existence of allosteric binding for flavonoids at a vicinal region next to the ATP-binding sites, and at the C-terminal of NBD2 of P-glycoprotein.[94] Later on, further studies identified a region outside the DBP, located at the lipid-protein interface, as a putative allosteric binding site for the modulators cis-(*Z*)-flupentixol and verapamil.[84,85] Other studies, however, proposed that it may be possible to overcome MDR by specifically targeting the ICHs, a specific motif within the ABC transporter structure located at the TMD-NBD interfaces (**Figure 1.2**).[37] Several structural, biochemical and computational studies have been providing insights about the role of the ICHs in the TMD-NBD communication. The crystallographic structure of Sav1866 transporter in a OF conformation showed short intracellular "coupling helices" located between specific TM helices, that connected to the NBDs through non-bonded

interactions. According to the authors, these ICHs are likely involved in the transmission of the conformational changes from the NBDs to TMDs upon ATP-binding and hydrolysis.[15] Following, the *Caenorhabditis elegans* (*C. elegans*)[95] homologue P-gp crystal structure revealed that the ICH2-NBD2 and ICH4-NBD1 interfaces act as "ball-and-socket" joints, wherein the ICH (the ball) engage into a cleft on the respective NBD surface (the socket), similar to that observed in the ABC importers.[96] The "ball-and-socket" joints revealed to be an important feature of the ABC transporters, acting as signal-transmission interfaces and enabling the conformational changes needed to shift P-gp from the IF to OF conformations.[97] Nonetheless, in the ABC exporters including P-gp, ICH1 and ICH3, also establish interactions with the respective NBD, thus increasing the TMD-NBD communication surface.[95]

In recent years, many mechanistic studies have provided valuable data about the biophysical process by which the ICHs are tightly coupled to the conformational shifts from the IF→OF transition. One of the first studies, using targeted MD[98] on P-gp revealed that, upon NBD engagement, the NBDs motions are transmitted to the TMDs through the ICHs. According to the author, the signal from the NBDs induces a great straightening of TMH4/5 (ICH2) and TMH10/11 (ICH4), thus contributing to the opening of the periplasmic pore, as seen in the fully OF structure. Moreover, this study also suggested that ICH2/ICH4 are more crucial than ICH1/ICH3 in driving conformational changes to the TMHs, because the number of contacts with the corresponding NBD were higher during all simulation time. Another study[99] identified some specific residue pairs at the ICH-NBD interfaces likely involved in the propagation of the conformational changes to the TMHs, namely E159/R467 (ICH1/NBD1), R262/E1084 (ICH2/NBD2), and D800/R1100 (ICH3/NBD2). Later, Prajapati and co-workers[100] also showed that ICHs move together with the NBDs, in agreement with Kim et al.,[36] allowing the transmission of the rotational and translational motions from the NBDs to the TM helices. Additionally, the same study also claimed that the predominant pushing force was exerted on the "crossing helices" TMHs 4/5 and TMH10/11 via ICH2 and ICH4, respectively, thus corroborating Wise's findings.[100]

Together with structural and mechanistic studies, biochemical approaches have also shown a remarkable contribution to clarify the role of ICHs on the P-gp function. Since the 1990s, several studies have demonstrated that the ICH-NBD interfaces have important roles for the P-gp activity as well as for the folding and maturation of the transporter.[101] One of the first studies identified four residues in the ICH1 that were critical for maintaining the P-gp function. As proof of concept, Currier and co-workers[102] replaced residues in the ICH1, TMH3, and TMH4 with the corresponding ones from the related

MDR2 transporter to find that, although the transporter was expressed, it was unable to confer drug resistance. However, by reverting the ICH1 mutations at positions 165, 166, 168, and 169 to the native MDR1 residues, the protein restored its MDR activity, being able to re-efflux drugs. Point mutations were also introduced in ICH1, TMH2, TMH3, the entire extracellular loop connecting TMH2/TMH3 (ECL2) and TMH4 by mean of a random mutagenesis approach, by Kwan and Gros.[103] Interestingly, the mostly sensitive regions to mutations were the ICH1 and ECL2, inducing a partial or complete loss of P-gp activity. In addition, the authors identified four key residues located at the ICH1 in the mouse MDR3 sequence (T169, R170, L171, and T172) as particularly sensitive to alterations and critical to the efflux function. Moreover, since only drug-stimulated ATPase activities were affected, the authors suggested that mutations in these particular residues could impair signal-transmission, from the transmembrane domains to the ATP-binding sites, or vice-versa. However, it is worth noticing that mutations in other ICH1 residues also caused the partial or complete loss of function, thus corroborating the structural/functional role of ICH1 in the efflux mechanism.

Other studies by Loo and Clarke[104] revealed that salt bridges, established between charged residues located at the ICH2-NBD2 interface, also play an important role in the maturation and folding of the P-gp efflux pump. Since P-gp is firstly synthesized as a 150 kDa core-glycosylated protein, any folding defect due to the introduced mutations will retain the 150 kDa immature form in the endoplasmic reticulum, being rapidly degraded.[105] In contrast, the fully folded transporter is further processed in the Golgi apparatus (modification of the carbohydrate groups) to yield the 170 kDa mature protein that is trafficked to the cellular surface.[106] Therefore, and following their initial studies on ICH2 mutations, Loo and Clarke subsequently reported that most of these mutations partially or completely inhibited P-gp maturation, describing the A266/F1086 residue pair as an important contact point for P-gp folding, because many mutations in these residues yielded the 150 kDa immature form of P-gp. In addition, this study also showed that the same residue pair A266/F1086 is critical for P-gp activity since cysteine mutants were inactive.[97] However, in contrast to previous studies, no mutation in ICH1 or ICH4 was found to be important for protein folding.[107] Another study also noted the importance of the ICH-NBD interfaces for P-gp function by showing that mutations in the tryptophan residues W158F/W799F, located at the ICHs 1 and 3, respectively, induced drug resistance to the fungicide drugs tested.[108] Additionally, other authors also observed severe effects on mating efficiencies in yeast colonies by affecting the transport of a-factor pheromone, a peptide required for mating, in a P-gp homologue from *Candida albicans*.[109] Two other mutations in ICH1 (D164C) and ICH3 (D805C) were also

described to have significant alterations in P-gp function.[110] While D805C completely abolished drug transport, the double mutant D164C/D805C showed about 70% decrease in daunorubicin and [N-ε-(4-nitrobenzofurazan-7-yl)-dLys8]-cyclosporin A (NBD-CsA) transport, only partially effluxing calcein-AM (~60%). Although not being critical for ATP-binding, the observed reduction in the ATPase hydrolysis rate (V_{max}) by these mutations led the authors to conclude that such residues must play a role in the TMD-NBD communication. Additionally, it was also found that while mutations in the NBD2 residue Y1087 (Y1087L and/or Y1087A) severely reduced P-gp activity (~90%), a more conservative mutation (Y1087F) impaired about 70% of the activity.[111] Since Y1087 interacts with D805 through hydrogen bonds,[112] other studies identified the residue pair D805/Y1087 as a key contact between ICH3 and NBD2 for coupling conformational changes between both domains.[113,114] In addition, it was also demonstrated that ATP-binding mediates conformational changes at the ICH2/ICH3-NBD2 interfaces because, in the mutant protein A259C (ICH2)/W803C (ICH3), cross-linking was impaired in the presence of ATP or AMP-PNP.[111]

Altogether, the ICHs located at the TMD-NBD transmission interfaces seem to have a critical function in the maturation, folding, and activity of the transporter, although with different contributions.[112] A more recent study[115] showed that, while the ICH2-NBD2 transmission interface is highly sensitive to point mutations, the maturation of P-gp was insensitive to all tested mutations at the ICH4-NBD1 interface. Specifically, by mutating residues in NBD1 (L443 and Y444) and NBD2 (F1086 and Y1087), it was concluded that the ATPase activity was severely impaired (Y444X, ~30%) or even fully inhibited (Y1087X). Thus, the authors suggested that the second transmission interface (ICH2-NBD2) could, in principle, be targeted with compounds suitable to inhibit the activity, thus improving the effectiveness of chemotherapy.

1.5.1 Targeting signal transduction by small molecules

Taking all the information collected above, targeting the ICHs located at the TMD-NBD interfaces are one of the best working hypotheses to develop new molecules that are able to modulate drug efflux by P-gp. By impairing such motifs from mediating ATP dependent conformational changes it would be possible to i) avoid competition with drug- and/or ATP-binding sites, thus vastly reducing the adverse effects identified previously during the clinical trials of past MDR modulator generations; ii) identify and develop new scaffolds for high-affinity modulators by *de novo* design, nature-inspired molecules, drug repurposing or fragment-based approaches; and iii) selectively block ABC transporters

expression, specifically in cancer cells, by incorporating such modulators in nanocarriers designed to specifically target cancer cells.[116]

Recently, it has been proposed the existence of allosteric binding sites, in close vicinity to the ICHs, that upon ligand binding may impair the TMD-NBD signal-transmission responsible for driving conformational changes leading to efflux.[117] In 2001, Isenberg et al.,[118] identified an [^{125}I]-iodoarylazidoprazosin binding site next to the second cytoplasmic loop, which includes ICH2 between TM helices 4 and 5. This site, also reported as a common binding site for [^{125}I]-iodomycin,[119] [^{125}I]-iodipine[120] and dexniguldipine,[121] is proof of concept that molecules are able to bind P-gp in regions distinct from those described for the DBP. However, the first study that identified a possible drug-binding site next to the transmission interfaces was conducted by Kim et al.,[122] in *the Arabidopsis thaliana* P-gp homologue. After identifying the 2-[4-(diethylamino)-2-hydroxybenzoyl]benzoic acid (BUM) and 1-N-naphthylphthalamic acid (NPA) as novel ABCB1-dependent auxin transport inhibitors, molecular docking predicted that NPA could bind to a pocket located between the ICHs and Q-loops of both NBDs, while BUM was predicted to bind only between the ICHs and Q-loop of NBD2, and in an additional location between NBD1 and NBD2. Remarkably, while the E502K mutation in the *Arabidopsis thaliana* sequence found in the N-terminal NBD completely abolished NPA binding, the mutation of phenylalanine residue at the C-terminal NBD (F792K) did not affect NPA binding but significantly reduced BUM competition. These findings clearly demonstrate that E502 and F792 (corresponding to E526 and F804 in the human P-gp sequence) are located in domains where these ligands bind to P-gp. Interestingly, Zolnerciks et al.,[123] additionally concluded that Q-loop is a central point where events in the DBP are coupled to the ATP catalytic cycle. Therefore, the binding of NPA and BUM to these regions is expected to alter the coupling between drug-binding at the TMDs and the ATP-binding/hydrolysis at the NBDs, as seen by Kim and co-workers[122] for the efflux inhibition of the hormone-like signaling molecule indolyl-3-acetic acid. Later, from the *in silico* screening of a large set of "drug-like" compounds, Brewer et al.,[124] predicted that some molecules may act as allosteric P-gp modulators interacting with residues located at the TMD-NBD interfaces. Furthermore, natural compounds such as flavonoids have also been described as promising modulators of MDR, binding in potential allosteric binding sites at the TMD-NBD interfaces of P-gp.[94,125] Later, molecular docking studies showed that, although narigenin derivatives preferentially bind at the DBP, they may also interact at the TMD-NBD interfaces, thus suggesting that the flavanone scaffold may be a suitable building block for developing novel allosteric P-gp modulators.[126] Similarly, thioxanthones — a group of natural analog compounds — have

been described as acting by a noncompetitive mechanism binding to regions located at the NBDs.[127] Finally, 1,4-dihydropyridine derivatives[126,128,129] and curcumin[130] were also predicted to interact with specific residues at the TMD-NBD interfaces of both NBDs.

1.6 Goals and Working Plan

Despite all the advances in cancer treatment, an increased number of chemotherapy failures are being registered, most of which by resistance to anticancer pharmacological action. Over-expression of ABC transporters represents one of the most relevant MDR mechanisms in cancer cells, and specifically concerning P-gp, three generations of efflux modulators were developed so far as an attempt to modulate its activity, but due to several problems in the clinical trials none reached the market. Therefore, a deep understanding of the molecular mechanisms of drug efflux, and a novel strategy to modulate drug efflux by P-gp are extremely important to develop high-affinity modulators with improved pharmacokinetic properties.

The poly-specificity of the DBP is considered by many the biggest drawback for the development of more selective and effective P-gp modulators, most of which are known to act as fast-turnover substrates or competitive modulators. However, after decades of research on P-gp the mechanisms of drug specificity, substrate recognition and/or signal-transmission between the transmembrane and cytoplasmic domains are still not completely understood.

Structural data and mechanistic studies on P-gp and CFTR identified a previously uncharacterized drug-binding site involving the ICHs, to where is thought to be possible to develop a new series of allosteric modulators.[117,131] In this work, a combination of computational and experimental approaches were undertaken to achieve the following goals:

I) *Acquisition of new insights on drug-binding and TMD-NBD signal-transmission mechanism*, by studying the wild-type (WT) P-gp and common variants — G185V,[132] G830V,[105] F978A[133] and ΔF335,[134] reported in the literature as having alterations in drug efflux. Herein, by means of MD simulations, the structural impact of such mutations on P-gp architecture was thoroughly characterized, in the presence and absence of molecules within the DBP;

II) *Characterization of the hypothetical allosteric DBS at the TMD-NBD interfaces,* by means of a computational Fragment-Based Drug Discovery (cFBDD) approach applying MD simulations. Fragments were derived from experimentally known ICH-NBD binders reported in literature and also from *in silico* data;[117]

III) *Identification of potential scaffolds*, able to bind in the previously identified DBS, and that may act as non-competitive P-gp MDR reversal agents. Based on the cFBDD findings, molecular docking studies were applied to a small *in-house* library of thioxanthone and flavanone derivatives. From the best top-ranked molecules, *in vitro* assays were undertaken to evaluate their effect on the P-gp ATPase activity. The ATPase assays were performed by the team of Professor Doctor Jean-Pierre Gillet, University of Namur, Belgium.

CHAPTER 2 - METHODOLOGY

2.1 Computational studies

2.1.1 Development of the human P-gp (hP-gp) homology model

Considering that no crystallographic or cryo-EM structures of hP-gp were available at the beginning of this thesis, a hP-gp homology model was developed using the murine P-gp (mP-gp) crystallographic structure (PDB ID: 4Q9H).[22] This template was chosen due to: i) the high sequence identity and similarity (87% and 94%, respectively) between the human and murine P-gp efflux pumps,[68] ii) its highest resolution, available at that time (3.40 Å), and iii) improvements in the resolution of several transmembrane helices (TMHs), first intracelular coupling helix (ICH1) and several extracellular loops, in respect to previous structures. The FASTA sequence of hP-gp was obtained from the Universal Protein Resource (UNIPROT) under the code P08183 (www.uniprot.org). The mP-gp crystallographic structure (PDB ID: 4Q9H), used as template, was retrieved from the Protein Data Bank (PDB, www.rcsb.org).[135,136] Both mP-gp crystallographic structure and the hP-gp FASTA sequence were loaded into the Molecular Operating Environment (MOE)[137] software and aligned by their sequence. The secondary structure of the linker motif, missing in all P-gp crystallographic structures so far, was retrieved from a previously published equilibrated mP-gp,[30] and used as template for modeling the respective sequence (A627-A688). Herein, the insertion of the linker in the novel structure was achieved by aligning the hP-gp homology model with the crystallographic template, followed by an "override" of the gap in the considered structure in which the refined mP-gp linker was used as template to obtain the human homologue. Afterward, 25 main-chain models sampling 25 side-chain orientations were performed, producing 625 models for the hP-gp structure in MOE (force field parameters used by default). The model with the lowest total potential energy was selected, protonated using the *Protonate 3D*[138] module available in MOE and exported as PDB file to be used GROMACS v5.0.7[139,140,141] package. This way, a hP-gp homology model in an inward-facing (IF) conformation (wild-type, WT) was obtained by incorporating the linker motif retrieved from a previously refined mP-gp structure.

2.1.2 Molecular Dynamics (MD) studies

2.1.2.1 Refinement of the hP-gp WT model: MD systems setup

To obtain a stable and reliable hP-gp model suitable to be used in further studies, the initial hP-gp homology model was inserted in an equilibrated lipid bilayer,[30] refined through MD simulations and validated using several approaches.

The topology of the hP-gp model was generated according to the GROMOS 54a7[142,143] parameter set of GROMOS96[144,145] force field. A previous equilibrated lipid bilayer patch comprising 469 molecules of 1-palmitoyl-2-oleoyl-phosphatidylcholine (POPC) was used with the lipid parameterization by Poger et al.,[146,147]. The hP-gp homology model was inserted into a lipid bilayer (with the longer P-gp axis perpendicular to the *xy* membrane plane) by matching the hydrophobic belt of the transmembrane domains (TMDs) with the POPC membrane. The relative position of the membrane was obtained from the Orientations of Proteins in Membranes (OPM)[148] database (http://opm.phar.umich.edu) and protein insertion was achieved through the *g_membed*[149,150] module in GROMACS. The protein embedded into the lipid bilayer was centered in a simulation box with dimensions *xyz* of 12.76 × 12.76 × 16.50 nm^3 and periodic boundary conditions (PBC)[151] were applied to all dimensions. Finally, the protein-membrane system was solvated with 63,841 water molecules with Single-Point Charge (SPC)[152] parameterization and 11 chlorine ions were added to neutralize the global charge of the system, using the *genion* module available in GROMACS.

2.1.2.2 MD simulations: equilibration and production runs

After the setup of the system, and to remove unfavorable clashes between atoms, an energy minimization run was applied using the steepest descent method. Then, the temperature of the membrane system (set at 303 K, above the lipids' gel to liquid-crystalline state phase transition temperature)[153] was equilibrated for 10 ps in the *NVT* ensemble — number of atoms (*N*), volume (*V*) and temperature (*T*) of the system were kept constant — spatially restraining all protein´s heavy atoms. Following, the POPC lipid bilayer was allowed to correctly adjust to the protein interface through a 20 ns *NpT* run, in which the number of atoms (*N*), pressure (*p*) and temperature (*T*) were kept constant, still keeping the protein's heavy atoms spatially restrained. Finally, three sequential 500 ps *NpT* runs were performed to progressively remove the protein's heavy atoms spatial restrictions (main-chain, backbone and alpha-carbons, respectively). This system was the starting point for a 200 ns fully unrestrained *NpT* production run.

2.1.2.3 Model quality assessment

The stability of the hP-gp model was monitored along the 200 ns MD run through the evolution of the root-mean-square deviation (RMSD) of the alpha-carbons (Cα), visual inspection and the MolProbity[154,155] evaluation server. During the simulation several additional MD runs were derived from the first one, and only those hP-gp models that were found to have a stable RMSD during the first 200 ns were extended to 400 ns of simulation time. Following, these models were submitted to extensive evaluation through

additional online servers as ERRAT,[156] PROCHECK[157,158] and SwissModel[159,160-162] Structure assessment tool. Lastly, the stability and quality of the final hP-gp models were also assessed considering the Ramachandran plot,[163] the RMSD of the Cα atoms, through visual inspection and by checking for correlations between molecular docking and experimental data. At the end, the most stable and reliable hP-gp WT model that performed best in the set of tests and validations, was further used in all subsequent studies described below.

2.1.2.4 Human P-gp variants: Insights on drug binding and signal-transmission mechanism

To deeply understand the mechanisms of drug specificity and efflux-related signal-transmission mechanism, the impact of four P-gp variants — G185V,[105,132,164-168] G830V,[105,167] F978A[133] and ΔF335[134,169-171] on P-gp architecture/organization and drug binding were comprehensively assessed by comparison the structural differences between the WT and each one of the variants. The considered P-gp mutations are located at the TMDs surrounding the three previously defined drug-binding sites (DBSs), and are experimentally linked with changes in efflux and substrate specificity — G185V (TMH3, H-site), G830V (TMH9, R-site), F978A (TMH12, M-site), and ΔF335 (TMH6, M-site) (**Figure 2.1**). The MD studies on these P-gp variants were firstly performed in the absence of ligands within the drug-binding pocket (DBP) (*apo* systems), in order to evaluate the structural impact of P-gp mutations on the P-gp architecture.
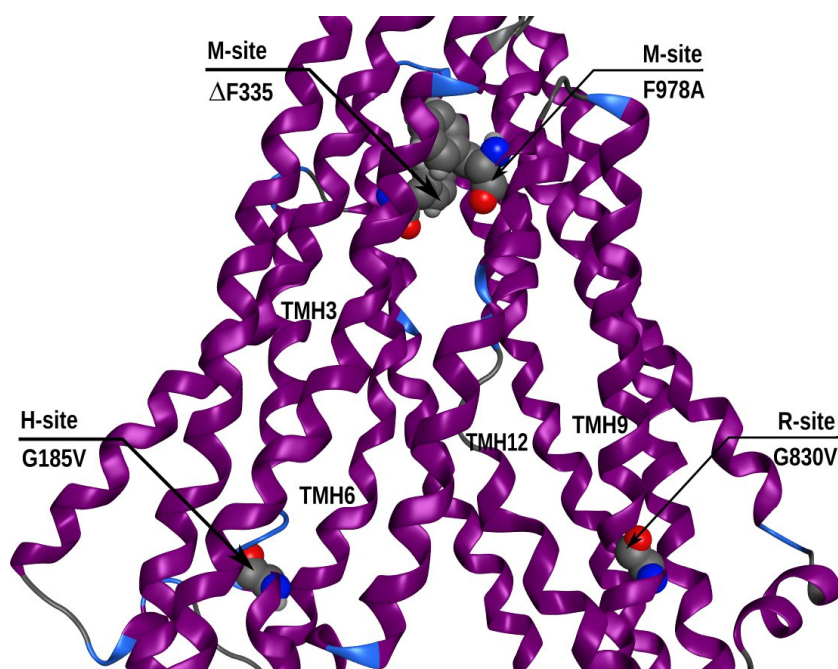
Figure 2.1. Location of the selected mutations in the transmembrane domains of the human P-gp model.

2.1.2.4.1 Human P-gp variants *apo* systems

From the final refined hP-gp WT model, the hP-gp variants mentioned above were built by mutating specific residues in the P-gp structure using MOE (*Protein Builder* module). Each P-gp variant was then re-embedded into a POPC membrane, solvated and the systems' charge neutralized with an adequate number of water molecules and counterions, respectively as described above (cf. *2.1.2.1*). Energy minimization runs comprising the whole systems were then performed, followed by a 10 ps *NVT* run at 303 K by spatially restraining all protein's heavy atoms. Fully unrestrained *NpT* runs followed for 100 ns. After 50 ns of simulation time, two replicates were derived for all P-gp variants and simulated for another 50 ns by randomly generating initial velocities, assigned from the correct temperature-dependent Maxwell-Boltzman distribution. This way, for each P-gp variant, three replicates were obtained, in a total of 200 ns of simulation time (**Figure 2.2**).
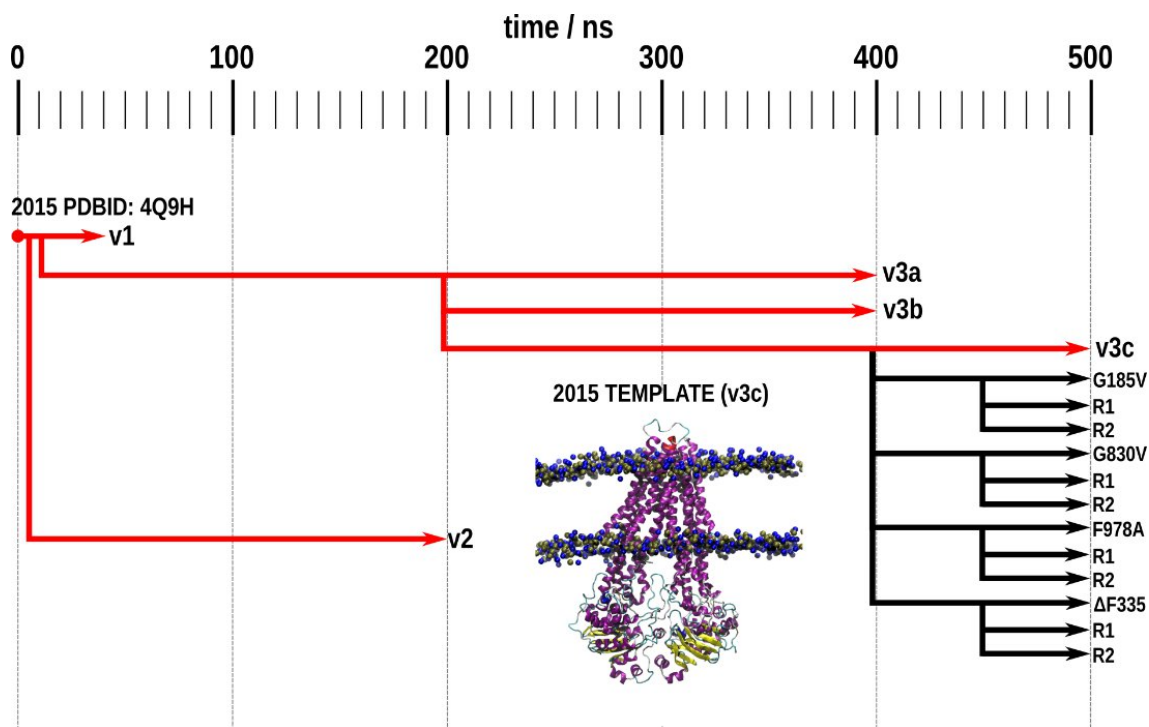
Figure 2.2. Schematic representation of the fully unrestrained *NpT* runs performed to obtain a stable human P-gp homology model and variants. The murine template used are colored as red (PDB ID:4Q9H), and the mutated structures in black.

2.1.2.4.2 Structural analysis of *apo* hP-gp WT and variants systems

The direct impact of mutations in the transmembrane region was assessed by evaluating the volume of the internal cavity and the packing of the transmembrane helices. The DBP volume (only considering the transmembrane helical bundle buried within the membrane) was estimated as the sum of the volumes of all water molecules found inside

this cavity through in-house developed python scripts. The effect of the mutations on TMHs rearrangement was analyzed through the *gmx bundle* module available in GROMACS. Moreover, the EPOS[BP 172,173] software was used to characterize the DBSs found within the DBP of the hP-gp WT model and variants. In addition, and based on the molecular docking results (cf. *Paper I; Results and discussion*), the top-ranked docking poses of each molecule in each DBS were then overlapped with the cavity search results identifying, this way, lining atoms (within a distance of 5 Å from the pocket probes) and calculating mean pocket volumes and polarities (ratio of the sum of N, O, and S atoms to the sum of N, O, S, and C atoms). Finally, and as ICHs located at the TMD-NBD interfaces are thought to play a key role in the transmission of the conformational changes needed to shift P-gp from the inward- to outward-facing conformations, the possible effect of mutations in the ICH-NBD interactions were evaluated by assessing the total number of contacts between the ICHs residues and the respective NBD using the *gmx hbond*[174] module. Additionally, the contact frequencies between ICH-NBD residues were also estimated by the *g_contacts*[175] module, only available in GROMACS v4.6. Visual inspections were performed with the Visual Molecular Dynamics (VMD)[176] and MOE software. For comparison purposes with the P-gp variants, the refined hP-gp WT model MD run was extended to additional 100 ns in a total of 500 ns of simulation time (**Figure 2.2**). All analysis were performed using the last 50 ns of each simulation in both WT and variants.

2.1.2.4.3 Human P-gp variants *holo* systems

Following the studies on drug specificity and TMD-NBD signal-transmission mechanism, the structural impact of these P-gp mutations on drug binding and efflux mechanism was assessed in the presence of molecules into DBP (*holo* systems), that are experimentally described as having changes in their efflux profiles in these variants. The ligands were chosen based on i) location of the mutation at the TMDs, ii) experimental and computational evidences for its putative DBS,[41,61,68,177-180] and iii) opposite drug efflux profiles induced by the same P-gp mutation (cf. *Paper II; Results and Discussion*). Therefore, colchicine (COL) and vinblastine (VIN) were chosen as substrates for the G185V as both are predicted to bind to the H-site as substrates. Similarly, doxorubicin (DOX) and actinomycin D (ACT), both predicted to interact with the R-site, were selected for the G830V (substrates, R-site). On the other hand, and as a significant impact on the H and R site properties and drug binding was reported for the F978A and ΔF335 variants (cf. *Paper I; Results and Discussion*), three different ligands were selected for the F978A variant, namely the P-gp modulator cyclosporine D (CYC, M-site), VIN (H-site) and ACT (R-site). Finally, and concerning what is known, there are no experimental reports about

the effect of the ΔF335 mutation in the efflux of molecules predicted to bind the H-site. Hence, CYC and valspodar (VLS), both predicted to bind at the M-site, were chosen while DOX was selected for the R-site. The top-ranked docking poses for the selected molecules were also inserted in the WT model for comparison purposes.

2.1.2.4.3.1 MD systems and simulations

The final configuration of the hP-gp WT (500 ns), and each P-gp variant (100 ns) were used as starting points for several short MD runs. The best-ranked docking poses of the ligands selected (cf. *Paper I; Results and Discussion*) were used to build the WT and variants *holo* systems. The molecules were parametrized using the PRODRG[181] online server and manually curated, according to the GROMOS96 54a7 force field and adjusting partial charges to the AM1-BCC[182] scheme calculated with antechamber v1.27.[183] Each protein-ligand complex was then embedded into a POPC membrane, water solvated and charge neutralized as described above (cf. *2.1.2.1*). After a short energy minimization run to remove clashes between the ligand and the protein, a total of 5 replicates were built for the WT model and each P-gp variant, per molecule. Short 20 ns MD run was applied to each system in a total of 1.5 μs of simulation time (0.1 μs each MD system). Three of the replicates were created by randomly generating initial velocities, assigned from a Maxwell-Boltzmann distribution at 303 K and, concerning the last two replicates, distinct snapshots were retrieved from the first MD simulation (at 18 ns and 19 ns, respectively).

2.1.2.4.3.2 Structural analysis of *holo* hP-gp WT and variants systems

The structural impact of mutations in the presence of ligands were assessed through the analysis of the TMH bundle using the *g_bundle* module available in GROMACS. Specific protein-ligand interactions were additionally estimated with the *g_contacts* module. Similarly to the *apo* systems, the total number of contacts and the contact frequencies between the ICH-NBD residue pairs were estimated by the *g_contacts* and *gmx hbond* modules. Finally, relative free-energies of binding (ΔG_{MD}) were calculated using the *g_mmpbsa*[184] tool, with polar solvation energies corrected by generating ion-accessibility and dielectric maps incorporating membrane environment (dielectric slab constant is set to 2.0 using the *draw_membrane2* program, http://www.poissonboltzmann.org) through in-house python scripts.[185] Only the last 10 ns from each replicate were considered as production runs and used for analysis.

2.1.2.5 Computational Fragment-Based Drug Discovery (cFBDD): Searching for putative drug-binding sites at the ICH-NBD interfaces

As several studies point to the existence of possible allosteric binding site(s) (aDBSs) near the ICHs, able to bind small molecules that potentially block the TMD-NBD efflux-related signal-transmission, a cFBDD approach based on known ICH-NBD binders was performed in order to accurately identify the hypothetical aDBSs at the TMD-NBD interfaces.

2.1.2.5.1 Fragment selection and parametrization

Fragments were obtained from predicted ICH-NBD binders such as: i) 2-[4-(diethylamino)-2-hydroxybenzoyl]benzoic acid (BUM) and 1-N-naphtylphtalamic acid (NPA), experimentally predicted to bind in a region between the ICHs and the Q-loop of the respective NBD,[122] or flavonoids as narigenin that were proposed to bind at the vicinal region of the ATP-binding sites,[94] and ii) from *in silico* studies of molecules that also have one energetic minimum at a similar regions (e.g. tariquidar, nicardipine, isoxazol-DHP and morphine) (although experimentally a photolabeling site for dihydropyridines is also proposed to exist in close proximity to such motifs).[117,121,186] Preference was given to aromatic systems with additional moieties (e.g. carboxylic acid, alcohols, amides, nitriles) and substitution patterns to maximize the coverage of the intended chemical space. The fragments obtained from each molecule are depicted in **Figure 2.3**. The selected fragments were drawn in MarvinSketch[187] v19.18 and exported to MOE[188] software for protonation and energy minimization, using the MMFF94x[189] force-field (adjusting hydrogen and lone pairs by default). Following, each fragment was parametrized in the PRODRG online server, and manually curated as previously described (cf. *2.1.2.4.3.1*).

2.1.2.5.2 Construction of the MD systems

Based on the previously published human P-gp homology model (cf. *Paper I; Results and discussion*), only the cytoplasmic portion of the N- and C-terminal (concerning NBD1 and NBD2, respectively) were used. The preparation of the initial structures was performed using VMD and MOE software packages. Each NBD was inserted in a simulation box with dimensions *xyz* of $10 \times 10 \times 9$ nm^3. Following, all systems were solvated with 27,975 (NBD1) or 27,994 (NBD2) SPC water molecules and 6 chlorine ions to neutralize the global charge of the system. For each NBD, five MD systems were built, each one comprising six fragment types and five copies of each (for a total of 30

fragments), by randomly inserted them in the surrounding water environment using *gmx insert-molecules* module available in GROMACS v2016.x software. Any overlapping waters were automatically removed, and any charge excess were neutralized by adding an adequate number of counter-ions (Na⁺ or Cl⁻). The fragments included in each MD system are depicted in **Table 2.1**.
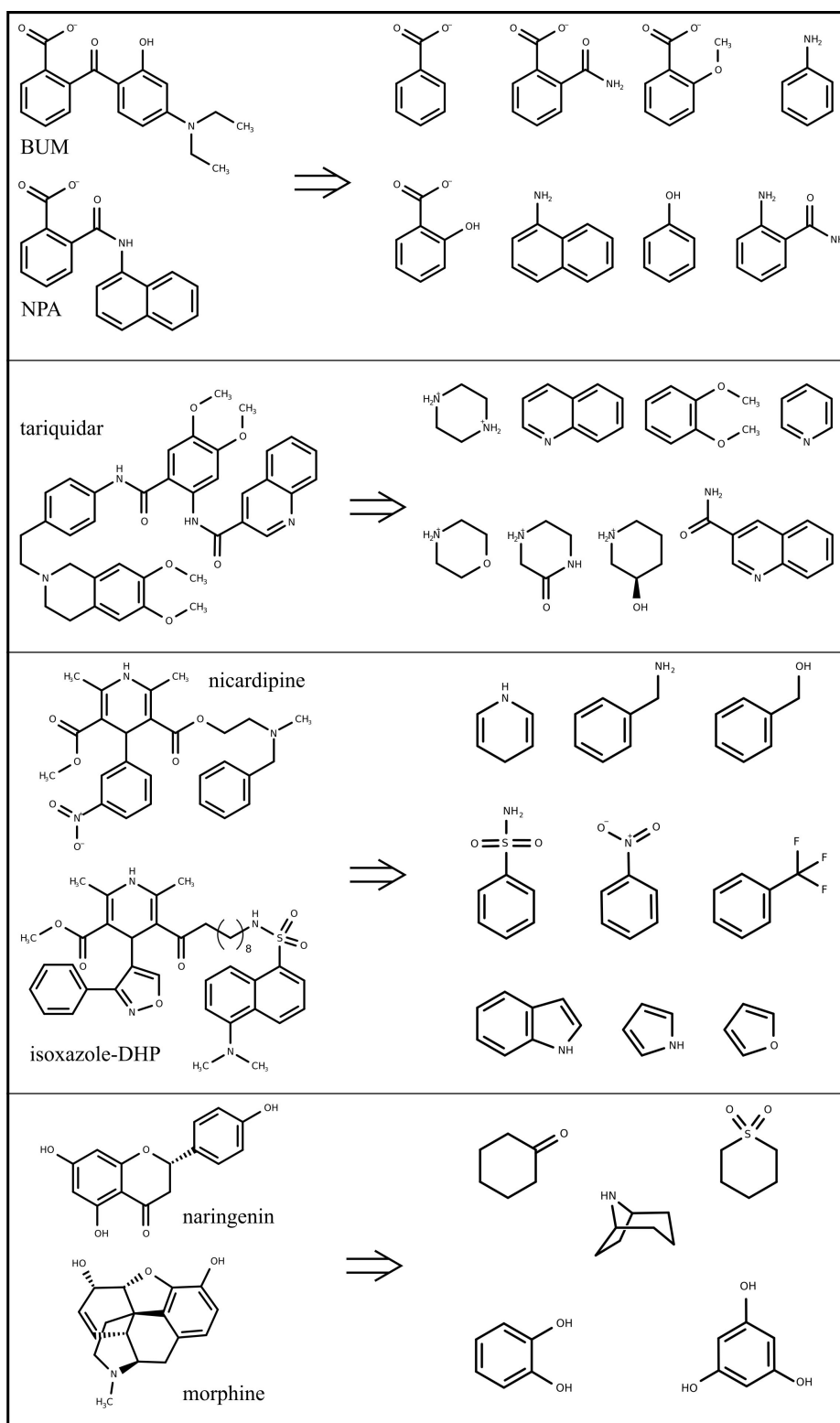

Figure 2.3. Fragments obtained from each molecule.

Table 2.1. Fragments included in each MD system.

System	Fragments (IUPAC name)	Formula	SMILES
1	Naphthalen-1-amine	$C_{12}H_{15}N$	NC1=CC=CC2=CC=CC=C12
	Aniline	C_6H_7N	NC1=CC=CC=C1
	Pyridine	C_5H_5N	C1=CC=NC=C1
	Benzoate	$C_7H_5O_2^-$	[O-]C(=O)C1=CC=CC=C1
	Phenol	C_6H_6O	OC1=CC=CC=C1
	1,4-dihydropyridine	C_5H_7N	C1C=CNC=C1
2	Benzene-1,2-diol	$C_6H_6O_2$	OC1=C(O)C=CC=C1
	Benzene-1,3,5-triol	$C_6H_6O_3$	OC1=CC(O)=CC(O)=C1
	1H-indole	C_8H_7N	N1C=CC2=C1C=CC=C2
	Nitrobenzene	$C_6H_5NO_2$	[O-][N+](=O)C1=CC=CC=C1
	1H-pyrrole	C_4H_5N	N1C=CC=C1
	Quinoline-3-carboxamide	$C_{10}H_8N_2O$	NC(=O)C1=CC2=CC=CC=C2N=C1
3	2-carbamoylbenzoate	$C_8H_6NO_3^-$	NC(=O)C1=C(C=CC=C1)C([O-])=O
	1,2-dimethoxybenzene	$C_8H_{10}O_2$	COC1=C(OC)C=CC=C1
	Furan	C_4H_4O	O1C=CC=C1
	Phenylmethanamine	C_7H_9N	NCC1=CC=CC=C1
	Quinoline	C_9H_7N	C1=CC2=CC=CN=C2C=C1
	Trifluormethylbenzene	$C_7H_5F_3$	FC(F)(F)C1=CC=CC=C1
4	2-aminobenzamide	$C_7H_8N_2O$	NC(=O)C1=C(N)C=CC=C1
	2-carboxyphenolate	$C_7H_5O_3^-$	OC1=C(C=CC=C1)C([O-])=O
	Cyclohexanone	$C_6H_{10}O$	O=C1CCCCC1
	2-methoxybenzoate	$C_8H_7O_3^-$	COC1=C(C=CC=C1)C([O-])=O
	Phenylmethanol	C_7H_8O	OCC1=CC=CC=C1
	Piperazine-1,4-diium	$C_4H_{12}N_2^{2+}$	C1C[NH2+]CC[NH2+]1
5	Morpholin-4-ium	$C_4H_{10}NO^+$	C1COCC[NH2+]1
	Benzenesulfonamide	$C_6H_7NO_2S$	NS(=O)(=O)C1=CC=CC=C1
	piperidin-1-ium-3-ol	$C_5H_{12}NO^+$	OC1CCC[NH2+]C1
	3-oxopiperazine-1,4-diium	$C_4H_9N_2O^+$	O=C1C[NH2+]CCN1
	8-azoniabicyclo[3.2.1]octane	$C_7H_{14}N^+$	C1CC2CCCC1[NH2+]2
	Thiane 1,1-dioxide	$C_5H_{10}O_2S$	O=S1(=O)CCCCC1

2.1.2.5.3 Simulated-annealing MD simulations (saMD)

After an energy minimization step, a short 10 ps *NVT* ensemble run was performed followed by a 5 ns *NpT* ensemble MD simulations for equilibrating temperature (303 K) and pressure (1 bar), respectively, while keeping all protein atoms spatially restrained. After, a 50 ns MD run was performed for each system, using a simulated annealing protocol applied to the fragments, solvent and ions to decrease the probability of non-specific binding to each of the NBDs. Herein, two heating cycles (303 K to 323 K) were performed, increasing the temperature during 1 ns (*t* = 0 ns and *t* = 25 ns) and decreasing over a period of 3 ns (*t* = 2 ns and *t* = 27 ns). PBC conditions were applied to all

dimensions. Due to the simulated annealing procedure, the protein´s alpha carbons were kept restrained throughout the whole MD simulation to prevent unfolding. For each system, five replicates were performed (25 MD systems per NBD, in a total of 1.25 µs of simulation time per NBD).

2.1.2.5.4 Analysis and identification of allosteric drug-binding site(s)

The last 20 ns of each replicate were concatenated using *gmx trajcat* tool available in GROMACS to obtain a single trajectory file of 100 ns per system (in a total of 500 ns per NBD). After loading the trajectory into VMD, occupancy volumetric maps (3D grids containing a value at each grid point, set to either 0 or 1, and based on the atomic coordinates of each fragment type), were computed using all frames with the *VolMap* tool in VMD, combined using average values (thus providing the fractional occupancy of each grid point) and exported as DX-formatted file. An isosurface cut-off of 0.1 was used to fine-tune the generated occupancy maps. Visual inspection of the fragments' occupancy maps was performed in VMD, and only those in close vicinity of each ICH-NBD interface were considered for the identification of possible binding sites next to the ICH domains. Finally, the properties of each putative aDBS such as lining residues, pocket volume, residues distribution and mean polarities were evaluated using the EPOS[BP] and MOE packages (default parameters), and compared with the modulator site (M-site).

2.1.2.5.5 Free-energy calculations and analysis of compound 23
(for further details, refer to: Letter, Results and Discussion)

After a short energy minimization run to prevent clashes between the ligand and the protein, a 100 ns MD run was performed for the top-ranked docking pose of compound **23** at the NBD2, using the final configuration of the NBD2 after the saMD simulations. From the initial MD simulation, four replicates of 50 ns MD run each were performed by selecting distinct snapshots along the simulation (based on the RMSD evolution during the first MD run), namely at 40 ns, 70 ns and 90 ns MD simulation times. Only the last 30 ns of MD simulation were considered as production run and used for analysis. The visual inspection of the binding mode(s) was performed in MOE, and relative free-energies of binding (ΔG_{bind}) were calculated using the *g_mmpbsa* tool in GROMACS.

2.1.2.6 MD simulation parameters

All MD simulations herein described were done with GROMACS v2016.x and v2020.x packages. All *NVT* equilibration runs were performed at 303 K using the Velocity-rescale (V-rescale)[190] thermostat. The Nosé-Hoover[191,192] thermostat and the Parrinello-

Rahman[193] barostat for temperature (303 K) and pressure (1 bar) control, were applied in all *NpT* runs. Pressure equilibration was achieved through a semi-isotropic (for membrane systems) or isotropic (concerning only the cytoplasmic domains) pressure coupling, with the systems' compressibility set to 4.5×10^{-5} bar^{-1}. All bond lengths were constrained using the LINCS[194,195] or SETTLE[196] (for water molecules) algorithms. PBC conditions were applied to all dimensions. The Particle Mesh Ewald (PME)[197,198] with cubic interpolation was employed, with a cut-off radius of 12 Å for both electrostatic and van der Waals interactions and a FFT grid spacing of 0.16 for long range electrostatics. Group-based and Verlet[139] cut-off schemes were applied for the calculation of non-bonded interactions on CPU or GPU, respectively.

2.1.3 Molecular Docking studies

2.1.3.1 Canonical drug-binding pocket

Molecular docking was performed using the final configuration of the hP-gp WT model and of all variants (G185V, G830V, F987A and ΔF335). This procedure aimed to: i) validate the refined hP-gp model, by comparing the number and location of the DBSs within the DBP with those observed in a previously refined mP-gp structure and ii) evaluate the impact of the studied P-gp mutations on drug binding by comparing each P-gp variant to the human WT model.

Molecular docking studies were carried out in the AutoDock VINA v1.1.2[199] docking software. Ligands were built in MOE, energy minimized to the closest local minima using the MMFF94x force-field (adjusting hydrogen and lone pairs by default), and saved in PDB format. For further usage in VINA, both receptor (P-gp) and ligands were converted to the PDBQT format with AutoDockTools v1.5.6rc[200] scripts. The chosen databases comprised P-gp substrates (*n* = 33), probes (*n* = 7) and modulators (*n* = 19), previously used in docking studies with the refined mP-gp structure.[41] The drug-binding site was defined through a docking box encompassing the whole internal cavity identified by Aller et al.,[16] with dimensions *xyz* of 32.25 × 26.25 × 37.50 Å^3, and centered at the DBP at coordinates *xyz* of 62.4 × 61.6 × 102.3 Å^3 (*xy* corresponds to the membrane plane) for the hP-gp WT model. For each P-gp variant, the docking box dimensions and box center were adjusted according to the DBP volume. Due to the large search volume (over 30,000 Å^3), VINA's 'exhaustiveness' parameter was manually set to 50 while 20 docking poses were generated. Visual inspection of the ligands' docking poses was made in MOE to allow the identification of individual docking zones, taking into account the

location of the three distinct DBSs within the DBP, and the lining residues predicted for each DBSs in the refined mP-gp structure.

2.1.3.2 Allosteric drug-binding site(s)

Molecular docking were performed in the AutoDock VINA v1.1.2 in both isolated NBDs. Considering the location of the selected occupancy maps (cf. *2.1.2.5.4,* and *Letter I, Results and Discussion*), a docking box was defined to include all regions between the ICHs and the whole ATP-binding site, with grid dimensions *xyz* of 18.8 × 18.8 × 18.8 Å^3 (NBD1) or 18.8 × 22.5 × 18.8 Å^3 (NBD2), and centered at the *xyz* coordinates 40.2 × 53.2 × 55.3 Å^3 (NBD1) or 63.4 × 47.6 × 55.7 Å^3 (NBD2), respectively. The database consisted in small in-house libraries of flavanone (*n* = 28)[126,201] and thioxanthone (*n* = 8)[202] derivatives. Additionally, due to experimental evidences that BUM and NPA bind in a region located at the ICH-NBD interfaces, these molecules were also included in the docking studies as references. Ten docking poses were generated for each ligand. Manipulation of the receptor and ligands was performed as previously described (cf. *2.1.3.1*). Visual inspection of the docking poses was made in MOE, and the protein-ligand interactions were calculated using BINANA v2.1[203] and LigPlot v4.5.3[204] software packages.

2.2 Experimental studies: ATPase assays

As most of the compounds tested in the docking studies described above (cf. *2.1.3.2*) bound at the ICH-NBD interfaces with favorable binding energies (cf. *Letter, Results and Discussion*), the ability of each test compound to inhibit P-gp ATPase activity was experimentally assessed. Accordingly, any compounds that interact with P-gp and promote its ATPase activity are considered P-gp activators, while molecules interacting with P-gp but reducing it are classified as P-gp inhibitors.[205]

2.2.1 General concepts

The ATPase activity of the P-gp ATP-dependent drug efflux pump was measured by using the P-gpGlo™ Assay Systems kit (Promega, The Netherlands) according to the manufactures' recommendation. Briefly, this assay relies on the ATP dependence of the light-generating reaction of firefly luciferase using recombinant human P-gp in a cell membrane fraction.[205] The recombinant human P-gp in cell membrane fraction (25 µg/well) was incubated with 5 mM of MgATP for a period of 40 minutes at 37 °C. The reaction was stopped by adding the luciferase-based ATP Detection Reagent. At this stage, luciferase reaction was initiated, and the luminescence signal was produced. Luminescence changes in direct proportion to the unmetabolized ATP concentration

(i.e. ATP not consumed by P-gp). Reactions where no ATP is consumed generate the brightest signals (e.g. sodium orthovanadate – Na_3VO_4) while reactions where ATP has been consumed have relatively lower luminescent signals, reflecting a lower non-hydrolyzed ATP concentration (e.g. in the presence of verapamil).

2.2.2 Selection of compounds

A small library of thioxanthone (n = 8) and flavanone (n = 28) derivatives (cf. *2.1.3.2*) was used in the ATPase assays. All the compounds were evaluated at the initial concentration of 100 µM (flavonoids) and 200 µM (thioxanthones). Additionally, the compounds BUM (1 mM) and spiropedroxodiol (50 µM) were also included in the ATPase assays as positive controls for allosteric[122] and competitive[206] P-gp inhibition, respectively. All compounds were dissolved in DMSO.

2.2.3 ATPase assays layout and analysis of the results

The assay was performed in a 96-well plate layout, which comprised four different sample types: i) non-treated (NT), ii) treated with Na_3VO_4, iii) treated with verapamil (VER), and iv) treated with test compound (TC) in the presence of verapamil (TC+VER). All the samples were tested in triplicate, in accordance with the technical bulletin. A first screening was performed to evaluate the ability of compounds to inhibit the P-gp drug-stimulated ATPase activity in the presence of 500 µM of verapamil, a P-gp substrate that stimulates P-gp ATPase activity.[207] In the following step, the compounds that showed the highest inhibition of the VER-stimulated ATPase activity were selected for calculating the respective maximal inhibitory concentration (IC_{50}) values by decreasing the test compounds' concentration using a two-fold dilution, while maintaining verapamil concentration unchanged (500 µM), in agreement with the experimental protocol provided by the manufacturer. The luminescence signal was read in average relative light units (RLU), and compared with the Na_3VO_4 – a strong P-gp inhibitor[208] – treated samples. The difference between the average luminescent signals from Na_3VO_4-treated samples (RLU_{Na3VO4}) and NT samples (RLU_{NT}) corresponds to the basal P-gp ATPase activity, and was calculated as follows:

$$RLU_{Na_3VO_4} - RLU_{NT} = \Delta RLU_{basal} \quad \text{(eq. 1).}$$

The difference between the average luminescent signals from Na_3VO_4-treated samples (RLU_{NA3VO4}) and VER-treated samples (RLU_{VER}) reflects the VER-stimulated ATPase, and can be determined as follows:

$$RLU_{Na_3VO_4} - RLU_{VER} = \Delta RLU_{VER} \quad \text{(eq. 2).}$$

Finally, the difference between the average luminescent signals from Na_3VO_4-treated samples (RLU_{NA3VO4}) and TC+VER-treated samples (RLU_{TC+VER}) reflects the ability of TC to inhibit VER-stimulated ATPase activity, and can be calculated by the same equation,

$$RLU_{Na_3VO_4} - RLU_{TC+VER} = \Delta RLU_{TC+VER} \quad \text{(eq. 3).}$$

CHAPTER 3 – PAPER I

Theoretical insights on helix repacking as the origin of P-glycoprotein promiscuity

P-glycoprotein (P-gp, ABCB1) over-expression is, currently, one of the most important multidrug resistance (MDR) mechanisms in tumor cells. Thus, modulating drug efflux by P-gp has become one of the most promising approaches to overcome MDR in cancer. Yet, more insights on the molecular basis of drug specificity and efflux-related signal-transmission mechanism between the transmembrane domains (TMDs) and the nucleotide binding domains (NBDs), are needed to develop molecules with higher selectivity and efficacy. Starting from a murine P-gp crystallographic structure at the inward-facing conformation (PDB ID: 4Q9H), we evaluated the structural quality of the herein generated human P-gp homology model. This initial human P-gp model, in the presence of the "linker" and inserted in a suitable lipid bilayer, was refined through Molecular Dynamics simulations and thoroughly validated. The best human P-gp model was further used to study the effect of four single-point mutations located at the TMDs, experimentally related with changes in substrate specificity and drug-stimulated ATPase activity. Remarkably, each P-gp mutation is able to induce transmembrane α-helices (TMHs) repacking, affecting the drug-binding pocket volume and the drug-binding sites properties (e.g. volume, shape and polarity) finally compromising drug binding at the substrate binding sites. Furthermore, intracellular coupling helices (ICHs) also play an important role since changes in the TMHs rearrangement are shown to have an impact in residue interactions at the ICH-NBD interfaces, suggesting that identified TMHs repacking affect TMD-NBD contacts and interfere with signal-transmission from the TMDs to the NBDs.

3.1 INTRODUCTION

Multidrug resistance (MDR) to anticancer drugs is, at the moment, a major contributor to chemotherapy failure.[209] In cancer, one of the most significant MDR mechanisms results from the over-expression of P-glycoprotein, a membrane efflux pump (P-gp, ABCB1). Thus, a deeper understanding on P-gp substrate specificity and efflux-related signal-transmission mechanism remains crucial for the development of more potent and selective compounds able to modulate drug efflux.[210] P-glycoprotein exports a broad range of structurally unrelated compounds through an ATP-dependent mechanism.[211] P-gp is organized in two homologous functional units (N- and C-terminal halves) with a pseudo-2-fold symmetry. Each halve comprises one transmembrane domain (TMD), formed by six transmembrane α-helices (TMHs), and one cytoplasmic nucleotide-binding domain (NBD). Both N- and C-terminal halves are connected by a small peptide sequence (the "linker"; residues 627-688). The TMHs are directly linked to the respective NBD by the intracellular loops, through the functional TM helices 6 (NBD1) and 12 (NBD2) and non-covalently by short intracellular coupling helices (ICHs), located between the structural TMHs 2/3 (ICH1-NBD1), 4/5 (ICH2-NBD2), 8/9 (ICH3-NBD2) and 10/11 (ICH4-NBD1) (**Figure 3.1**).
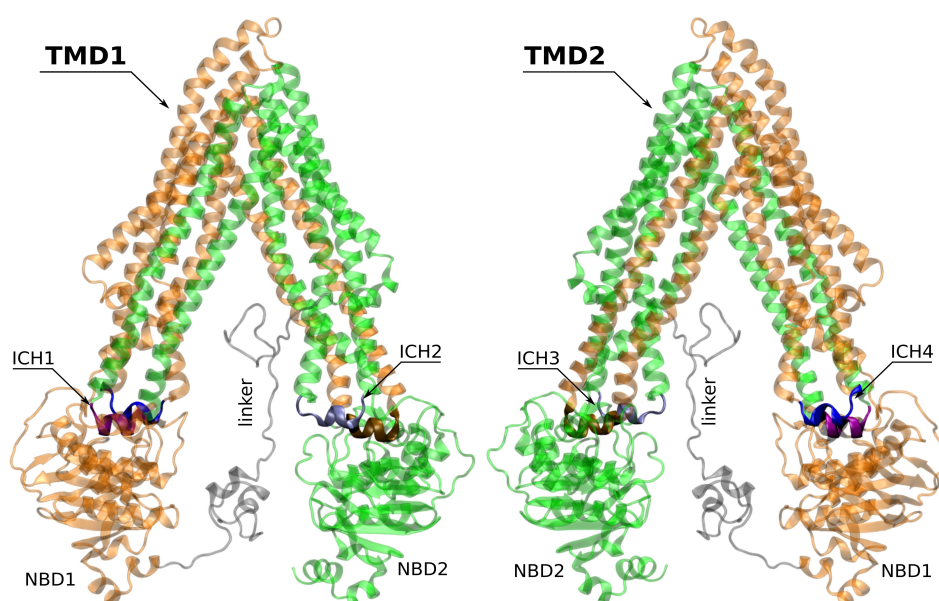
Figure 3.1. Structural representation of the human P-glycoprotein in an inward-facing conformation. The 12 transmembrane α-helices (TMHs) are divided into two transmembrane domains (TMD1; orange and TMD2; green), being physically linked to the respective nucleotide-binding domain (NBD) by coils bridging TMH6/NBD1 and TMH12/NBD2 and also by non-covalent interactions involving short intracellular coupling helices: ICH1 (purple) /ICH4 (blue) with NBD1 and ICH2 (silver) /ICH3 (brown) with NBD2. The DBP is a large cavity between both TMDs. Figures were created with MOE from the final human P-gp homology model.

These ICHs were found to be important for the maturation and folding of the P-gp transporter, being also involved in the signal-transmission pathway between the TMDs and NBDs.

The drug-binding pocket (DBP) is a large cavity formed by the TMHs of both N- and C-terminal P-gp halves and is capable of recognizing and accommodating several structurally distinct substrates. Each NBD contains the catalytic site for ATP-binding and hydrolysis. Several experimental[16,20,24,39,40,48,51,53,55,60-62,65,72,97,115,179,212-221] and computational[30,41,88,98,100,222-225] studies were performed to better understand the details of drug specificity and efflux. However, the information gathered from these experiments are often scarcely related and the mechanisms of drug recognition/specificity and efflux are still unsolved. Therefore, for additional insights on P-gp efflux mechanism, the study of P-gp variants that are experimentally related with altered drug-resistance phenotypes, and changes in the ATPase activity is an interesting approach.

The first mutation in the human P-gp was identified in the colchicine-selected multidrug-resistance cell line (KB-C1)[132] and comprises a substitution of a glycine by a valine residue in position 185 (G185V; TMH3). This mutation is reported to confer increased resistance to colchicine, etoposide, doxorubicin and puromycin while decreasing the resistance to vinblastine, vincristine, taxol and actinomycin D. The same study reported, unexpectedly, a 3.6-fold decrease in colchicine binding and a 3.8 to 5.5-fold increase in vinblastine binding when compared with the WT transporter, leading the authors to conclude that G185V mostly affects the dissociation of drugs from P-gp rather than its initial association. Later studies additionally concluded that the basal ATPase activity and the degree of activation by substrates was also increased in G185V, with only minor changes on nucleotide binding.[165-167] As it was also reported that G185V mutant is able to transport colchicine and etoposide in an energetically more efficient way, the residue at position 185 was identified as pivotal for transmitting conformational changes between the catalytic sites and colchicine-binding site.[168]

Similar alterations were observed in another glycine variant (G830V; TMH9), obtained by site-directed mutagenesis, in the presence of verapamil and colchicine.[105,167] Herein, increased resistance to colchicine (3.3-fold) and decreased resistance to actinomycin D (0.29-fold) was reported, but no significant change on vinblastine resistance was observed when compared with the WT protein. Its verapamil-stimulated ATPase activity was found to be only slightly increased (1.4 to 1.7-fold), but no change in the vinblastine-stimulated ATPase activity was found to occur. Nonetheless, the pattern of

drug-stimulated ATPase correlated well with the relative drug-resistance profiles in transfected cells. No data on basal ATPase activity was reported.

Likewise, mutations involving the phenylalanine residue at position 978 were also identified to play an important functional role in P-gp. By mutating F978 to an alanine (F978A) a significant alteration of the drug resistant pattern was observed, conferring little or no resistance to either colchicine or doxorubicin but not changing the resistance to vinblastine or actinomycin D.[133] This was further suggested to be due to a defect in the mutant protein, being unable to transport colchicine and having a reduced capability to transport vinblastine, even with an increased expression of the transporter. Quite interestingly, when testing other substitutions (F978S, F978L or F978Y), only the latter restored similar drug-resistance profiles for all tested molecules, but nonetheless azidopine labeling was indistinguishable from WT protein. Interestingly, by including the previously characterized G185V together with F978A/S was also unsuccessful in fully restoring the ability of conferring resistance to colchicine. In addition, little or no drug-stimulated ATPase activity by colchicine or verapamil was reported, being suggested that F978A mutants have either a decreased affinity for substrates and/or an impairment in coupling drug binding to ATPase activity.[105] Recent studies, using a F978C mutant, additionally showed complete absence of stimulation of ATP hydrolysis by several drugs, proposing that F978 residue is part of a common translocation route, crucial for the propagation of conformational changes following ATP hydrolysis and important for the translocation process of high-affinity substrates.[226] Again, no data was reported on basal ATPase activity for this mutant.

Finally, the deletion of a phenylalanine residue at position 335 (ΔF335; TMH6) was reported in a multidrug-resistant variant cell line (DxP), by co-selection with doxorubicin and the cyclosporine D analogue, valspodar (PSC833), a potent P-gp modulator.[134,169,170] The transfected cells harboring the ΔF335 P-gp variant showed to be resistant to many chemotherapeutic drugs such as doxorubicin, daunorubicin, etoposide and paclitaxel. On the other hand, this variant conferred little resistance to vincristine, vinblastine and actinomycin D as well as a decreased ability to bind or transport cyclosporine, valspodar, vinblastine, actinomycin D and rhodamine-123, suggesting that the region surrounding the F335 residue is an important binding site for these compounds.[170,171] An enhancement of photoaffinity binding by [^{125}I]-iodo arylazidoprazosin in the presence of verapamil or PSC833 and a decrease of [^{3}H]-cyclosporine binding was also referred. Moreover, by deleting the F335 residue a 2-fold increase in the basal ATPase was observed but, when in the presence of drugs, only verapamil-stimulated ATPase activity occured. Interestingly, this mutant also

presented a substantial decrease of 8-[α-^{32}P]azido-ATP labeling when compared with the WT P-gp, which lead the authors to suggest that the ATPase activity of the mutant protein may depend on the helix conformation defined by F335.

In this work, a human P-gp homology model in the *apo* inward-facing state conformation was generated based on the most recent murine P-gp crystallographic structure available at the beginning of our study (PDB IDs: 4Q9H).[22] The P-gp model was refined through molecular dynamics (MD) simulations and validated using several approaches. The final human P-gp model was further used to understand the possible structural impact of the mutations described above on P-gp architecture/organization.

3.2. MATERIAL AND METHODS

3.2.1 Human P-gp homology modeling

The FASTA sequence of human P-gp was obtained from the Universal Protein Resource (UNIPROT) under the code P08183 (www.uniprot.org). The murine P-gp crystallographic structure (PDB IDs: 4Q9H), used as template, was retrieved from the Protein Data Bank (PDB; www.rcsb.org). Both the murine P-gp crystallographic structure and the FASTA sequence human P-gp were loaded into the MOE software and aligned by their sequence. The "linker" secondary structure, missing in all P-gp crystallographic structures so far, was obtained from a previously equilibrated murine P-gp and used as template for modeling the respective sequence (A627-A688). Herein, the insertion of the "linker" in the novel structures was achieved by aligning the human P-gp homology model with the crystallographic template, followed by an "override" of the gap in the considered structure in which the murine P-gp linker was used as template to obtain the human homologue. Afterwards, 25 mainchain models sampling 25 sidechain orientations were performed, producing 625 models for the human P-gp structure in MOE (force field parameters used by default). The model with lowest total potential energy was selected, protonated using the *Protonate 3D* module and exported as PDB file to be used by GROMACS v5.0.7. This way, an initial human P-gp homology model in an inward-facing conformation incorporating the "linker" obtained from a previously equilibrated murine P-gp, was obtained.

3.2.2. Construction of the protein membrane system

The topology of the human P-gp model was generated according with the GROMOS96 54a7 force field. A previously equilibrated 1-palmitoyl-2-oleoyl-phosphatidylcholine

(POPC) membrane patch was used with the lipid parameterization by Poger et al.,[146,147] . The human P-gp homology model was inserted into a lipid bilayer (longer P-gp axis perpendicular to the xy membrane plane) to match the hydrophobic thickness of TMDs and membrane. The relative position of the membrane was obtained from the Orientations of Proteins in Membranes (OPM) database (http://opm.phar.umich.edu) and protein insertion was achieved through the *g_membed* module in GROMACS. The protein embedded into the lipid bilayer was centered in a simulation box with dimensions *xyz* of 12.76 × 12.76 × 16.50 nm^3 and periodic boundary conditions (PBC) were applied to all dimensions. Finally, the system was solvated and neutralized with an adequate number of water molecules and counter-ions using other GROMACS' modules.

3.2.3 Molecular Dynamics: equilibration and production run

Firstly, an energy minimization run comprising the whole system was applied using the steepest descent method. Then, the temperature of the membrane system (303 K) was equilibrated for 10 ps in the *NVT* ensemble, while spatially restraining all protein´s heavy atoms. Following, the POPC lipid bilayer was allowed to correctly adjust to the protein interface through a 20 ns *NpT* run, still keeping the protein's heavy atoms restrained. Finally, three sequential 500 ps *NpT* runs were performed to progressively remove the protein's heavy atoms spatial restriction (mainchain, backbone and alpha-carbons, respectively). This system was the starting point for a 200 ns fully unrestrained *NpT* production run (**Figure S1**, *Annex S2*).

3.2.4 Model quality assessment

The stability of the P-gp model was monitored along the MD run through the evolution of the root mean square deviation (RMSD) of the Cα atoms, by visual inspection and using the MolProbity evaluation server. After 200 ns of simulation time, more exhaustive evaluations were performed, through additional servers namely ERRAT, PROCHECK and SwissModel Structure assessment tool. Moreover, the stability and quality of the human P-gp model were also assessed considering the Ramachandran plot and by checking for correlations between molecular docking and experimental data. The evaluation of the recently published human cryo-EM P-gp structure was also performed, for comparison purposes.

3.2.5 Construction of the human P-gp mutated structures and systems

From the final refined human P-gp homology model, four human P-gp variants (G185V, G830V, F978A and ΔF335) experimentally linked with changes in efflux and substrate specificity were built using MOE. Each P-gp variant was then embedded into a POPC membrane, water solvated and charge neutralized as described above. Energy minimization runs comprising the whole system were applied followed by a 10 ps *NVT* run at 303 K by spatially restraining all protein's heavy atoms. Fully unrestrained *NpT* runs followed for 100 ns (**Figure S1**, *Annex S2*). After 50 ns of simulation time, two system replicates were obtained for all P-gp variant systems, each one simulated for another 50 ns by randomly generating initial velocities, assigned from the correct temperature dependent Maxwell-Boltzman distribution, and starting with the final configuration obtained at the end of the first 50 ns. This way, for each P-gp variant, three replica systems were therefore simulated in a total of 200 ns of simulation time.

3.2.6 Simulation parameters

All *NVT* equilibration runs were performed at 303 K using the Velocity-rescale thermostat (V-rescale). The Nosé-Hoover thermostat and the Parrinello-Rahman barostat for temperature (303 K) and pressure control (1 bar), were applied in all *NpT* runs. Due to the presence of membranes, pressure equilibration was achieved through a semi-isotropic pressure coupling, with the systems' compressibility set to 4.5×10^{-5} bar^{-1}. All bond lengths were constrained using the LINCS and SETTLE (for water molecules) algorithms. The Particle Mesh Ewald (PME) with cubic interpolation was employed, with a cut-off radius of 12 Å for both electrostatic and van der Waals interactions and an FFT grid spacing of 0.16 for long range electrostatics. Group-based and Verlet cut-off schemes were applied for the calculation of non-bonded interactions on CPU or GPU, respectively.

3.2.7 Structural analysis of the human P-gp variants

To evaluate the impact of each mutation in the volume of the internal DBP (only considering the transmembrane helical bundle buried within the membrane) and to allow a fast comparison between WT and variants, the DBP was estimated as the sum of the volumes of all water molecules found inside this cavity through *in-house* python scripts. To assess the effect of the mutations on TMDs rearrangement, the TM bundle was analyzed through the *g_bundle* module available in GROMACS. The total number

of contacts between the ICHs residues and the respective NBD were calculated using the *g_hbond* module and the contact frequencies were estimated by the *g_contacts* module. Furthermore, the EPOS[BP] software (default parameters) was used to characterize the DBSs found within the DBP of the human WT P-gp model and variants. The top-ranked docking poses of each molecule in each zone were then overlapped with the cavity search results identifying, this way, lining atoms (within a distance of 5 Å from the pocket probes) and calculating mean pocket volumes and polarities (ratio of the sum of N, O, and S atoms to the sum of N, O, S, and C atoms). Visual inspections were performed with VMD and MOE software. All analysis described above were performed using the last 50 ns of each simulation.

3.2.8 Docking studies

Molecular docking was performed using the final human WT P-gp model and the generated variants (G185V, G830V, F987A and ΔF335). The chosen databases comprised P-gp substrates (*N* = 33), probes (*N* = 7) and modulators (*N* = 19), previously used in docking studies with our refined murine P-gp structure. The ligands binding location was defined by a docking box comprising the whole internal cavity identified by Aller et al., with dimensions *xyz* of 32.25 × 26.25 × 37.50 Å^3, and centered at the DBP (*xy* corresponds to the membrane plane). Due to the large search volume (over 30.000 Å^3), Vina's 'exhaustiveness' parameter was manually set to 50 and twenty docking poses were generated. Visual inspection of the best ranked docking poses was made in MOE to allow the identification of individual docking zones.

3.3 RESULTS AND DISCUSSION

3.3.1 Human P-gp homology model development

Considering the high sequence identity and similarity with the human P-gp efflux pump (87% and 94%, respectively), a human P-gp homology model was obtained using the murine P-gp crystallographic structure of 2015 (PDB ID: 4Q9H) as template.

The murine P-gp crystallographic structure was chosen based on: i) it was the most recently published murine P-gp crystallographic structure in the beginning of our study, ii) it shows improvements in the resolution of several TMHs, ICH1 and some extracellular loops in respect to previous structures, stressing the quality of the starting template to the development of a reliable human P-gp model, and iii) oppositely to the recently published human P-gp structure, it was obtained without any ligands or

antibody complexes that could influence the native arrangement of the transmembrane helical domains.

After 14 ns of simulation time, and unlike the refined murine P-gp structure, a shift of the linker's upper loop downwards was observed in our human P-gp model (v1 model). Thus, to assess the correct position of the "linker" in the human P-pg model, two snapshots retrieved at 8 ns (v2 model) and 14 ns (v3 model) of simulation time were the starting points for two additional MD runs (**Figure S1**, *Annex S2*).

At the end of 200 ns MD simulations, a shift of the linker's upper loop downwards was also observed in both v2 and v3 models confirming the possibility that this new position acquired by the "linker" structure is favored in our model. This result also demonstrates the high flexibility of the "linker" region offering a plausible explanation for its absence in all crystallographic structures so far. Although an adjustment of the NBDs was observed in all P-gp models, both v1 and v2 models also revealed distortions in the "linker" secondary structure, namely through the formation of an α-helix in its middle coil. Additionally, a kink in the TMHs 6, 10 and 12 was found in both models compromising the DBP bottom and the portals. Together with the low scores obtained in the evaluation servers (**Table S1**, *Annex S2*), the identification of structural alterations were the underlying reasons leading to the rejection of both v1 and v2 models. However, a stable human P-gp model was obtained after 200 ns of simulation time, although with a POPC lipid molecule located in the portal 4/6 (v3a model), in agreement with the computational findings of Tajkhorshid and colleagues.[227] Interestingly, no secondary structure in the middle coil of the linker structure was found in this model.

As it is experimentally known that some lipids play an important role in the stability and function of protein membranes including P-gp,[31,228,229] it was important to assess the impact of the lipid molecule in the protein stability. Therefore, two new system snapshots were retrieved from the v3a model at 200 ns of simulation time as starting points for two additional 200 ns MD runs (in a total of 400 ns of each system), where the POPC molecule was removed from the system (v3b) or moved into the lipid bilayer (v3c) (**Figure S1**, *Annex S2*).

3.3.2 Human P-gp homology model validation

Although lower scores were observed in the initial human P-gp model obtained with MOE, significant structural improvements were achieved in the refined models, after molecular dynamics (MD) simulations (**Table S1**, *Annex S2*). Although the SwissModel

online server still utilizes an algorithm that is not adequate for evaluating membrane proteins,[230,231] all our structures scored above -5.0, a solid indicator on the good quality of the herein developed models.[231]

Another important tool to assess the structural quality of a given structure is through the analysis of the Ramachandran plots of each v3 P-gp models (**Figure S2**, *Annex S2*). Substantial improvements were observed in the refined P-gp models with only 16 (v3a), and 19 (v3b) and 18 (v3c) outliers, respectively versus 34 outliers found in the initial P-gp homology model. The outliers were found mainly in coils, being the only exceptions A348 residue, located at the TMH6 (v3a and v3b), and the F359 (TMH11) and N81 (TMH1) residues in the v3c model.

The RMSD evolution of the P-gp models during MD simulations (**Figure S3**, *Annex S2*) showed that any considerable changes in protein's conformation occurred during the first 200 ns of simulation time. Moreover, a comparison between the refined v3 models showed that all models were stable during the last 200 ns of MD runs with the POPC molecule initially found in the portal 4/6 (v3a model) not affecting the protein stability. However, visual inspection of the three P-gp models revealed differences in the TMHs spatial positions when compared to the starting MD structures (200 ns). While no changes were observed in both v3b and v3c models, the POPC molecule found in the portal 4/6 in the v3a model produced deviations in the spatial arrangement of the TMHs 4 and 6. Moreover, in the v3b model (obtained by removing the POPC molecule from the system's topology), changes in the positions of most TMHs were observed, including helices 4/6 and 10/12 forming both DBP portals. When the POPC was moved into the lipid bilayer in the v3c model, only slight differences in the spatial position of the TM helices 1, 3 and 4 were observed.

All v3 models were also validated using docking studies to assess the number of DBSs within the DBP. The molecules tested were clustered in substrates, probes and modulators, in agreement with experimental[232,233] and *in silico*[41] data. Three DBSs were found inside the DBP in the v3c model, in agreement with the previously refined murine P-gp structure and experimental data, while in both v3a and v3b models the three DBSs were found to be less defined (data not shown).

Altogether, our results showed significant improvements in the quality of human P-gp models after MD simulations indicating that the presence of the "linker" and lipid bilayer are undoubtedly important for the stability of the transporter, in agreement with other studies.

The analysis of the online evaluation servers, Ramachandran plot and RMSD did not reveal significant changes among the refined human P-gp models. However, as differences in the TMHs spatial position and number of DBSs within the DBP were found in both v3a and v3b models, the v3c model was considered the most stable and suitable human P-gp homology model to study the human P-gp variants. This model will be named as wild-type (WT) (**Figure 3.2**). To compare the data gathered from the analysis of P-gp variants with the human WT model, v3c MD run was extended from 400ns to 500ns. This way, after a total of 700 ns of simulation time, a stable human P-gp model with high quality and robustness was obtained.
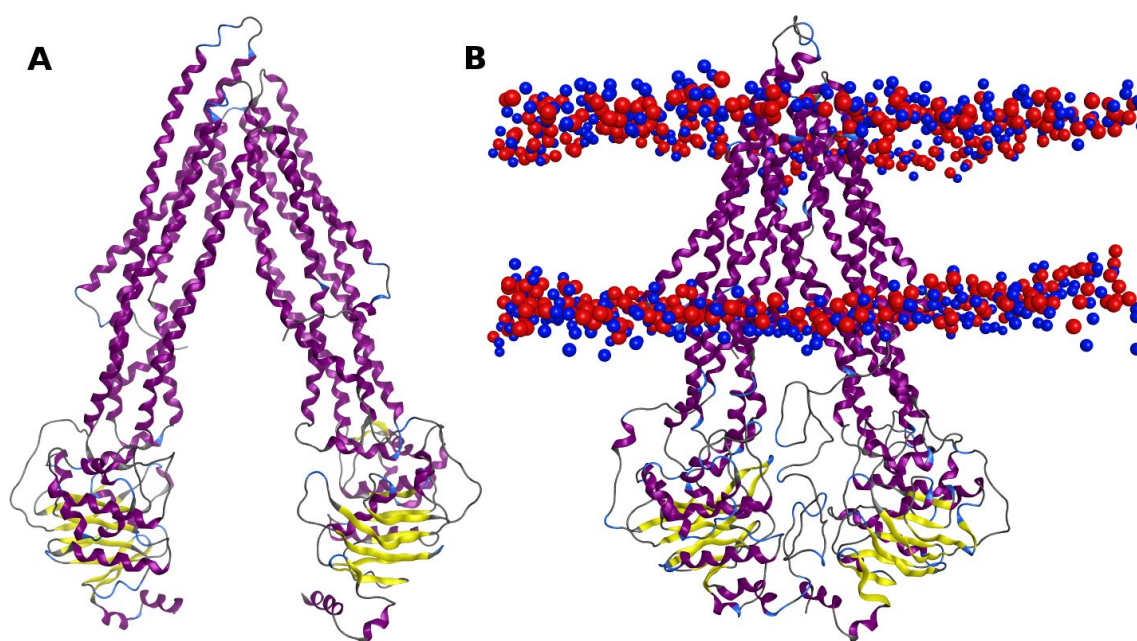

Figure 3.2. (A) Representation of the murine P-gp crystallographic structure (PDB ID: 4Q9H) used as template vs (B) human WT P-gp model (v3c) in the presence of the "linker" and a POPC membrane. The lipid bilayer boundaries are represented through the phosphate (red) and nitrogen (blue) atoms of the lipid headgroups.

During the development of this work, the first human P-gp cryo-EM structure in a nucleotide-free inward-facing conformation was made available (PDB ID: 6QEX).[43] According to the online servers (**Table S1**, *Annex S2*), our human WT P-gp model (v3c) presents better scores than the recently published 6QEX structure, indicating that the quality of our P-gp model is not lower than the quality of the 6QEX structure. Additionally, in the recently available QMEANBrane server,[234] our human WT model also displayed similar local scores (0.746; WT vs 0.783; 6QEX) but more favorable membrane insertion energy (**Figure S4**, *Annex S2*). Structural alignment between the human WT P-gp model and 6QEX structure showed similar spatial positions of the transmembrane and cytoplasmic domains when compared with the initial template (**Figures S5 and S6**, *Annex S2*). Nevertheless, significant differences in the secondary

structure of the TMH4 (S237–A248) and TMH10 (S880-K885) were found between our P-gp model and 6QEX structure. When comparing the RMSD of each TMH (**Table S2**, ***Annex S2***), TMH4 and TMH10 are the helices that present higher RMSD values between the WT model and 6QEX. Interestingly, a chimeric human-mouse P-gp cryo-EM structure (PDB ID: 6FN4)[42] presents similar alterations at the same regions of TMHs 4 and 10 and high RMSD values when compared to the human WT P-gp model.

In order to understand the nature of the TMH4 and TMH10 alterations in the human P-gp cryo-EM structure, four mouse P-gp structures in the *apo* state inward-facing conformation – PDB IDs: 4M1M,[235] 4Q9H[22], 5KPI[23] and 6GDI[38] – and in the presence of ligands at the DBP – PDB IDs: 4M2S[235] and 4Q9L[22] – were used for comparison. Visual inspection showed that both TMH4 and 10 are continuous α-helices in all *apo* and *holo* P-gp crystallographic structures considered, as observed in our human WT P-gp model. Moreover, Kim et al., recently published a human P-gp cryo-EM structure in the ATP-bound, outward-facing conformation (PDB ID: 6C0V),[36] demonstrating that the TMHs 4 and 10 must be continuous α-helices to ensure the correct closure of the cytoplasmic pore when shifting from the inward to outward-facing conformation upon NBD dimerization. Furthermore, structural alignment between each P-gp crystallographic structure and the human WT P-gp model also revealed lower RMSD values for both TM helices 4 and 10, in contrast with what was observed in the 6EQX structure (**Table S2**, ***Annex S2***). As P-gp is a highly flexible protein, we cannot exclude that the recent cryo-EM structure captured some degree of dynamic transitions between several conformations of TM helices 4 and 10, but the above results also imply that v3c is a robust and reliable human WT P-gp model that can be further used in mechanistic and drug discovery studies. The comparison with the experimental data retrieved from the P-gp crystallographic structures indicate that our human WT model, although preserving the secondary structure of the TMDs, retains enough flexibility in the TM helices 4 and 10, thus being comparable to that observed in the human 6QEX structure.

3.3.3 Structural analysis of the human P-gp variants

To gain additional insights on drug specificity and efflux-related signal-transmission mechanism, the structural impact of four P-gp mutations – G185V, G830V, F978A and ΔF335 – experimentally linked with changes in efflux and substrate specificity, were analyzed. The mutations were selected according to their location within the DBP namely, at the substrate binding sites (SBSs) H and R sites (G185V and G830V, respectively) and at the modulator binding site (M-site) (F978A and ΔF335) (**Figure S**7,

Annex S2). Herein, while both H and R sites were initially characterized by their interaction with Hoechst 33342 or Rhodamine-123, respectively, the modulator M site was identified from the localization of the co-crystallized ligands QZ-SSS and QZ-RRR in the first crystallographic structure of murine P-gp. Regarding SBSs, both were later characterized by molecular docking and experimentally confirmed by electron microscopy.

One of the properties that can be altered by mutations in the transmembrane region of P-gp pump is the volume of the internal cavity. An estimation of the probability distribution function of the DBP volumes $P_{(V)}$ (**Figure S8**, *Annex S2*) in the human WT P-gp model and variants showed that all mutations induced a reduction of the DBP volume, more pronounced in the F978A variant. Nevertheless, while G830V and F978A variants still sampled a wide range of DBP volumes during MD simulations (as observed in WT), G185V and ΔF335 variants showed narrower volume distributions suggesting a different structural cohesion of the transmembrane domains.

Thus, to investigate this hypothesis, the TMDs arrangement was analyzed through *g_bundle* and compared to the human WT P-gp model. Accordingly, this tool gives information about the bundle of axes (e.g. TMHs) such as the distance, length, and *z*-shift of the axis mid-points with respect to the average center of all axes as well as the total, lateral and radial tilt with respect to the average axis (*xy* corresponds to the membrane plane and *z* to the longest protein axis). The statistically significant changes in the bundle parameters are summarized in *Annex S2* (**Tables S3-S8** and **Figures S9-S20**).

Overall, the results show distinguishable changes in the TMHs repacking in all P-gp variants, including the helices where the respective mutation is located. Moreover, all mutations showed significant changes in the bundle parameters of the TMHs 4/6 and 10/12 that form the DBP portals and in the "crossing helices" 4/5 and 10/11, that directly link the TMD1 to NBD2 and TMD2 to NBD1. However, significant differences were found in the TMHs repacking between the mutations lying at the SBSs (G185V and G830V) and the mutations located at the M-site (F978A and ΔF335).

Although being located at opposite helices (**Figure S7**, *Annex S2*), the G185V and G830V mutations surprisingly showed similar structural changes in the TMHs rearrangement involving both TMDs. However, G830V mutation located at the TMH9 in the C-terminal P-gp halve seemed to have a stronger impact in the TMHs reorientation when compared to the G185V mutation in TMH3 (**Tables S5-S7**, *Annex S2*). On the other hand, mutations at the M-site showed a completely distinct behavior. While the

F978A mutation (TMH12) seemed to preferentially affect the TMHs rearrangement of the C-terminal P-gp halve where the mutation is located (**Tables S4, S5, S7 and S8**, ***Annex S2***), the deletion of F335 residue (ΔF335, TMH6) apparently did not have a severe impact in the TMHs rearrangement as could be initially expected.

Altogether, the analysis of the helical bundle indicates that i) all P-gp mutations affect the TMHs repacking including the TMHs that form the DBP portals, which may compromise the access of drugs to the internal cavity, ii) all P-gp mutations affect the bundle of the "crossing helices" 4/5 and 10/11, described as important helices in the NBD dimerization process upon ATP-binding[36] and iii) mutations located at the SBSs have a different impact on helical repacking than those located at the M-site. Finally, as a result of the TMHs reorientation, all the selected mutations induce a reduction of the DBP volume in relation to WT (**Figure S6**, ***Annex S2***), but while both glycine and ΔF335 mutations induce a slight decreased in the pocket volume, the partial repacking caused by the F978A mutation may explain the severe reduction of the DBP volume found in this variant.

3.3.4 Interactions between coupling helices and nucleotide-binding domains

Since the selected P-gp variants are experimentally linked with altered drug-resistance profiles and changes in either the basal or drug-stimulated ATPase activity, the residue interactions at the ICH-NBD interfaces thought to be involved in signal-transmission and efflux-related conformational changes were evaluated and compared to the human WT P-gp model. The total number of contacts was estimated using the *g_hbond/gmx hbond* module and is depicted in ***Annex S2*** (**Figures S21 and S22**).

For both glycine variants, only at ICH3-NBD2 interface a significant decrease in the total number of contacts was observed. Oppositely, both M-site variants displayed a distinct behavior. While the F978A variant showed a decrease in the total number of contacts at the ICH3-NBD2 interface (similar to the glycine mutants), ΔF335 showed significant changes in the total number of contacts in three of the four interfaces, namely ICH2-NBD2, ICH3-NBD2 and ICH4-NBD1. As all mutants seemed to induce changes in the ICH-NBD total number of contacts, we further identified which residue pairs were involved. Mean residue-residue contact frequencies ≥ 0.5 and variations above 10% were considered significant and are summarized in ***Annex S2*** (**Table S10**). For clarity purposes, each ICH-NBD will be analyzed separately in the following section.

Concerning the ICH1-NBD1 interface, all P-gp variants showed an increase of the mean contact frequencies between I160 (ICH1) and L443 (Walker A, NBD1), both located in regions identified to be involved in ATP-binding.[103,117,217] Oppositely, only in F978A and ΔF335 a decrease in contact frequencies involving D164 (ICH1) and R404 (NBD1) was observed. Quite interestingly, D164 was also reported to be part of an extensive interaction surface between the TMDs and NBDs, with D164C mutants additionally revealing lower cell surface expression.[110] Regarding the ICH1-NBD1 hydrogen bond network, a significant decrease in the hydrogen-bond lifetimes (life column, Table S10, *Annex S2*) was also observed for all mutants.

The other interface at this nucleotide-binding domain is the ICH4-NBD1. Herein, the most affected residue pairs were R905–S434/Q438/Q441, with a greater decrease of contact frequencies in both glycine mutants; V908/R467 (increased in all mutants); and S905/Y401 and S909–Q4441/R467/V472, mainly increased in the F978A and ΔF335 mutants. Again, mutational studies implied both S905 and S909 in the activation and ATPase stimulation when in the presence of drugs and/or lipids.[217,218] Interestingly, and specifically concerning the ΔF335, new contacts between V907/F480, L910/R547, E913/R464 and Q914/R464 explain the increase in the total number of contacts reported above. Regarding the hydrogen bond network between the ICH4 and NBD1 residues, a significant decrease in the average HB number, lifetime and energy of HB formation was observed only in both glycine mutants (**Table S10**, *Annex S2*).

At the opposite NBD, all mutations seem to induce a general decrease in the overall contacts frequencies, with the ICH2 residues I265 and F267, involved in P-gp maturation and activity[104,115] and the NBD2 residue R1110, being the most affected ones. However, this decrease is partially mitigated by a reinforcement of the HB network in all mutations and, specifically for the ΔF335, through new contacts between F267 (ICH2) and G1134/R1188/A1189/R1192 (NBD2) or G269 (ICH2) and N1136 (NBD2) (**Table S10**, *Annex S2*). Finally, and regarding residue F1086, identified through *in vitro* studies as important for coupling of ATP-binding to conformational changes in the TMDs,[97] almost all contacts frequencies decrease except when paired with R262 (G185V and F978A) or I265 (ΔF335).

Finally, for the ICH3-NBD2 interface, all variants showed, in general, a negative variation in the contact frequencies, most particularly between V801/S802 (ICH3) and Y1087 (NBD2), thought to play an important role in P-gp activity and assembly.[97] In the same way, all mutations induced a decrease of the contact frequencies between the

D805 (ICH3), an important residue thought to be involved in the TMD-NBD communication,[50] and Y1044, located at the A-loop of NBD2.

Altogether, these results indicate that although mutations in the TMDs of the human P-gp affect directly the transmembrane region with changes in the DBP volume and DBSs features, they also induce changes in the residue interactions at the ICH-NBD signal-transmission interfaces involved in the TMD-NBD communication. However, a comparison among P-gp variants showed that both mutations at the SBSs (G185V and G830V) induce identical changes in the total number of contacts between the ICH and NBD residues, while the impact of the M-site mutations (F978A and ΔF335) seems to be dependent of the TMH where the mutation is located, as observed in the helical bundle. Finally, the analysis of the mean contact frequencies in the human WT P-gp model and its variants, identified some residue pairs potentially involved in TMD-NBD communication, indicating an interaction network between the ICHs and NBDs residues, including those that directly interact with ATP, in agreement with several other studies.[45,50,97,98,217]

3.4 DOCKING RESULTS

3.4.1 Identification of Drug-Binding Sites in the human P-gp model

Based on the docking poses of known P-gp substrates, probes and modulators, Ferreira and co-workers[41] defined the location of the DBSs within the DBP, firstly identified by Shapiro et al.,[39,40,61] and Aller et al.,[16] using a refined murine P-gp structure. To acquire a deeper knowledge on P-gp substrate binding, docking studies targeting the DBP of the human P-g WT model were undertaken. As twenty docking poses were generated per molecule, to simplify the results only the top-ranked binding energies (ΔG_{dock}) at each DBS will be compared to the data obtained from previous studies.

Overall, most of the molecules tested bind at the three DBSs found within the DBP of the human WT model (**Table S9**, *Annex S2*), as observed in the murine P-gp structure. Nevertheless, higher number of molecules interacted with the H-site in respect to the refined murine P-gp. However, considering the standard error reported for VINA (2.85 kcal·mol^{-1})[199] no conclusions could be made regarding the affinity of these compounds since they have similar ΔG_{dock} among the DBSs within the DBP of the WT model and similar ΔG_{dock} than those reported in the refined murine P-gp structure.

Altogether, these results show that most P-gp substrates, probes and modulators interacted with the three DBSs (**Figure 3.3**) suggesting that the DBP of the human WT P-gp model discriminate ligands differently than the internal cavity of the refined murine P-gp structure.
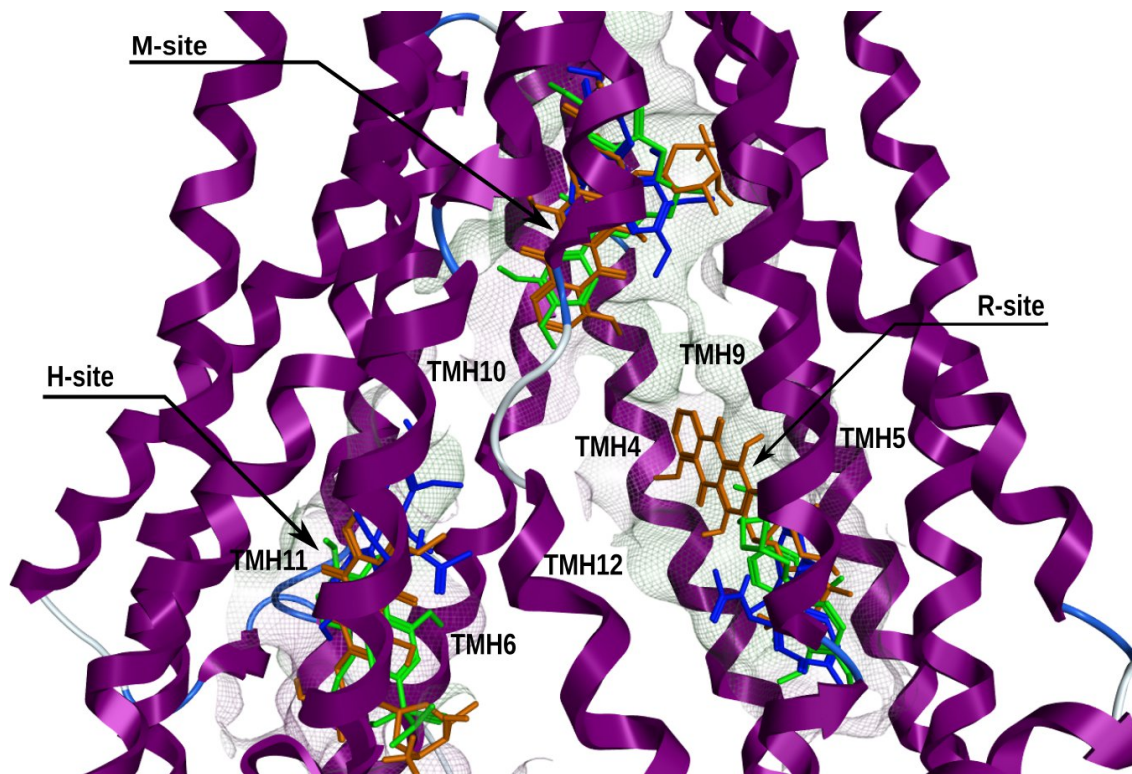

Figure 3.3. Representation of the DBSs found within the DBP of the human WT P-gp model. The three DBSs are defined by the best-ranked docking poses at each binding cavity of well-known P-gp substrates and modulators e.g. verapamil (green), doxorubicin (dark orange) and colchicine (blue).

3.4.2 Identification of Drug-Binding Sites in the human P-gp variants

To assess if the TMHs repacking found in the P-gp variants affect drug binding, docking studies using known P-gp substrates and probes were performed as previously described and compared to the WT model. The molecules top-ranked binding energies (ΔG_{dock}) at each DBS are depicted in *Annex S2* (**Table S9**).

The total repacking of the transmembrane α-helices (TMHs) observed in both substrate binding site (SBS) variants (G185V and G830V) led to changes on substrate binding mostly affecting the R and H sites. Although more dramatic in the G830V variant, most of the compounds tested that showed to interact with the three DBSs in the WT model, did not bound at the H-site or did not interact with both SBSs. Additionally, and in contrast with what was observed for the G185V variant, some of the substrates that docked at the three DBSs in WT, did not interact with both M and H sites in the G830V variant. Although, no clear conclusions could be drawn about the possible changes in

the molecule's affinity upon the TMHs rearrangement observed in these variants, nonetheless it becomes clear that the total helical repacking induced by the glycine mutations had a dramatic influence on the availability of each drug-binding site (DBS) to the evaluated set of molecules.

On the other hand, although the modulator binding site (M-site) variants (F978A and ΔF335) also showed changes on substrate binding mostly affecting the H or both SBSs, these mutations did not have a severe impact on drug binding as observed in the glycine variants. Nevertheless, due to the standard error reported for VINA and although no clear correlation could be obtained about the possible changes in drug affinity upon the TMHs repacking observed in these variants, it is quite interesting to note that specific mutations at the M-site have such a large influence on the SBSs rather than what would be initially expected, at the M-site.

Altogether, the analysis of the docking results demonstrates that all P-gp variants present changes on drug binding as a result of the total or partial TMHs repacking induced by these mutations, mostly affecting the R and H sites.

3.4.3 Characterization of Drug-Binding Sites

To better understand the changes on substrate binding observed in the P-gp variants, the pocket volume, residues distribution and mean polarity of each DBS were assessed for the human WT model and variants using the EPOS[BP] software. The results were further compared to the human WT model and the refined murine P-gp structure.

Overall, both M and R sites showed similar pocket volumes among the human WT model (M-site, 1284 Å^3 and R-site, 1902 Å^3) and the refined murine P-gp structure (M-site, 1300 Å^3 and R-site, 1900 Å^3). In contrast, the H-site is considerable smaller (1232 Å^3) in human WT than in the refined murine structure (2200 Å^3). Additionally, and much like the refined murine structure, the M-site of the human WT showed to be the most hydrophobic DBS with higher number of aromatic residues in contrast with the H and R sites, that presented a higher percentage of polar residues. When comparing the site's residues distribution between species, no significant changes were found in the M and R sites, although the human WT H-site is more hydrophilic than the previously reported for the murine P-gp structure. However, despite the differences in volume and residues distribution found in the human WT H-site, this SBS was capable of binding more compounds than those observed in the refined murine P-gp structure.

Regarding P-gp variants, the total repacking induced by both glycine mutations led to changes in the residues' side-chain facing the H-site, with a remarkable increase in the

content of hydrophobic (G185V, +53%; G830V, +75%) and aromatic side-chains (G830V, +63%) in respect to WT. Nevertheless, as a result of the stronger impact of G830V mutation on TMHs repacking, dramatic changes in the DBSs volume were also found. While a reduction around 50% was observed in the volume of R-site, the volumes of both M and H sites increased about 37% and 58%, respectively, in respect to WT. Therefore, for these mutations the data suggest that a slight increase in the hydrophilicity of the substrates may hamper its binding to the more hydrophobic sites.

Although the F978A mutation preferentially affected the TMHs repacking of the C-terminal P-gp halve, alterations in the residues' side-chain facing the H-site were also observed with significant variations in the percentage of hydrophobic (+89%) and aromatic side-chains (+26%). When compared to the other P-gp variants, the ΔF335 variant seemed to show minor variations in the volume and residues' side-chain facing all DBSs, although the H-site remained the most affected showing a decrease of its volume around 20% and an increase in the percentage of hydrophobic (+12%) and aromatic (+20%) side-chains. Similar to the glycine variants, lowering the hydrophobicity and reducing the number or aromatic rings is expected to decrease the binding to the mutated binding sites.

Altogether, these results indicate that the changes on drug binding found in all P-gp variants are related with alterations in the DBSs properties upon the rearrangement of the TMHs that delimited each DBSs. The visual inspection of each DBS showed that all mutations also induced changes in the DBSs structure, being more severe in the H-site. While in the WT model, the pocket entrance of the H-site is completely opened to the DBP with the formation of a cleft capable of accommodating substrates, all mutations induced distortions that affected either the pocket entrance and/or the cleft, compromising the binding of substrates (**Figure S23**, *Annex S2*).

In sum, the characterization of the DBSs indicates that the total or partial TMHs repacking in response to mutations in the transmembrane domains (TMDs) seems to affect the shape, volume and residues distribution of the three DBSs within the drug-binding pocket (DBP). Interestingly, although changes on drug binding were also observed in the M and R sites, the H-site was the most affected by all mutations becoming unable to interact with most of the P-gp substrates and probes tested.

3.5 FINAL DISCUSSION

Modulating drug efflux by P-gp pump is one of the promising strategies to reverse MDR in cancer cells. Nevertheless, the lack of information about the molecular basis underlying drug specificity and efflux-related signal-transmission mechanism between the TMD-NBD domains impairs the development of more potent and selective compounds able to overcome MDR. Therefore, to provide additional insights on drug specificity and efflux mechanism, the impact of four P-gp mutations (G185V, G830V, F978A and ΔF335), experimentally linked with changes in efflux, basal and drug-stimulated ATPase activity, were comprehensively assessed.

In this work, a human P-gp homology model was developed based on a murine P-gp crystallographic structure in the *apo* state inward-facing conformation (PDB ID: 4Q9H). The initial homology model obtained was refined through MD simulations in the presence of a "linker" retrieved from a previously equilibrated murine P-gp structure, inserted in a POPC membrane also used in previous studies and thereafter validated. By comparison with a recently published cryo-EM P-gp structure,[43] our homology model was found to maintain an adequate reliability and robustness crucial for the herein proposed analysis.

The final human P-gp model was then used to thoroughly characterize the effect of the P-gp mutations mentioned above in the structure of the transporter. The mutations are located at the transmembrane region surrounding the M (F978A and ΔF335), H (G185V) and R (G830V) sites. Taken together, these mutations seem to induce TMHs repacking affecting the DBP portals by changing the "crossing helices" 4/5 and 10/11 important to NBD dimerization also compromising the access of drugs to the internal cavity, and reducing the DBP volume. Additionally, as a result of the TMHs repacking significant changes in the volume, shape and polarity of the DBSs within the DBP were also observed in all P-gp variants, mostly affecting the binding of substrates at the H and R sites. It is noteworthy that although F978A and ΔF335 mutations lie at the M-site, they affect in a similar manner the SBSs properties and drug binding as observed in both glycine variants, suggesting a communication pathway between the M-site and the SBSs through the functional TMH6 and TMH12, in agreement with some experimental studies.[62,221,236] Therefore, we hypothesize that changes in the structure and polarity of the DBSs induced by the TMHs repacking i) provide a possible explanation for P-gp promiscuity, reported in the literature and ii) suggest that small variations in both substrates and modulators may be enough to impair substrate binding and/or to enhance the modulators' activities. Interestingly, several examples on

the latter are already described in literature as suitable approaches in enabling molecules to evade efflux[237] or even to switch the activity of known substrates into high-affinity compounds able to inhibit P-gp's ATPase activity.[238]

Even though the mutations described above directly affected the transmembrane region, they also induced changes in the total number of contacts at the ICH-NBD interfaces, suggesting that the TMHs rearrangement is involved in the TMD-NBD communication, in agreement with several experimental. Furthermore, all P-gp variants showed significant changes in the mean contact frequencies of specific residue pairs, mainly located at the ICH2/ICH3-NBD2, an important transmission interface to couple drug binding to ATPase activity, being also critical for P-gp folding. Thus, another interesting approach to the modification of substrates/inhibitors is the development of allosteric modulators able to specifically interact at the ICH-NBD interfaces and impair the signal-transmission between the TMD and NBDs. Currently, only two scaffolds are currently known to interact in such domains, namely dihydropyridines[118] (TMD-NBD1) and flavonoids (NBD2),[94] and additional efforts must be taken in the future to explore this hypothesis.

Nevertheless, although all P-gp mutations induced similar structural effects on the transporters' architecture, it seems clear that mutations at the M-site (F978A and ΔF335) have a completely different impact on P-gp structure than the mutations located at the SBSs (G185V and G830V). Both glycine mutations induced a total TMHs repacking affecting drug binding at the SBSs. Additionally, the glycine mutations do not have a significant impact in the total number of contacts as expected, suggesting that G185 (TMH3) and G830V (TMH9) residues have equivalent roles in P-gp function and possibly more involved in drug binding.

On the other hand, the effects of the M-site mutations show to be dependent of the affected TMH. The F978A mutation (TMH12) preferentially induced a partial TMH repacking, affecting drug binding at the SBSs while ΔF335 mutation (TMH6) dramatically change the residue interactions at the ICH-NBD interfaces. These results support the hypothesis that while the F978 residue is likely involved in drug binding as reported in some experimental studies, F335 residue is involved in TMD-NBD communication.

3.6 CONCLUSION

Overall, this work provides clear evidence that mutations at these specific TMHs i) are responsible for inducing a repacking of the TMHs, changing the DBP volume and drug binding sites, mostly affecting drug binding at the SBSs and ii) also impact the ICH-NBD signal-transmission interfaces, suggesting that a perturbation in the TMDs (e.g. mutations or binding of substrates) induce a TMHs rearrangement that are transmitted to the NBDs through changes in the residue interactions between the ICHs and the respective NBD, in agreement with experimental studies.

Nevertheless, as protein conformational changes may occur at least on the timescale of microseconds[239] and the computational power available is limited, it should be clear that our findings are based in theoretical models aiming to predict the possible structural impact of single-point mutations on P-gp architecture/organization. Additionally, as these P-gp mutations are experimentally related with changes in drug-stimulated ATPase activity upon binding of specific substrates, more studies are needed to assess their direct impact on drug and ATP-binding. Therefore, further studies including molecules with altered efflux properties are undergoing, in an attempt to clarify these issues.

CHAPTER 4 – PAPER II

Long-range communication between transmembrane- and nucleotide-binding domains does not depend on drug binding to mutant P-glycoprotein.

The modulation of drug efflux by P-glycoprotein (P-gp, ABCB1) represents one of the most promising approaches to overcome multidrug resistance (MDR) in cancer cells, however the mechanisms of drug specificity and signal-transmission are still poorly understood, hampering the development of more selective and efficient P-gp modulators. In this study, the impact of four P-gp mutations (G185V, G830V, F978A and ΔF335) on drug-binding and efflux-related signal-transmission mechanism was comprehensively evaluated in the presence of ligands within the drug-binding pocket (DBP), which are experimentally related with changes in their drug efflux profiles. The severe repacking of the transmembrane helices (TMH), induced by mutations and exacerbated by the presence of ligands, indicates that P-gp is sensitive to perturbations in the transmembrane region. Alterations on drug-binding were also observed as a consequence of the TMH repacking, but were not always correlated with alterations on ligands binding mode and/or binding affinity. Finally, and although all P-gp variants *holo* systems showed considerable changes in the intracellular coupling helices/nucleotide-binding domain (ICH-NBD) interactions, they seem to be primarily induced by the mutation itself rather than by the presence of ligands within the DBP. The data further suggest that the changes in drug efflux experimentally reported are mostly related with changes on drug specificity rather than effects on signal-transmission mechanism. We also hypothesize that an increase in the drug-binding affinity may also be related with the decreased drug efflux, while minor changes in binding affinities are possibly related with the increased drug efflux observed in transfected cells.

4.1 INTRODUCTION

Over-expression of membrane efflux pumps is intimately related to multidrug resistance (MDR) phenomenon in cancer cells. Among them, P-glycoprotein (P-gp, ABCB1) is the most studied so far, and it is capable to extrude a wide range of neutral and charged hydrophobic compounds), through an ATP-dependent mechanism.

P-glycoprotein architecture comprises two transmembrane domains (TMDs), each one formed by six transmembrane α-helices (TMHs) packed in a pseudo 2-fold symmetry, and two cytoplasmic nucleotide-binding domains (NBDs), with both P-gp halves connected through a small peptide sequence (the "linker"; residues 627-688). The transmembrane and cytoplasmic domains of the P-gp efflux pump are physically linked to the respective NBD by coils bridging TMH6/NBD1 and TMH12/NBD2, and by short intracellular coupling helices (ICHs), located between TMHs 2/3 (ICH1-NBD1), 4/5 (ICH2-NBD2), 8/9 (ICH3-NBD2) and 10/11 (ICH4-NBD1), the latter involved in the TMD-NBD communication through non-covalent interactions (see *Paper I, Results and Discussion*).[240,241] Additionally, some studies indicated that ICHs may also play important roles in P-gp folding and maturation. Oppositely, the drug-binding pocket (DBP) is located within the TMHs of both N- and C-terminal P-gp halves, and contains, at least, three distinct drug-binding sites (DBS): the modulator site (M-site) located at the top of DBP, and the substrate-binding sites (SBSs) H and R, named after Hoechst33342 and Rhodamine-123 respectively, located in the opposite side next to the inner leaflet of the lipid bilayer.

Although decades of research focus on the molecular basis of drug promiscuity and efflux, the mechanisms of drug specificity and efflux-related signal-transmission are still unclear, thus hampering the development of novel, more potent and selective P-gp modulators able to overcome MDR. Herein, one interesting approach is the study of P-gp variants experimentally related to altered drug-resistance phenotypes and/or changes in the basal and drug-stimulated ATPase activity. Our previous work (see *Paper I, Results and Discussion*) reported the impact of four hP-gp variants involved in MDR – G185V, G830V, F978A and ΔF335 – on the P-gp structure in the absence of ligands within the DBP (*apo* systems). In this paper, in order to understand if the P-gp conformational changes induced by these P-gp mutations affect drug-binding and efflux-related signal-transmission mechanism, their structural impact was assessed in the presence of molecules experimentally described as having alterations in the efflux profiles and bound to the reported DBSs (*holo* systems).

Similar to our previous work, as the glycine mutations G185V (TMH3) and G830V (TMH9) are located at the SBSs H and R, respectively, both variants will be named SBS-variants. On the other hand, the F978A (TMH12) and ΔF335 (TMH6) mutations, both lying at the M-site, will be referred as M-site variants. Following, the ligands were chosen considering: i) the location of the mutation at the TMDs, ii) the experimental and/or computational evidences for its putative DBS, and iii) opposite changes in the drug efflux profiles by the same P-gp mutation (see *Paper I, Introduction*). Thus, colchicine (COL) and vinblastine (VIN), both predicted to bind at the H-site, were the substrates chosen for the G185V (H-site). Likewise, doxorubicin (DOX) and actinomycin D (ACT), both predicted to interact with the R-site, were the substrates selected for the G830V (R-site). On the other hand, since the mutations at the M-site (F978A and ΔF335) seem to strongly affect the SBSs properties and drug-binding (*see Paper I, Introduction*), the structural effects of these mutations were analyzed in the presence of ligands in all three sites. Therefore, the P-gp modulator cyclosporine A (CYC, M-site), and the substrates ACT (R-site) and VIN (H-site) were the molecules selected for the F978A mutation. However, to the best of our knowledge, there is no information about a P-gp substrate that binds at the H-site and has altered drug-resistance profile in transfected cells harboring the ΔF335 mutation. That said, DOX (R-site) and two P-gp modulators, CYC and valspodar (VLS), both predicted to bind at the M-site, were the ligands chosen for this P-gp variant.

4.2 MATERIAL AND METHODS

4.2.1 Initial structures

From our previous study, five MD *apo* systems were used as template to build the MD *holo* systems, the refined and validated hP-gp wild-type (WT) and four hP-gp variants models (G185V, G830V, F978A and ΔF335). In each one the transporter was embedded in a lipid bilayer with 469 molecules of 1-palmitoyl-2-oleoylphosphatidylcholine (POPC), using a lipid parameterization by Poger et al.,[146,147] as previously described (*see Paper I, Material and Methods*).

From our docking studies previously published, the top-ranked docking poses for the chosen ligands were selected and used to assemble the WT and variants MD *holo* systems. The molecules were parameterized in the PRODRG online server and manually curated, according to the GROMOS96 54a7 force field, using AM1-BCC partial charges calculated with Antechamber v1.27 . The protein-ligand-membrane systems were then solvated with 63,810 ± 31 single-point charge (SPC) water molecules and

neutralized with 11 chlorine ions, for a total of 16 MD *holo* systems. These MD systems were the starting point for several short MD runs.

4.2.2 Free-energy calculations and analysis

After a short energy minimization run to minimize clashes between the ligand and the protein, a total of five replicates of 20 ns MD runs were performed for each system in a total of 1.5 μs of simulation time (*n* = 16, 0.1 μs each system). For each replicate initial velocities were assigned from a Maxwell-Boltzmann distribution at 303 K. To enhance sampling, for the last two replicates distinct snapshots were retrieved from the first MD run (at 18 ns and 19 ns, respectively). Only the last 10 ns of simulation time were considered as production runs and used for further analysis. Relative free-energies of binding (ΔG_{MD}) were calculated using the *g_mmpbsa* tool, with an implicit membrane correction for polar solvation energies. As in a previous paper, the tools *gmx bundle, gmx hbond* and *g_contacts* (only available in GROMACS v4.6.7) were used to further evaluate changes in the transmembrane (TM) helical bundle, hydrogen bond (HB) networks, and protein interactions, respectively.

4.2.3 MD simulation parameters

All MD simulations were done with GROMACS v2016.x package. All *NVT* equilibration runs were performed at 303 K using the Velocity-rescale (V-rescale) thermostat. The Nosé-Hoover thermostat and the Parrinello-Rahman barostat for temperature (303 K) and pressure (1 bar) control, were applied in all *NpT* runs. Due to the presence of a membrane, pressure equilibration was achieved through a semi-isotropic pressure coupling, with the systems' compressibility set to 4.5×10^{-5} bar^{-1}. All bond lengths were constrained using the LINCS or SETTLE (for water molecules) algorithms. PBC conditions were applied to all dimensions. The Particle Mesh Ewald (PME) with cubic interpolation was employed with a cut-off radius of 12 Å for both electrostatic and van der Waals interactions, and an FFT grid spacing of 0.16 for long range electrostatics. Group-based and Verlet cut-off schemes were applied for the calculation of non-bonded interactions on CPU or GPU, respectively.

4.3 RESULTS AND DISCUSSION

The analysis that will be presented and discussed in the following subsections relates to the behavior of the human P-gp WT and variants studied, containing different allocrites inside the drug-binding sites.

4.3.1 Structural analysis of transmembrane domains of the hP-gp variants

To give additional insights on drug specificity and efflux-related signal-transmission mechanism, the structural impact of four selected P-gp mutations (G185V, G830V, F978A and ΔF335), experimentally related to changes in drug efflux and substrate specificity, were analyzed in the presence of molecules bound to each site inside the DBP (*holo* systems). The molecules chosen are described in literature as having altered efflux profiles in these P-gp variants —G185V (COL, VIN), G830V (ACT, DOX), F978A (ACT, VIN, CYC), and ΔF335 (CYC, VLS, DOX).

Mutations (*see Paper I, Results and Discussion*) and/or the presence of ligands within the DBP may directly affect the TMDs architecture. Thus, the rearrangement of the transmembrane helices (TMHs) was assessed using the *gmx bundle* for the P-gp variants *holo* systems and compared to the WT (*holo* system). All the significant changes in the bundle parameters are summarized in the ***Annex S3*** (**Tables S1-S24 and Figures S1-S48**). To organize and simplify the results described below, the studied P-gp systems will be encoded and referred as "variant-DBS1_ligand-DBS2", where DBS1 is DBS where the mutation is located and DBS2 the DBS where the molecule is reported to bind.

Overall, the bundle parameters calculated for all P-gp variants were found to have distinct shifts from the WT, when in the presence of the respective ligands. Most remarkably, the changes were larger when the helices that form the DBP portals (4/6 and 10/12) and the "crossing helices" (4/5 and 10/11), directly connecting the TMD1 to NBD2 and TMD2 to NBD1 respectively, are involved. However, distinguishable structural effects in the TMH rearrangement were still observed between P-gp variants. The SBS-variants G185V and G830V, in the presence of the respective ligands, revealed deeper changes in the TMH repacking affecting both P-gp halves. Firstly, COL in the G185V-H_COL-H system induced a stronger impact in the TMDs architecture than VIN in the G185V-H_VIN-H system. Similarly, the F978A variant showed a striking TMH repacking, more severe in the F978A-M_CYC-M and F978A-M_ACT-R systems. On the other hand, the total repacking observed in the ΔF335 variant seemed to be similar in all *holo* ΔF335 systems.

The structural impact of a specific ligand in two different P-gp variants was determined (**Table 4.1**). Herein, a stronger effect in the TMH repacking was observed in the G185V-H_VIN-H and G830V-R_DOX-R systems than those observed for the M-site variants in the F978A-M_VIN-H and ΔF335-M_DOX-R systems, respectively. On the other hand,

ACT induced a similar impact in the TMH rearrangement in both G830V-R_ACT-R and F978A-M_ACT-R systems.

Table 4.1 Structural impact of mutations in the TMH repacking for the *holo* hP-gp variants.[a]

MUTATION-DBS	MOLECULE-DBS	REPACKING EFFECTS
G185V-H	COL-H	+++
	VIN-H	++
G830V-R	DOX-R	+++
	ACT-R	+++
F978A-M	CYC-M	+++
	ACT-R	+++
	VIN-H	+
ΔF335-M	CYC-M	+
	VLS-M	+
	DOX-R	+

[a] Positive variation – lowest impact (+); highest impact (+++). The P-gp variants *holo* systems were compared both with the WT and between them.

Moreover, CYC seemed to have larger shifts in the bundle parameters in the F978A-M_CYC-M system than for the ΔF335-M_CYC-M system. Finally, when compared to the P-gp variants *apo* systems (*see Paper I, Results and Discussion*), more severe changes in the TMDs architecture were observed for all *holo* P-gp variants. Nonetheless, the presence of ligands in the M-site variants seemed to have a larger effect in the TMH repacking than those reported for the respective *apo* systems.

Altogether, the analysis of the helical bundle indicates that the studied P-gp mutations, when in the presence of ligands within the DBP, i) severely affect the native spatial position of the helices that form the DBP portals, which may alter the access of drugs to the internal cavity or their release to the membrane,[242] and ii) induces changes in the helical bundle concerning the "crossing helices", crucial for the NBD dimerization process upon ATP-binding. Additionally, the presence of molecules seems to induce a more severe impact in the TMHs rearrangement than those observed in the respective *apo* systems,[240] indicating a TMDs adjustment and response to the presence of ligands.

4.3.2 Characterization of protein-ligand interactions

Since all the above mutations changed the TMDs architecture in variable degrees, their impact on the ligand-binding affinity was further evaluated. Relative free-energies of

binding (ΔG_{bind}) were estimated using the *g_mmpbsa* tool and compared to the WT *holo* systems (**Figures S49-S52**, *Annex S3*).

Although a severe TMH repacking was observed in all P-gp variants, the binding affinity was affected differently (**Table 4.2**).

Table 4.2 Structural impact of mutations in the ligand-binding affinity for the P-gp variants *holo* systems.

Mutation-DBS	Molecule-DBS	Binding affinity impact [a]
G185V-H	COL-H	NS
	VIN-H	↑↑↑
G830V-R	DOX-R	NS
	ACT-R	NS
F978A-M	CYC-M	↑↑
	ACT-R	↑
	VIN-H	↑↑↑
ΔF335-M	CYC-M	↑↑
	VLS-M	↑↑
	DOX-R	NS

[a] Positive variation – lowest impact (↑); highest impact (↑↑↑); not significant (NS). The P-gp variants *holo* systems were compared both with the WT and between them.

While no significant changes were found in the G185V-H_COL-H system, a substantial increase in binding affinity was observed in both G185V-H_VIN-H and F978A-M_VIN-H systems. Oppositely, no considerable effects were observed in the calculated ΔG_{MD} values for the G830V-R_DOX-R and ΔF335-M_DOX-R systems. No significant changes in binding affinities were also found for the G830V-R_ACT-R system. On the other hand, although observing a slight increase in the ACT-binding affinity in the F978A-M_ACT-R system, no clear conclusions could be made about the possible changes in the molecule's affinity. Finally, an increase on the ligand-binding affinities were observed in the F978A-M_CYC_M, ΔF335-M_CYC-M and ΔF335-M_VLS-M systems.

To understand how the severe TMH rearrangement observed in all P-gp variants affected ligand-binding affinity differently, specific protein-ligand interactions were additionally assessed using the *g_contacts* (from GROMACS v4.6.7) and *gmx hbond* tools, and again compared to the WT protein *holo* systems. Mean contact frequencies ≥ 0.5 and variations above 10% were considered as significant and are depicted in *Annex S3* (**Tables S25-S27**).

According to our previous studies, the TMH repacking induced by these P-gp mutations led to severe changes in both structure and residues distribution at the DBSs, thus having a strong impact on size and polarity (*see Paper I, Results and Discussion*). Herein, and concerning the G185V-H_COL-H system, colchicine, initially located at the bottom of the H-site, shifted from its original binding position in WT, losing interactions with TMHs 10/12 while establishing new ones with TMHs 6/11 (**Figure 4.1**).
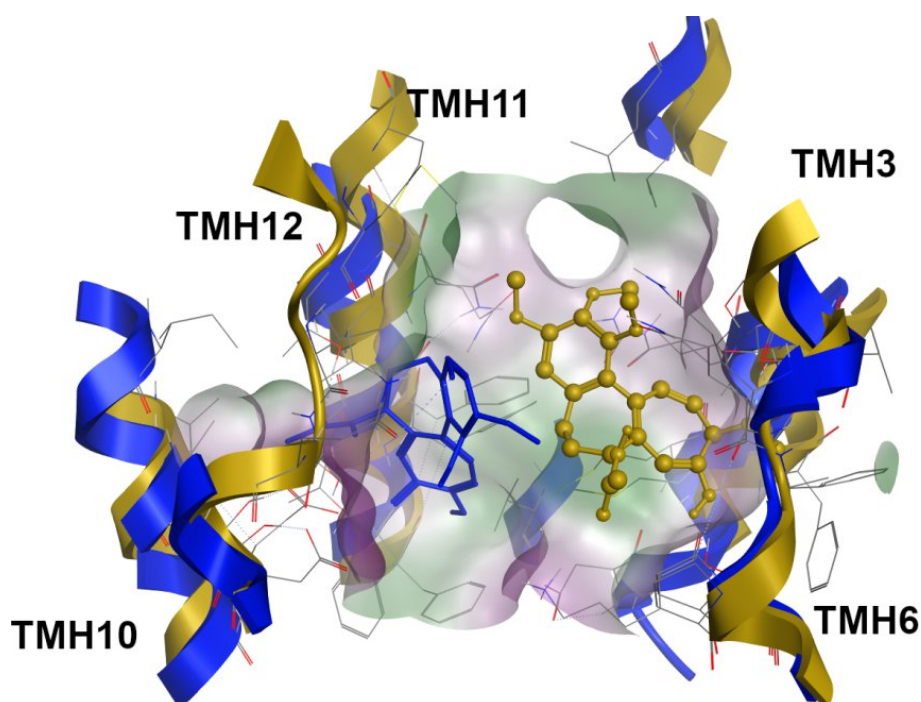
Figure 4.1 Superimposition of the colchicine-bound systems at the H-site for WT (blue, licorice) and G185V variant (dark yellow, ball-and-stick).

However, and according to the ratio between the number of significant protein-ligand contacts and the average of the total number of contacts (hereby named contact efficiency ratio – CE, **Table S28**, *Annex S3*, similar values were observed in both WT-H_COL-H and G185V-H_COL-H systems, thus indicating that although COL established new contact points with residues from other TMHs, no significant impact in binding affinity was observed for this ligand. In the same way, a similar CE ratio was also observed between the WT-R_DOX-R and the variant *holo* systems G830V-R_DOX-R and ΔF335-M_DOX-R.

Regarding doxorubicin, a common binding mode between the WT and both G830V and ΔF335 variants was observed, in which the tetracyclic core is found in close contact with hydrophobic surface patches – by interacting with residues from TMHs 7 and 9 – in the respective pockets (**Figure 4.2**).
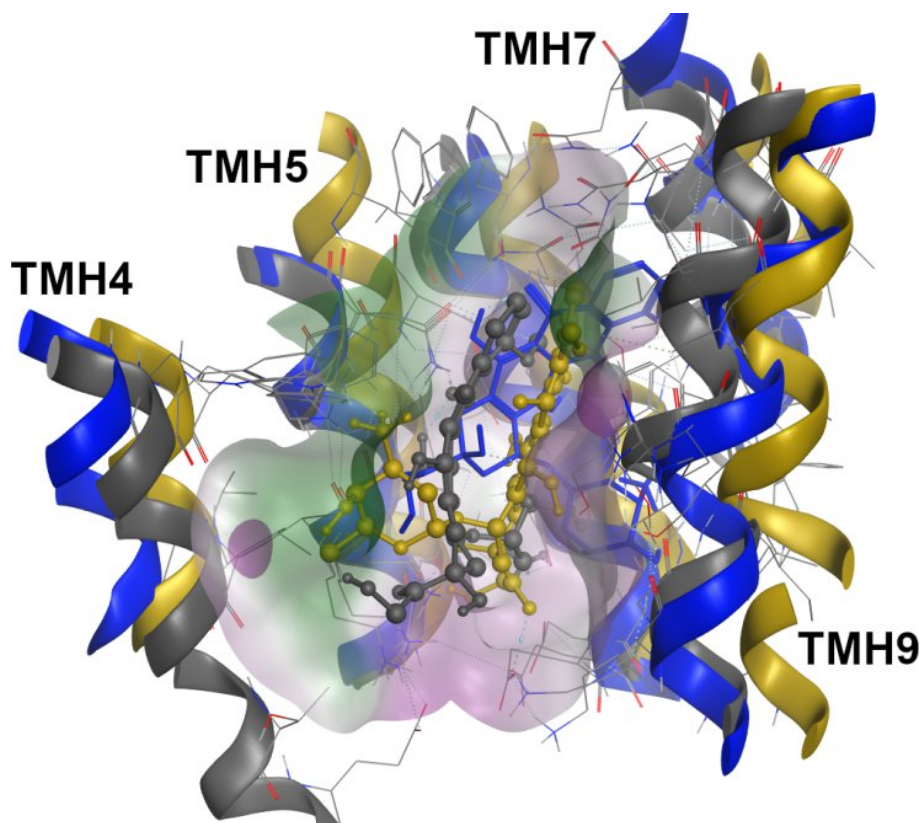
Figure 4.2 Superimposition of the doxorubicin-bound systems at the R-site for the WT (blue, licorice), G830V (dark yellow, ball-and-stick) and ΔF335 (gray, ball-and-stick) variants.

These results indicate that the binding mode for DOX is not substantially affected by these mutations, and no significant alterations in the DOX binding affinity were observed for these variants. Oppositely, a great increase in the vinblastine contact efficiency ratio was observed in both G185V-H_VIN-H and F978A-M_VIN-H systems. The visual inspection showed that VIN was deeply buried in the H-site in both P-gp variants, which may account for stronger interactions and, concomitantly, to the higher binding affinity observed for this molecule. Quite interestingly, a complete distinct binding mode was observed for ACT in the G830V-R_ACT-R system. The visual inspection revealed severe alterations in the ACT-binding mode, with a large region of the molecule protruding from the R-site towards the middle of the DBP and, thus exposed to the surrounding environment (**Figure 4.3**).
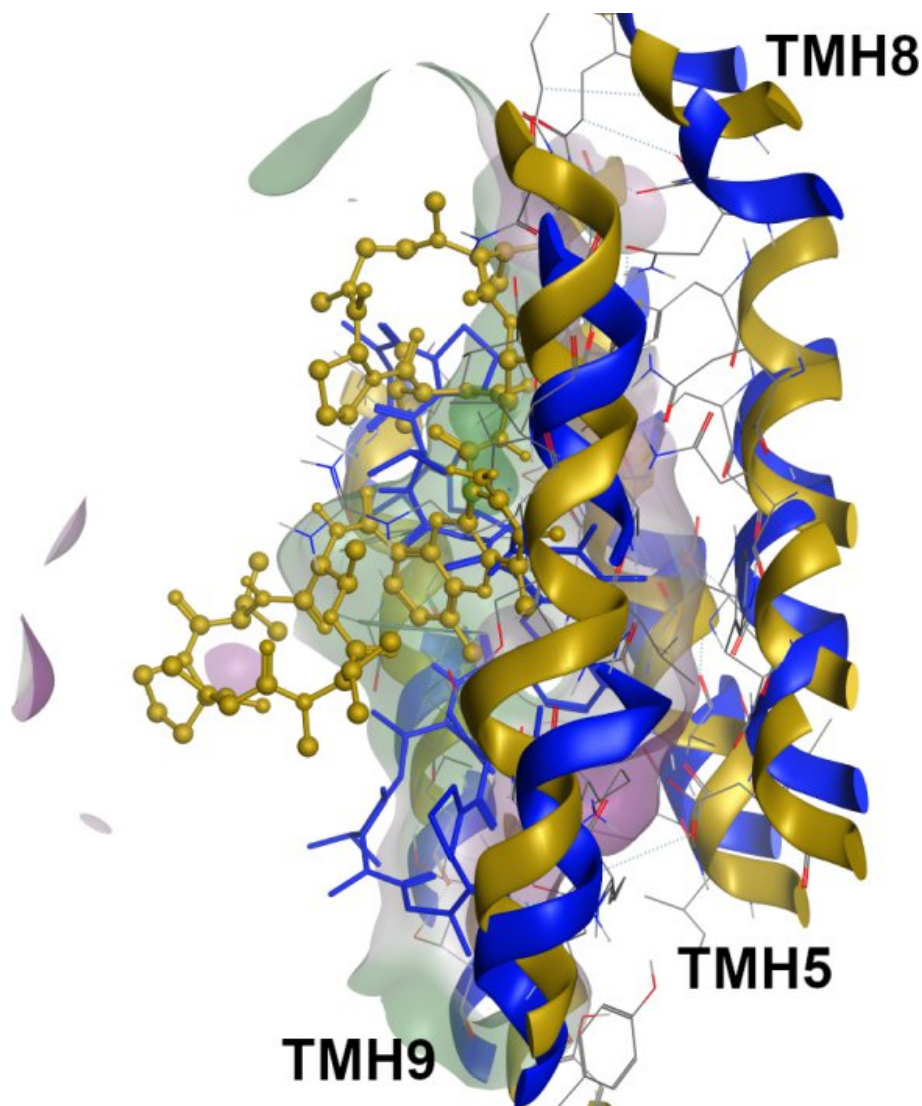
Figure 4.3 Superimposition of the actinomycin-bound systems at the R-site for the WT (blue, licorice) and G830V variant (dark yellow, ball and stick).

Since the contact efficiency (CE) ratio was higher in the G830V-R_ACT-R system than in WT-R_ACT-R but without significant change in binding energy, the loss of contacts between a large region of the molecule and the protein may explain the minor impact in the ACT ΔG_{bind} values observed for this variant, i.e the molecule changes the interaction pattern loosing interactions by making fewer but stronger ones. Oppositely, a visual inspection of the F978A variant showed that ACT remained deeply buried in the R-site as observed in WT, thus showing a similar contact efficiency ratio and, as expected, no changes in the ACT binding affinity.

Lastly, an increase in the CE ratio was observed for both F978A-M_CYC-M and ΔF335-M_CYC-M systems. Again, visual inspection showed that the structural changes at the M-site in the F978A variant prompted CYC to change the conformation of its macrocycle,

shifting one of the isoleucine side-chains into the more hydrophobic environment at the center of the macrocyclic core (**Figure 4.4**) and therefore improving the CYC binding affinity.
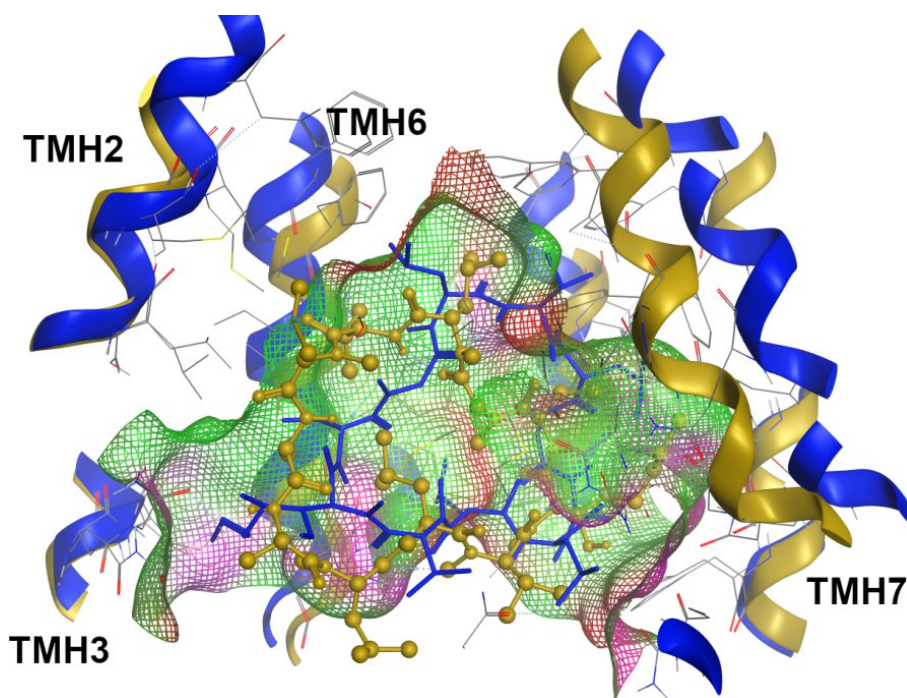
Figure 4.4 Superimposition of the cyclosporine-bound systems at the M-site for the WT (blue, licorice) and F978A variant (dark yellow, ball-and-stick).

Regarding the ΔF335-M_CYC-M system, no considerable changes in the CYC binding mode was observed in respect to WT, with all hydrophobic groups being protected from the solvent. However, the increase in the CYC binding affinity observed for these variants allow to infer that small changes of conformation in a large ligand may play important roles in binding affinity. Likewise, valspodar in the ΔF335-M_VLS-M system was also found to be deeply buried inside the M-site, which favors the reinforcement of the VLS-ΔF335 interactions, and consequently, its binding affinity. These results could be explained by the lower impact in the M-site volume and polarity reported for the F978A and ΔF335 variants *apo* systems, when compared to the SBSs where the structural impact of these mutations was higher (*see Paper I, Results and Discussion*).

Altogether, the above data indicate that these P-gp mutations seem to have an impact on the binding energy, ligand-binding modes and/or protein-ligand interactions. Moreover, although the mutations have an undeniably distinct but strong structural impact in the TMH repacking, similar effects on the binding energy observed for the same molecule across two different mutations (VIN-H in G185V and F978A; ACT-R in G830V and F978A; DOX-R in G830V and ΔF335; CYC-M in F978A and ΔF335), suggest that the molecular properties of these ligands may play important contributions to their

binding affinity rather than the severity of the TMH repacking induced by these mutations. This remark seems to be independent of the DBS where the molecule binds.

4.3.3 Interactions between coupling helices and nucleotide-binding domains

All the herein considered P-gp variants are experimentally described as having altered drug efflux profiles and changes in the basal and/or drug-stimulated ATPase activity (*see Paper I, Introduction*). Therefore, as the ICH-NBD residue interactions are increasingly considered to be key players in signal-transmission and efflux-related conformational changes, a thorough evaluation of the ICH-NBD contacts was performed and compared to the WT *holo* systems. The total number of contacts was estimated by the *gmx hbond* module (**Figures S53-S60**, *Annex S3*).

Overall, all P-gp *holo* systems revealed changes in the mean total number of contacts at the ICH-NBD signal-transmission interfaces. However, while in most cases a reinforcement in the total number of contacts was observed at the ICHs-NBD1 interfaces, a negative impact in the total number of contacts was observed in the ICHs interacting with NBD2 (**Table 4.3**).

Table 4.3 Structural impact of mutations in the total number of contacts at the ICHs-NBD interfaces for the P-gp variants *holo* systems.

		ICHs-NBD1		ICHs-NBD2	
Molecule-DBS	**Mutation-DBS**	**ICH1**	**ICH4**	**ICH2**	**ICH3**
G185V-H	COL-H	↑↑ [a]	NS	NS	↓ [b]
	VIN-H	NS	↑	NS	↓
G830V-R	ACT-R	↑↑	NS	↓	↓
	DOX-R	↑	↑	↓	↓
F978A-M	VIN-H	↑↑	↑↑	NS	NS
	ACT-R	↑	↑	↓↓	NS
	CYC-M	↑	↑↑	NS	NS
ΔF335-M	DOX-R	↑↑	↑↑	↓↓	NS
	CYC-M	↑↑↑	NS	↓	↓
	VLS-M	↑↑	↑	NS	NS

[a] Positive variation — the lowest impact (↑), the highest impact (↑↑↑); [b] negative variation (↓) — the lowest impact (↓), the highest impact (↓↓↓); NS — not significant. The P-gp variants *holo* systems were compared both with the WT and between them.

As all P-gp mutations in the presence of ligands seemed to affect the total number of contacts at the ICH-NBD interfaces, we aimed for the identification of which residue pairs were mostly involved through the evaluation of their mean contact frequencies, using the *g_contacts* tool. Only mean contact frequencies ≥ 0.5 and variations above 10% were considered meaningful (**Tables S28-S32**, *Annex S3*). For clarity purposes, the analysis of each ICH-NBD interface will be described separately:

4.3.3.1 ICH-NBD1 interfaces

Starting from the N-terminal P-gp halve, the first ICH-NBD interface is the ICH1-NBD1. Herein, all P-gp variants *holo* systems showed a significant increase in the mean contact frequencies between I160 and the NBD1 residue L443, described as important for ATP-binding.[103,117,217] On the contrary, a negative impact in the residue interactions was observed in most of the P-gp variants *holo* systems between D164-K405, also involved in the TMD-NBD communication pathway.[110] When concerning the hydrogen bond network, a negative effect in the HB lifetime was also observed in both G830V-R_ACT-R and F978A-M_ACT-R systems as well as in all ΔF335 variant *holo* systems. A decrease in the HB lifetime was also inferred from the G185V-H_VIN-H, but when in the presence of colchicine a significant reinforcement of the HB network in the G185V-H_COL-H system was observed.

On the other hand, all the P-gp variants showed clear differences in the residue interactions at the ICH4-NBD1, the second signal-transmission interface involving NBD1. Regarding the SBS-variants, similar effects in the residue interactions were observed in the G185V *holo* systems, mainly reinforcing the contact frequencies between Q912-R467, with R467 being involved in the propagation of the conformational changes.[32] Oppositely, a strong negative impact in the mean contact frequencies was observed for the residue pair T911-R467. A negative variation in the contact frequencies was also verified to occur for the residue pairs S909-Q441 and Q912-R464, with S909 being involved in the activation and ATPase stimulation when in the presence of drugs and/or lipids.[217,218] In the G830V-R_ACT-R system, the presence of ACT again induced a strong negative effect in the contacts involving the R905-Q438, T911-R467, and Q912-R464 residue pairs. Quite interestingly, mutational studies already had suggested that R905 is a pivotal residue in the activation and ATPase stimulation when in the presence of drugs and/or lipids.[217,218] Surprisingly, an opposite trend was found for both M-site variants. Both the F978A and ΔF335 *holo* systems showed a significant reinforcement in the mean contact frequencies between the R905-Y401/Q438 and V908/T911-R467 residue pairs, with Y401 (NBD1) again being important for ATP-binding.[36] Regarding

hydrogen bonds, the G830V-R_ACT-R system was the only one that showed a dramatic increase in the HB lifetime while a negative impact in the calculated HB parameters was observed in all other P-gp variants *holo* systems.

4.3.3.2 ICH-NBD2 interfaces

At the C-terminal P-gp halve, the first signal-transmission interface is the ICH2-NBD2. Herein, all P-gp variants *holo* systems showed a decrease in the mean contact frequencies between the residues A266 (ICH2) and F1086 (NBD2), a critical residue pair for coupling of ATP-binding to conformational changes in the TMDs.[97] Nonetheless, substantial differences in residue pair interactions were observed between the P-gp variants. For both G185V-H_COL-H and G185V-H_VIN-H systems, the presence of these ligands affected the R262-F1086 (positive variation) and I265-R1110 (negative variation) residue pairs, being R262 (ICH2) an important residue in the propagation of the conformational changes.[32] However, a noteworthy reinforcement in the mean contact frequencies between T263-Q1118 and T263-D1200 residue pairs was observed only for both G185V *holo* systems. Accordingly, Q1118 is an important residue for ATP-binding/hydrolysis and Mg^{2+} binding since mutations in this residue, located within the Q-loop, blocked drug-stimulated ATPase activity.[100,243] In contrast, the presence of ACT and DOX in the G830V-R_ACT-R and G830V-R_DOX-R systems had an overall negative impact in the mean contact frequencies involving the R262-Q1081 and T263-S1117 residue pairs. Regarding the M-site variants, a negative variation in the mean contact frequencies was observed in the F978A variant *holo* systems, namely between R1110 (NBD2) and residues I265/F267 (ICH2), described as involved in P-gp maturation and activity.[104,115] Similarly, the presence of DOX in the ΔF335-M_DOX-R system negatively affected most of the identified ICH-NBD contact points.

Changes in the ICH-NBD hydrogen-bond network were additionally observed for the G185V and M-site variants *holo* systems. While the G185V-H_COL-H system showed a decreased in the HB lifetime, VIN on the other hand reinforced the calculated HB parameters in the G185V-H_VIN-H system. Opposite effects were observed between CYC and VLS in their respective P-gp variants: while the HB lifetime was negatively affected by the presence of CYC in both F978A-M_CYC-M and ΔF335-M_CYC-M systems, the presence of VLS in the ΔF335-M_VLS-M system instead reinforced the HB network.

Finally, most of the P-gp *holo* systems also showed negative effects in the residue interactions at the ICH3-NBD2, the second signal-transmission interface interacting with NBD2. Most particularly, a significant decrease in the mean contact frequencies was

observed between the residue pairs V801/S802 and the NBD2 residue Y1087, which is thought to play an important role in P-gp activity and folding.[97] Moreover, the Y1087 residue was also predicted to be the key residue in the ICH3-NBD2 communication.[112] Similarly, and concerning the HB network, a general decrease in the HB lifetime was observed in the *holo* systems of the G185V and M-site variants.

Altogether the results indicate that the studied P-gp mutations *holo* systems i) have long-range effects in the residue interactions at the ICH-NBD signal-transmission interfaces, ii) induce asymmetries in the ICH-NBD residue interactions and iii) affect the mean contact frequencies of important residue pairs experimentally related to drug-stimulated ATPase activity, P-gp maturation and/or folding.

4.4 MECHANISTIC INSIGHTS

P-glycoprotein over-expression is one of the most relevant multidrug-resistance mechanisms in cancer cells. To develop compounds with high selectivity and efficacy towards P-gp, a deeper knowledge underlying the mechanisms of drug specificity and efflux-related signal-transmission is needed. In our previous work, a refined hP-gp WT model and four P-gp variants (G185V, G830V, F978A, and ΔF335), experimentally related to changes in substrate specificity, basal and drug-stimulated ATPase activity were built in the *apo* state inward-facing conformation, and their structural impact in the TMDs and ICH-NBD interactions was thoroughly evaluated (*see Paper I, Results and Discussion*). In the present study, these P-gp models were used to assess the structural effects of these mutations on drug-binding and efflux-related signal-transmission mechanism in the presence of ligands within the DBP, that are also experimentally described as having altered efflux patterns.

Taken together, our data indicate that these P-gp mutations in the presence of ligands within the DBP induce structural effects in the TMDs by affecting i) the DBP portals 4/6 and 10/12 with the possibility for compromising the access and release of substrates, and ii) the "crossing helices" 4/5 and 10/11 able of compromising the NBD dimerization upon ATP-binding. Although these findings are in agreement with those reported in our P-gp *apo* systems (*see Paper I, Results and Discussion*), the impact on the TMH repacking seems to be stronger in the P-gp *holo* systems, suggesting that P-gp is sensitive to the presence of ligands. Additionally, our results are also in agreement with the experimental studies performed by Loo and Clarke. Accordingly, substrates bind to P-gp through a "substrate-induced fit mechanism", where the size and shape of the

substrate induce rearrangements in the TMHs.[66] Furthermore, the rearrangement of the TMH residues' side-chains due to repacking affect the protein-ligand interactions and/or binding affinity. On the other hand, the presence of mutations and/or ligands induces variations in the total number of contacts at the ICH-NBD interfaces, possibly affecting the TMD-NBD communication, and consequently altering drug efflux. Lastly, the impact in the mean contact frequencies of specific ICH-NBD residue pairs, experimentally described as involved in the signal-transmission mechanism, folding and ATP-binding, offers a possible explanation for the alterations in drug efflux reported for these P-gp variants.

Interestingly, most of the P-gp variants *holo* systems showed asymmetrical changes in the total number of contacts at the ICH-NBD interfaces, namely a reinforcement in the number of contacts at the ICHs-NBD1, while a decrease in the contacts was observed in the ICHs connecting NBD2. Two experimental studies recently demonstrated that i) the presence of cholesterol in the membrane induces asymmetries between NBDs, involving the ICH-NBD interfaces,[244] and ii) there are common regions affected in a similar manner, but diverging in the post-hydrolysis state, especially in ICHs 3 and 4. Concerning the former, if taken into account that cholesterol is also a P-gp substrate,[244] the latter corroborates our hypothesis that the presence of the selected ligands within the DBP exacerbate the asymmetries in the contacts at the ICH-NBD interfaces. The fact by which similar changes in the residue's interactions at the ICH-NBD interfaces were also reported in these P-gp variants *apo* systems (*see Paper I, Results and Discussion*), especially concerning the M-site variants, additionally suggests that i) such alterations can be induced solely by the mutation, independently of the presence of molecules within the DBP, and ii) such alterations can be further exacerbated by the presence of molecules bound at the SBSs.

Nevertheless, important structural differences were observed between the P-gp variants *holo* systems. Regarding the G185V (TMH3) mutation, experimental data showed that it conferred increased resistance to COL in transfected cells, with an increase in the COL-stimulated ATPase activity, and a 3.6-fold decrease in COL binding. Oppositely, a decreased resistance to VIN was reported for this P-gp variant, with a 3.8- to 5.5-fold increase in the VIN binding (*see Paper I, Introduction*). Although no clear conclusions could be made about the possible changes in the COL binding affinity in the G185V-H_COL-H system, an indubitably increase in the VIN binding affinity was observed in the G185V-H_VIN-H system, in agreement with the experimental data.

Nonetheless, the presence of COL induced a reinforcement of the residue's interactions and HB network at the ICH1-NBD1, that was not observed in the presence of VIN. As ICH1 establishes interactions with the A-loop residues of NBD1 and the adenine group of ATP,[50] we hypothesized that the reinforcement of the ICH1-NBD1 interactions may be related to the increase in COL-stimulated ATPase activity reported in the experimental studies and, therefore, with the increased resistance to COL. Furthermore, a recent study emphasized the role of the conserved ICH1-NBD1 interface as critical for the cross-talk between the TMD-NBD domains in ABCG2, another relevant ABC transporter in the MDR in cancer cells.[245] On the other hand, in our models, the decrease in the VIN efflux experimentally observed in this P-gp variant, is solely due to the sharp increase in the VIN binding affinity. A previous experimental study performed by Loo and Clarke[133] focused on another single point mutation, F335A, which correlated the increase in the VIN binding affinity within the DBS with a decrease in the VIN efflux. According to the authors, the release of VIN during the transport cycle was postulated to be impaired due to its increased affinity at the DBS, and the herein published model corroborates those assumptions.

Located at the opposite P-gp halve, the G830V mutation (TMH9) in the presence of DOX (G830V-R_DOX-R) did not seem to significantly affect DOX binding affinity, as observed in the G185V-H_COL-H system. Although no more conclusions could be retrieved from our data, it clearly suggests that even minor effects in the ligands binding affinity may be related with the increased drug efflux. In contrast, and despite the overall increase in the contact efficiency ratio – which was expected to contribute for increasing the ACT-binding affinity – the severe changes in ACT binding mode observed in the G830V-R_ACT-R system seemed to act as a "counter-weight", having as consequence a non-significant change in the ΔG_{bind} values. Therefore, we suggested that in this specific case, the decrease in the ACT efflux is not related with the increase binding affinity as observed with VIN, but to the ACT unfavorable binding mode at the R-site, which may reduce the ability for the G830V variant in transport ACT. This hypothesis could be supported by experimental studies, where only a slight decrease resistance to ACT (0.29-fold) in respect to WT was observed in this P-gp variant.[105]

Overall, it becomes clear that the binding of substrates at the H and R sites has similar effects on drug efflux in both glycine variants. Additionally, these findings also corroborate our previous studies, where we suggested that G185 and G830 residues, located at opposite halves have equivalent roles in P-gp function, likely more involved in drug binding (*see Paper I, Results and Discussion*). Finally, and although no significant changes in the DOX-binding affinity in the ΔF335-M_DOX-R system, both F978A

(TMH12) and ΔF335 (TMH6) mutations tend to induce an increase on ligand-binding affinities, which seems to be independent of the DBS where the molecules interact. Therefore, we hypothesize that for certain ligands, the increase in drug binding affinity within the DBS is a possible mechanism underlying the decreased drug efflux.

CHAPTER 4 – LETTER

Probing the allosteric modulation of P-glycoprotein: A medicinal chemistry approach towards the identification of non-competitive P-gp inhibitors.

A medicinal chemistry approach combining *in silico* and *in vitro* methodologies was performed aiming at identifying and characterizing putative allosteric drug-binding sites (aDBS) at the interface of the transmembrane- and nucleotide-binding domains (TMD-NBD) of P-glycoprotein. Two aDBSs were identified, one in each NBD, by means of *in silico* fragment-based molecular dynamics, and characterized in terms of size, polarity, and lining residues. From a small library of thioxanthone and flavanone derivatives, experimentally described to bind at the TMD-NBD interfaces, several compounds were identified to be able to decrease the verapamil-stimulated ATPase activity. An IC50 of 81 ± 6.6 μM is reported for a flavanone derivative in the ATPase assays, providing evidence for an allosteric efflux modulation in P-glycoprotein. Molecular docking and Molecular Dynamics gave additional insights on the binding mode on how flavanone derivatives may act as allosteric inhibitors.

5.1 INTRODUCTION

Over-expression of membrane efflux pumps as P-glycoprotein (P-gp, ABCB1) is tightly related to the multidrug resistance (MDR) phenomenon in cancer cells, and chemotherapy failure. Modulating drug efflux by P-gp is, currently, still one of the most promising strategies to beat MDR in cancer. P-gp architecture consists of two transmembrane domains (TMDs) and two cytoplasmic nucleotide-binding domains (NBDs) organized in a pseudo 2-fold symmetry. Four short intracellular coupling helices (ICHs), located between the transmembrane helices (TMHs) 2/3 (ICH1-NBD1), 4/5 (ICH2-NBD2), 8/9 (ICH3-NBD2) and 10/11 (ICH4-NBD1), mediate the communication between both domains through non-covalent interactions. The drug-binding pocket (DBP) lies within the TMDs and it is capable of interacting with several structurally unrelated scaffolds. To date, three distinct generations of P-gp modulators were categorized, but all failed to demonstrate their efficacy and safety in the clinical environment. Furthermore, the development of more selective and effective P-gp modulators, using structure-based approaches, has also being hampered by the poly-specificity of the DBP.[246] Thus, novel strategies for P-gp efflux modulation are extremely important to reverse MDR in cancer cells.

Recently, specific motifs within the ABC architecture – the intracellular coupling helices (ICHs) – have been described as possible targets for small molecules. Targeting such motifs has several advantages, because they i) are highly conserved among the ABC transporter superfamily, ii) are involved in the propagation of the conformational changes from the inward- to outward-facing conformations, thus acting as TMD-NBD signal-transmission interfaces, iii) play important roles in the P-gp activity, folding and maturation, and iv) if targeted by small molecules it would be possible to avoid competition with the DBP, thus vastly reducing the adverse effects identified previously during the clinical trials of past MDR modulator generations. Additionally, this novel approach is also suitable for application to other members of the ABC transporter superfamily.[117]

Possible allosteric drug-binding sites (aDBSs) have been proposed at the ICH-NBD interfaces, able to bind small molecules that potentially block the TMD-NBD signal-transmission, responsible for driving conformational changes leading to efflux. Small molecules as flavonoids, thioxanthones and 1,4-dihydropyridine derivatives were predicted to interact in these regions. However, the first study that clearly identified a possible drug-binding region next to the ICHs was performed by Kim et al.,[122] in the

Arabidopsis thaliana P-gp homologue. The authors predicted that the ABCB1-dependent auxin transport inhibitors 2-[4-(diethylamino)-2-hydroxybenzoyl] benzoic acid and 1-*N*-naphthylphthalamic acid (BUM and NPA, respectively) bind in a pocket located between the ICHs and the Q-loops of NBDs. Furthermore, while NPA interacts in both NBDs, BUM preferentially binds at the NBD2 and in an additional region between NBD1 and NBD2.

5.2 RESULTS AND DISCUSSION

5.2.1 Identification and characterization of possible allosteric drug-binding sites at the nucleotide-binding domains

Therefore, and aiming to accurately identify any putative aDBSs at the ICH-NBD interfaces, a computational fragment-based drug discovery (cFBDD) approach was undertaken. The fragments were derived from i) molecules that are experimentally described to bind at the ICH-NBD interfaces such as BUM/NPA or flavonoids e.g. narigenin, and ii) *in silico* studies of molecules that also have one energetic minimum at similar regions, e.g. tariquidar, nicardipine, isoxazol-DHP and morphine (although experimentally a photolabeling site for dihydropyridines is also proposed to exist in close proximity to such motifs) (*see chapter 2, Methodology*). Nevertheless, preference was given to aromatic systems in which additional moieties (e.g. carboxylic acid, alcohols, amides, nitriles) and substitution patterns were chosen to maximize the coverage of the intended chemical space. The fragments obtained from each molecule are depicted in **Table S1** in the *Annex S4*.

Several simulated-annealing Molecular Dynamics (saMD) simulations were performed only using the cytoplasmic portion of the N- and C-terminal (concerning NBD1 and NBD2, respectively). For each NBD, five MD systems were built, each one comprising six fragment types and five copies of each fragment (for a total of 30 fragments), by randomly insertion in the surrounding water environment, using the GROMACS *gmx insert-molecules* tool. The fragments included in each MD system are depicted in **Table S2** in the *Annex S4*. Five replicates of each system were performed in a total of 25 MD runs per NBD. Following, occupancy volumetric maps were generated from the MD trajectories to identify possible hotspots at the ICH-NBD interfaces using the *VolMap* tool in VMD.

Overall, the occupancy maps (**Figure 5.1**) showed two important hotspots at the ICH-NBD interfaces in each NBD comprising i) both ICHs and the A-loop motif of the respective NBD, or ii) the ICH4/2 (NBD1 or NBD2, respectively) and the Q-loop/Walker A (WA) motifs in each NBD. In addition, while most fragments preferred to bind between the ICHs and the A-loop of NBD1, a closer matching distribution of the fragments was observed for NBD2.

Apart from occupancy maps, fragment types within each hotspot were also considered crucial to identify which chemotypes are preferred. Concerning the hotspot located in-between ICHs and closer to the A-loop motif of NBD1, 6-membered aromatic ring systems are preferred, but fused-ring systems are also tolerated, with a preference towards aromatic amines, hydroxy and methoxy groups (often in an *ortho* substitution pattern for the latter). Oppositely, heterocyclic rings such as pyrrole and pyridine, together with the benzylamine and benzylalcohol substructures, preferentially bind at the second hotspot, located closer to the ICH4/Q-loop/WA motifs.

On the other hand, a higher heterogeneity was observed for NBD2. Concerning the hotspot next to both ICHs and closer to the A-loop motif, single- (1,4-dihydropyridine) or fused-heterocyclic ring systems (quinoline, indole) are tolerated, as well as amine/amide groups and hydroxyl substituents (the latter in a meta substitution pattern). Interestingly, at the ICH2/Q-loop/WA hotspot, a preference for positively charged fragments was identified, namely the piperazine, 3-hydroxy-piperidine, or morpholine scaffolds (not observed in NBD1), together with other moieties such as benzylamine, benzyl alcohol, 1,4-dihydropyridine or trifluoromethylphenyl, and fused-ring systems (3-amidoquinolines, indoles).

These results indicate that although the same fragments' database was applied to both NBDs, different scaffolds were observed in the equivalent hotspots of the NBDs, thus suggesting an apparent degree of specificity between NBD1 and NBD2. More importantly, as some fragments derived from known binders as BUM (phenol, 2-hydroxybenzoate) and NPA (naphtalen-1-amine),[122] isoxazole-DHP (1,4-dihydropyridine),[119,121,186] or naringenin (benzene-1,2-diol, benzene-1,3,5-triol),[94,247] were also found to bind to these hotspots, the results are found to be in good agreement with experimentally available information.
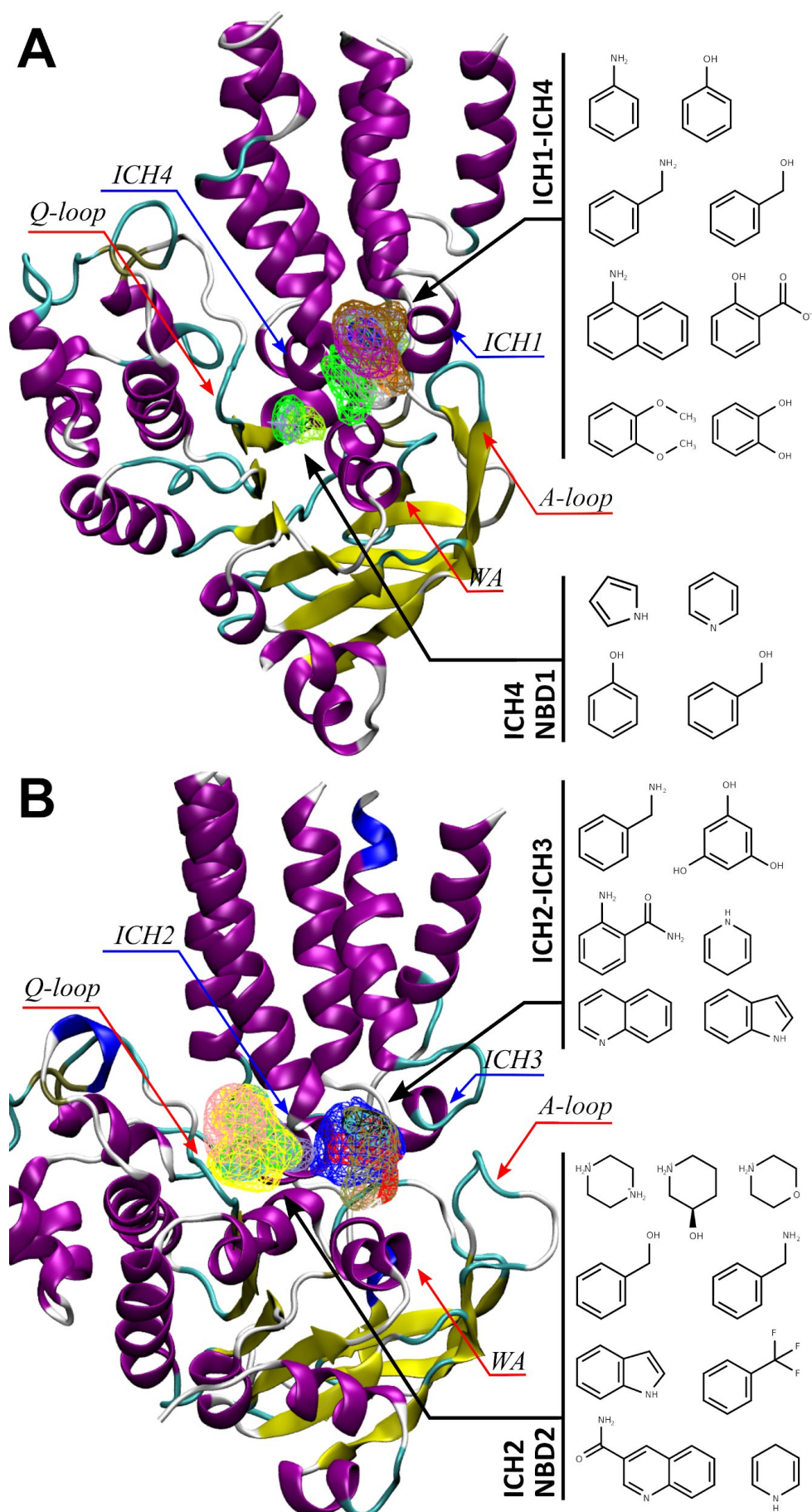
Figure 5.1 Occupancy maps and identified fragments at each hotspot for (A) NBD1 and (B) NBD2.

Upon the hotspots' identification, and to better characterize these novel regions, lining residues, pocket volume, residues distribution and mean polarity were assessed using the EPOS[BP] and MOE packages, and compared with the modulator site (M-site), located at the top (outer leaflet) of the DBP.

These results suggest that the two hotspots are probably part of a larger aDBS, herein named as aDBS1 for NBD1 and aDBS2 for NBD2. The mean polarities and residues distribution showed that both aDBSs are more polar than the M-site, containing 47 ± 1% of polar residues, against 30% of the M-site. Moreover, both allosteric sites also seem to have similar volumes (683 ± 21 Å^3) but, when concerning the solvent accessible surface area, aDBS2 seems to be slightly more exposed (~231 Å^2 *vs* ~198 Å^2 for aDBS1).

5.2.2 Molecular Docking of a small library of in-house compounds

Afterward, an evaluation on drug binding to this specific location was performed by molecular docking. This is a fast technique that can be swiftly used to assess the ability of small molecules in binding to the proposed allosteric drug-binding site. Moreover since it is an independent technique from the fragment-based molecular dynamics reported above, it can also provide a statistical check for results reliability. Based on the scaffolds found in each hotspot at the ICH-NBD interfaces, a small in-house library of thioxanthone[202] and flavanone derivatives[126,206] was selected for this purpose (**Table S3**, ***Annex S4***). Additionally, since BUM and NPA were experimentally predicted to bind in a pocket located between the ICHs and Q-loop motif of NBDs, these compounds were also included as references in our docking studies. The docking box was defined to include all occupancy regions between ICHs and the whole ATP-binding site (**Figure 5.1**). Ten docking poses were generated per molecule, but to simplify the results only the top-ranked binding energies (ΔG_{dock}) at each aDBS will be described.

Overall, all compounds tested showed favorable ΔG_{dock} values in both NBDs (**Table S4 and S5**, ***Annex S4***), similar to those observed at the M-site, ranging from -5.8 kcal/mol (BUM, NBD1) to -8.5 kcal/mol (compound **15**, NBD2), thus indicating that both aDBSs found are hypothetically druggable, i.e. capable of accommodating molecules with favorable binding energies. Moreover, the analysis of the protein-ligand contacts confirmed that the tested molecules overlapped the occupancy maps in both NBDs, interacting with most of the lining residues identified (**Tables S3 and S4**, ***Annex S4***) and independently corroborating the location of the aDBS in each NBD.

However, regarding NBD1, most of the compounds tested (including BUM and NPA), were found to protrude from the hotspot next to the ICH4/Q-loop/WA motifs, and thus partially overlapping the phosphate groups of ATP (or the coordinating magnesium ion). On the contrary, most of the molecules tested in NBD2, including BUM and NPA, did not overlap the ATP-binding site. Yet, these results are in good agreement with previous studies concerning flavonoids, where a partial overlapping between a flavonoid-binding region and the ATP-binding site was inferred from experimental data.[94,247] Additionally, and when compared with thioxanthones, flavanone derivatives also have higher probabilities of occupying the hotspot in-between ICH2 and ICH3, which can be prone to "lock" ICH2 and ICH3 together.

Altogether, the above data is consistent with the existence of an allosteric drug-binding site in each TMD-NBD interface, vicinal to where ATP binds, involving the ICHs and the A-loop, Q-loop, and WA motifs of the respective NBD. In addition, the favorable binding energies obtained for all molecules in the molecular docking studies led to assume that, indeed, both aDBSs are expected to be druggable.

5.2.3 Verapamil-stimulated ATPase assays

As most of the compounds tested in our docking studies bind at this vicinal region, with or without partially overlapping the ATP-binding site, ATPase assays were conducted to evaluate their effect on drug-stimulated P-gp ATPase activity. Herein, the tested compounds were incubated with 500 μM of verapamil, a P-gp substrate that stimulates the P-gp ATPase activity.[207] The ATPase assays were carried out in recombinant human P-gp membranes, using the Pgp-Glo™ Assay System (Promega, The Netherlands), and according to the manufacturer's experimental protocol (*see chapter 2, Methodology*).[205] Firstly, an initial screening of the in-house library of thioxanthone and flavanone derivatives was performed. Two reference compounds, BUM – experimentally validated as an allosteric inhibitor of the *A. thaliana* P-gp homologue[122] – and the triterpene spiropedroxodiol – a potent competitive P-gp efflux inhibitor[206] – were also included in the ATPase assays for comparison purposes. The basal P-gp ATPase activity and verapamil-stimulated ATPase activity in the absence or presence of compounds were estimated against the sodium orthovanadate (Na_3VO_4) treated samples, which is a strong P-gp inhibitor,[208] in accordance with the Pgp-Glo™ Assay System technical bulletin. The data were obtained in relative light units (RLU) but for clarity purposes, the results were normalized, using as reference the basal P-gp ATPase activity values (**Figure 5.2** and **Table S6/Figures S1-S2**, *Annex S4*).
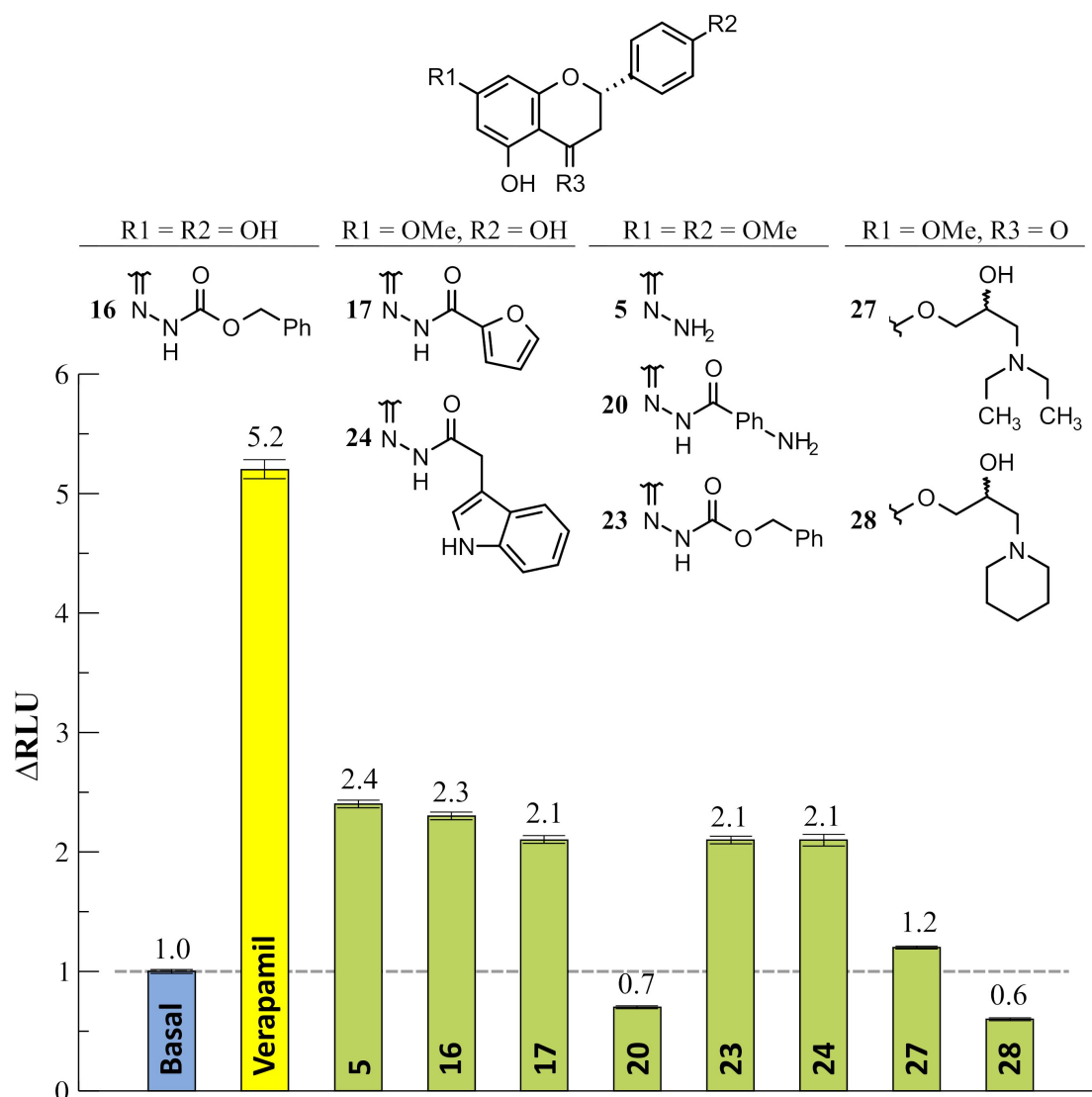
Figure 5.2 Normalized ATPase activity values (only for the compounds chosen for IC_{50} determination). The chemical structures of the selected compounds are depicted.

The data showed that all the tested compounds (except the thioxanthone **TX3,** and the flavanones **7** and **14**) were able to inhibit the verapamil-stimulated ATPase activity. Herein, while most thioxanthones had a weak activity on inhibiting ATPase, compounds **5, 16**, **17**, **20**, **23**, **24**, **27**, and **28,** from the flavanone derivatives set, were classified as moderate to good ATPase inhibitors. The reference compounds BUM and spiropedroxodiol were also found to inhibit the verapamil-stimulated ATPase activity, but with distinct potencies. As anticipated, BUM, described as a potential ICH-NBD binder,[122] had a stronger impact on decreasing the verapamil-stimulated ATPase activity than spiropedroxodiol, a potent efflux P-gp inhibitor that showed competition with Rhodamine-123 (R123) inside the DBP at the nanomolar range.[126,206]

Interestingly, all the other compounds with lower ATPase inhibition rates (compounds **1-4**, **6-15**, **18**, **19**, **21**, **22**, **25** and **26**) were previously described as competitive efflux

modulators in the R123 accumulation assay,[126,206] with fluorescence activity ratios (FAR) ranging from 2.10 (**7**) up to 10.67 (**8**) at 2.0 µM. The only exceptions are compounds **27** (FAR, 4.06; ΔRLU, 1.2) and **28** (FAR, 3.08; ΔRLU, 0.6). This reinforces our hypothesis that the selected flavanones are expected to act through a different mechanism, most probably as allosteric inhibitors.

Next, the compounds that had the highest inhibitory effect in verapamil-stimulated ATPase activity (**Figure 5.2**), were chosen for calculating the respective half maximal inhibitory concentration (IC_{50}) values (**Table 5.1**).

Table 5.1 IC_{50} values determined from ATPase inhibition experiments (n = 3).

Compound	Classification	IC_{50} (µM)
5	inhibitor	188 ± 34
16	non-inhibitor	—
17	non-inhibitor	—
20	inhibitor	> 200
23	inhibitor	89 ± 6.6
24	inhibitor	> 200
27	inhibitor	> 200
28	inhibitor	> 200

Test compounds were evaluated at a range of concentrations for their capability to inhibit 500 µM of verapamil-stimulated ATPase activity, according with the experimental protocol. The obtained IC_{50} results allowed the identification of the flavanone derivative **23** (**Figure 5.3**) as a promising hit for novel P-gp allosteric inhibitors, with an IC_{50} of 81 ± 6.6 µM.
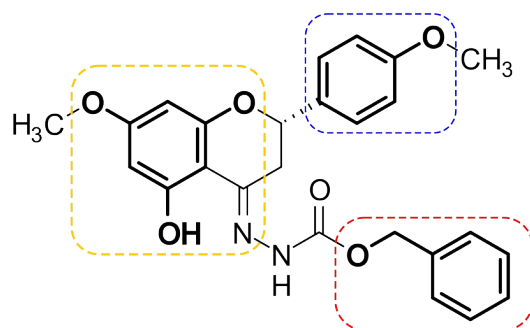
Figure 5.3 Structure of compound **23**. Matching fragments with cFBDD are highlighted (red, phenylmethanol; purple, *p*-methoxyphenyl; yellow, benzene-1,3,5-triol).

5.2.4 Molecular Dynamics studies

As flavonoids are described to interact preferentially with NBD2,[94] multiple MD runs were performed using the top-ranked docking pose of compound **23** obtained at the NBD2 as a starting configuration. Five replicates were carried out, in a total of 300 ns of simulation time, and their relative free-energies of binding (ΔG) were estimated using the *g_mmpbsa* tool (*see chapter 2, Methodology*). The results showed two possible binding modes (BMs) for compound **23** with similar binding energies (**Figure 5.4**). The first one (BM$_1$) was obtained in three out of five simulations, having an ΔG_1 of -27 ± 2.4 kcal/mol, with A259/A260 (ICH2) and Q-loop residue Q1118 as the residues with higher contributions for binding.

The second binding mode (BM$_2$) was found to occur in the remaining two simulations, in which compound **23** was found between the ICH2 and ICH3, with an average ΔG_2 of -26 ± 0.3 kcal/mol. Herein, residues L258, A259, I261 (ICH2), F804, and T811 (ICH3) had the highest contribution for ligand binding. Quite interestingly, F804 corresponds to the phenylalanine residue at position 792 (F792) in the *A. thaliana* sequence and was the residue specifically described to be experimentally involved in BUM binding.
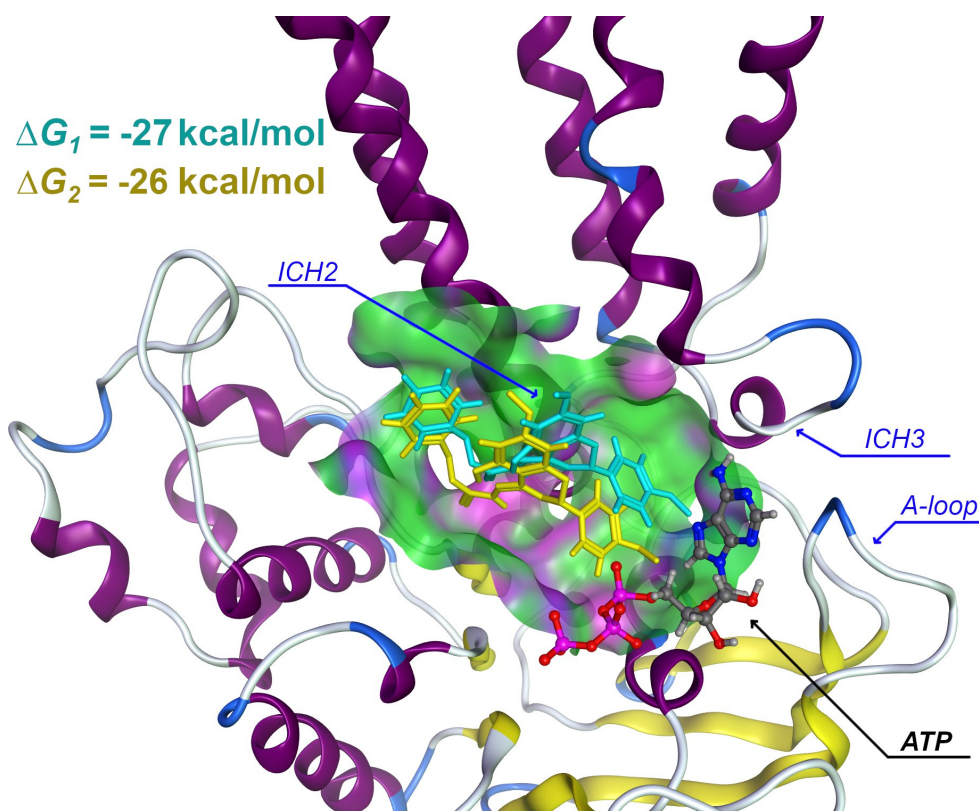
Figure 5.4 Averaged final configurations for compound **23**, obtained from MD simulations (yellow, *n* = 3 poses; cyan, *n* = 2 poses). Pocket surface is colored by *ActiveLP* (pink, H-bonding; green, hydrophobic; blue, mild polar). ATP pose was obtained from PDB ID: 6C0V and is depicted for comparison purposes only.

5.3 Hypothesis on allosteric inhibition by compound 23

Altogether, the above data allowed to infer that compound **23** might not have a preferential binding mode, and thus we hypothesize that the observed ATPase inhibitory activity result instead from a contribution of two different binding modes, corresponding also to two different mechanisms. The analysis of BM_1 for compound **23** suggests that the signal-transmission mechanism may be impaired by disrupting the contacts between the ICH2 and the Q-loop of NBD2. Accordingly, some studies identified the glutamine located within the Q-loop (Q1118) to have an important role for ATP-binding/hydrolysis and Mg^{2+} binding, as mutations in this residue blocked drug-stimulated ATPase activity.[123] Furthermore, the Q1118 residue is also described as a key residue in coupling events occurring in the DBP to ATP binding.[51,248]

More interestingly, BM_2 additionally suggests an agreement with a mechanism proposed by Loo and Clarke.[111] As demonstrated by the authors, the cross-linking of the mutant A259C/W803C (residues located at ICHs 2 and 3, respectively) inhibited the P-gp drug-stimulated ATPase activity possibly by impairing the conformational changes that occurred at the ICH2/ICH3 interface upon ATP binding. Therefore, we hypothesize that compound **23** may induce a similar "locking" of ICHs 2 and 3, impairing the signal-transmission mechanism triggered by verapamil at the DBP through the inhibition of the conformational changes promoted by ATP binding.

Finally, the possibility for compound **23** to act as a non-competitive inhibitor is reinforced by a previous published work made in our group using the same flavanone derivatives dataset. Quite interestingly, compound **23** – corresponding to compound **31** in previous work[126] – had much higher cytotoxicity in the ABCB1-overexpressing L5178Y (L5178Y-MDR) than in parental cell lines (7.25 μM vs > 100 μM, respectively), but was unable to promote R123 accumulation in resistant cells (FAR of 0.94 and 1.13 at 2.0 and 20 μM, respectively). Yet, compound 23 was still able to synergistically enhance doxorubicin cytotoxicity. These findings led the authors to suggest that this particular compound modulates P-gp drug efflux by a different mechanism. Herein, and based on our findings, we hypothesize that compound **23** is able to act as an allosteric inhibitor by targeting the ICHs-NBD2 interface and impairing signal-transmission, thus blocking the efflux-related conformational changes that ultimately leads to drug efflux.

Summarizing, in this letter we have provided further evidences that i) an allosteric binding site at the TMD-NBD interface exists within the P-gp architecture, as demonstrated by the binding of small fragments derived from known molecules, ii) the proposed binding site is in agreement with previous literature, slightly overlapping the ATP-binding site but

with different characteristics in each of the NBDs, iii) tested thioxanthone and flavanone derivatives bind to the proposed binding site with favorable affinities, thus inferring that the aDBSs are druggable and iv) the flavanone scaffold may be a suitable building block for the design of novel allosteric P-gp modulators.

The above results demonstrate a "proof-of-concept" for allosteric efflux modulation of ABC transporters, at least regarding P-glycoprotein. This opens future perspectives concerning the development of a new generation of P-gp efflux modulators bypassing problems found with modulators targeting the poly-specific DBS. Moreover, it also enables the usage of computational techniques such as fragment expansion and high-throughput virtual screening to guide an expedite discovery of new P-gp modulators.

CHAPTER 6 - CONCLUSION

Over-expression of membrane efflux pumps such as P-gp, MRP1, and BCRP, is one of the most significant mechanisms in MDR in cancer cells, thus contributing for the decrease of the intracellular concentration of anticancer drugs, and chemotherapy failure. MDR cancer cells can be re-sensitized to the pharmacological action of chemotherapeutics by co-administration with compounds able to decrease drug efflux – efflux modulators. Specially concerning P-gp, three generations of modulators were developed but none of them proved to be effective to reverse MDR in cancer and safe in the clinical environment. The poly-specificity of the DBP, able to bind and transport a wide range of structurally unrelated compounds, is pointed to the major disadvantage for the development of more potent and selective compounds, able to modulate P-gp drug efflux. Therefore, more insights on the molecular mechanisms of drug efflux, and a novel approach to modulate drug efflux by P-gp are critical for the development of high-affinity modulators with improved pharmacokinetic properties.

In the recent decades, the ICHs located at the TMD-NBD interfaces have been described as key players in the signal-transmission mechanism, responsible for driving the conformational changes needed for drug efflux. By targeting the ICHs by small molecules it would be possible to i) circumvent the competition with the DBP, thus reducing the adverse effects identified in the previously generations of MDR modulators, ii) apply to other efflux pumps of the ABC transporter superfamily, and iii) impair the P-gp expression, maturation and/or folding. Moreover, evidence of the existence of putative allosteric DBSs at the ICH-NBD interfaces has also been reported in the literature, thus opening new working hypotheses for the modulation of P-gp by an allosteric mechanism.

Therefore, the work described in this thesis aimed to i) obtain new insights on drug specificity and efflux-related signal-transmission mechanism, ii) identify and characterize putative allosteric DBSs at the ICH-NBD interfaces, and iii) identify potential scaffolds able to bind in the previously identidied DBS(s), and block the signal-transmission between the TMDs and NBDs.

In the first part of this project, a human P-gp homology model was built using a murine P-gp crystallographic template, refined by means of Molecular Dynamics simulations and validated using several approaches. Then, the equilibrated human P-gp model was used as template for building new systems with the most common mutations reported in literature (G185V, G830V, F978A and ΔF335), experimentally related to changes in drug efflux and drug-stimulated ATPase activity. The impact of these four mutations, located in each DBS — G185V (H-site), G830V (R-site), F978A (M-site), and ΔF335 (M-site) – in the P-gp architecture, was evaluated using Molecular Dynamics. The studies were

conducted in the absence (*apo* systems) and in the presence (*holo* systems) of molecules within the DBP, that experimentally proved to have altered efflux profiles in these variants. From these studies, it becomes clear that the P-gp efflux pump is sensitive to perturbations at the TMDs by mutations and/or the presence of ligands within the DBP, having both a direct impact in the TMHs repacking, with the possibility for compromising the access and release of drugs, and the NBD dimerization.

As a consequence of the TMHs rearrangement, a severe impact in the DBSs properties such as volume, shape and polarity was observed, mostly affecting the binding affinity and binding mode of the ligands tested, in particular those that interacted with the H and R sites. These findings suggest that the TMH repacking induced by perturbations in the TMDs, is responsible for driving changes in drug specificity, and offer a plausible explanation for the DBP poly-specificity. Moreover, this new data is also in agreement with previous literature claiming that changes in the membrane environment may induce similar effects in the P-gp architecture, and thus affecting drug specificity and signal-transmission mechanism.

Quite interestingly, the TMHs repacking observed in both *apo* and *holo* systems affect the residues interactions at the ICH-NBD signal-transmission interfaces, specially those reported in the literature as having important roles in the drug-stimulated ATPase activity, P-gp maturation and/or folding, thus emphasizing the critical role of the ICHs in the TMD-NBD communication.

Hence, the second part of this project consisted in exploring how druggable this region is, and if small molecules bound to those potential allosteric sites can act as P-gp modulators. Thereby, by means of computational Fragment-Based Drug Discovery, two allosteric DBSs were identified, located in-between the ICHs and the A-loop, Q-loop and the Walker A motifs in each NBD. Molecular docking data indicates that both allosteric DBSs are able to bind small molecules as thioxanthone and flavanone derivatives with favorable binding energies. In addition, and despite a slight overlap between the proposed allosteric DBSs and the ATP-binding site, most of the compounds evaluated in the ATPase assays proved to successfully reduce the P-gp verapamil-stimulated ATPase activity, thus inferring that these putative allosteric DBSs are hypothetically druggable.

Finally, this work demonstrates the "proof-of-concept" of allosteric efflux modulation for P-gp, being one of the scaffolds — a flavanone derivative — identified as possible "hit" for the design of allosteric P-gp modulators. By means of Molecular Dynamics simulations, two mechanisms for allosteric inhibition of the signal-transmission are

proposed: i) by disrupting the contacts between ICHs and Q-loop motif of NBD, and/or ii) by "locking" both ICHs in each TMD-NBD interface, thus impairing the conformational changes required for drug efflux. Again, these findings are in total agreement with several hypothetical models earlier proposed in literature.

Additionally, the identification of the flavanone scaffold as possible allosteric P-gp modulator also opens perspectives for the development of new scaffolds using computational techniques such as fragment expansion, *de novo* drug design and structure-based virtual screening of commercial libraries, such as Enamine, MCULE or MolPort, to expedite and guide the discovery of new allosteric P-gp modulators.

REFERENCES

1. Bukowski, K., Kciuk, M. & Kontek, R. Mechanisms of multidrug resistance in cancer chemotherapy. *Int. J. Mol. Sci.* **21**, 3233 (2020).

2. Juliano, R. L. & Ling, V. A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants. *Biochim. Biophys. Acta.* **455**, 152–162 (1976).

3. Leier, I. *et al.* The MRP gene encodes an ATP-dependent export pump for leukotriene C4 and structurally related conjugates. *J. Biol. Chem.* **269**, 27807–27810 (1994).

4. Allikmets, R. *et al.* A human placenta-specific ATP-binding cassette gene (ABCP) on chromosome 4q22 that is involved in multidrug resistance. *Cancer Res.* **58**, 5337–5339 (1998).

5. Doyle, L. A. *et al*. A multidrug resistance transporter from human MCF-7 breast cancer cells. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 15665–15670 (1998).

6. Miyake, K. *et al.* Molecular cloning of cDNAs which are highly overexpressed in mitoxantrone-resistant cells: Demonstration of homology to ABC transport genes. *Cancer Res.* **59**, 8–13 (1999).

7. Chartrain, M. *et al*. Melanoma chemotherapy leads to the selection of ABCB5-expressing cells. *Plos One.* **7**, e36762 (2012).

8. Frank, N. Y. *et al.* ABCB5-mediated doxorubicin transport and chemoresistance in human malignant melanoma. *Cancer Res.* **65**, 4320–4333 (2005).

9. Montanari, F. & Ecker, G. F. Prediction of drug-ABC-transporter interaction-Recent advances and future challenges. *Adv. Drug Deliv. Rev.* **86**, 17–26 (2015).

10. Wang, J. Q. *et al.* ATP-binding cassette (ABC) transporters in cancer: A review of recent updates. *J. Evid. Based Med.* **14**, 232–256 (2021).

11. Szakács, G. *et al.* Targeting the Achilles Heel of Multidrug-Resistant Cancer by Exploiting the Fitness Cost of Resistance. *Chem. Rev.* **114**, 5753–5774 (2014).

12. Gottesman, M. M. Mechanisms of cancer drug resistance. *Annu. Rev. Med.* **53**, 615–627 (2002).

13. Reyes, C. L. & Chang, G. Structure of the ABC transporter MsbA in complex with ADP.vanadate and lipopolysaccharide. *Science* **308**, 1028–1031 (2005).

14. Chang, G. *et al.* Retraction of 'Structure of MsbA from E. coli: A homolog of the multidrug resistance ATP binding cassette (ABC) transporter' [*Science* **293**, 1793–1800 (2001)]. *Science* **314**, 1875 (2006).

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

16. Aller, S. G. *et al.* Structure of P-Glycoprotein Reveals a Molecular Basis for Poly-Specifc Drug Binding. *Science.* **323**, 1718–1722 (2009).

17. Sonveaux, N., Vigano, C., Shapiro, A. B., Ling, V. & Ruysschaert, J-M. Ligand-mediated tertiary structure changes of reconstituted P-glycoprotein: A tryptophan fluorescense quenching analysis. *J. Biol. Chem.* **274**, 17649–17654 (1999).

18. Lu, P., Liu, R. & Sharom, F. J. Drug transport by reconstituted P-glycoprotein in proteoliposomes. Effect of substrates and modulators, and dependence on bilayer phase state. *Eur. J. Biochem.* **268**, 1687–1697 (2001).

19. Loo, T. W., Clarke, D. M. Identification of residues in the drug-binding site of human P-glycoprotein using a thiol-reactive substrate. *J. Biol. Chem.* **272**, 31945–31948 (1997).

20. Loo, T. W. & Clarke, D. M. 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 (2000).

21. Loo, T. W. & Clarke, D. M. 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 (2001).

22. Szewczyk, P. *et al.* Snapshots of ligand entry, malleable binding and induced helical movement in P-glycoprotein. *Acta Crystallogr. D Biol. Crystallogr.* **71**, 732–741 (2015).

23. Esser, L. *et al.* Structures of the Multidrug Transporter P-glycoprotein Reveal Asymmetric ATP Binding and the Mechanism of Polyspecifcity. *J. Biol. Chem.* **292**, 446–461 (2017).

24. Verhalen, B. *et al.* Energy transduction and alternating access of the mammalian ABC transporter P-glycoprotein. *Nature.* **543**, 738–741 (2017).

25. Le, C. A., Harvey, D. S. & Aller, S. G. Structural definition of polyspecific compensatory ligand recognition by P-glycoprotein. *IUCrJ.* **7**, 663–672 (2020).

26. Nicklisch, S. C. *et al.* Global marine pollutants inhibit P-glycoprotein: Environmental levels, inhibitory effects, and co-crystal structure. *Science Adv.* **2**, e1600001 (2016).

27. Romsicki, Y. & Sharom, F. J. The ATPase and ATP-binding functions of P-glycoprotein. Modulation by interaction with defined phospholipids. *Eur. J. Biochem.* **256**, 170–178 (1998).

28. Orlowski, S., Martin, S. & Escargueil, A. P-glycoprotein and 'lipid rafts': Some ambiguous mutual relationships (floating on them, building them or meeting them by chance?). *Cell Mol. Life Sci.* **63**, 1038–1059 (2006).

29. Klappe, K., Hummel, I., Hoekstra, D. & Kok, J. W. Lipid dependence of ABC transporter localization and function. *Chem. Phys. Lipids.* **161**, 57–64 (2009).

30. Ferreira, R. J., Ferreira, M. J. U. & dos Santos, D. J. V. A. Insights on P-glycoprotein's efflux mechanism obtained by molecular dynamics simulations. *J. Chem. Theory Comput.* **8**, 1853–1864 (2012).

31. Ferreira, R. J., dos Santos, D. J. V. A. & Ferreira, M.-J. U. P-glycoprotein and membrane roles in multidrug resistance. *Future Med. Chem.* **7**, 929–946 (2015).

32. Ward, A. B. *et al.* 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 (2013).

33. Sato, T. *et al.* Functional role of the linker region in purifed human P-glycoprotein. *FEBS J.* **276**, 3504–3516 (2009).

34. Ferreira, R. J., Ferreira, M. J. U. & dos Santos, D. J. V. A. Assessing the stabilization of P-glycoprotein's nucleotide-binding domains by the linker, using molecular dynamics. *Mol. Inf.* **32**, 529–540 (2013).

35. Kodan, A. *et al.* Structural basis for gating mechanisms of a eukaryotic P-glycoprotein homolog. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 4049–4054 (2014).

36. Kim, Y. & Chen, J. Molecular structure of human P-glycoprotein in the ATP-bound, outward-facing conformation. *Science.* **359**, 915–919 (2018).

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

38. Tonghin, N. *et al.* Novel features in the structure of P-glycoprotein (ABCB1) in the post-hydrolytic state as determined at 7.9 Å resolution. *BMC Struct. Biol.* **18**, 17 (2018).

39. Shapiro, A. B. & Ling, V. Positively cooperative sites for drug transport by P-glycoprotein with distinct drug specifcities. *Eur. J. Biochem.* **250**, 130–137 (1997).

40. Shapiro, A. B. & Ling, V. Transport of LDS-751 from the cytoplasmic leafet of the plasma membrane by the rhodamine-123- selective site of P-glycoprotein. *Eur. J. Biochem.* **254**, 181–188 (1998).

41. Ferreira, R. J., Ferreira, M. J. U. & dos Santos, D. J. V. A. Molecular docking characterizes substrate-binding sites and efflux modulation mechanisms within P-glycoprotein. *J. Chem. Inf. Model.* **53**, 1747–1760 (2013).

42. Alam, A. *et al.* Structure of a zosuquidar and UIC2-bound human-mouse chimeric ABCB1. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E1973–E1982 (2018).

43. Alam, A., Kowal, J., Broude, E., Roninson, I. & Locher, K. P. Structural insight into substrate and inhibitor discrimination by human P-glycoprotein. *Science.* **363**, 753–756 (2019).

44. Nosol, K. *et al.* Cryo-EM structures reveal distinct mechanisms of inhibition of the human multidrug transporter ABCB1. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 26245–26253 (2020).

45. Zhang, Y., Gong, W., Wang, Y., Liu, Y. & Li, C. Exploring movement and energy in human P-glycoprotein conformational rearrangement. *J. Biomol. Struct. Dyn.* **37**, 1104–1119 (2019).

46. Xing, J., Huang, S., Heng, Y., Mei, H. & Pan, X. Computational insights into allosteric conformational modulation of P-glycoprotein by substrate and inhibitor binding. *Molecules.* **25**, 6006 (2020).

47. Barbieri, A. *et al.* Structure of ABCB1/P-glycoprotein in the presence of the CFTR potentiator Ivacaftor. *Membranes.* **11**, 923 (2021).

48. Rosenberg, M. F. *et al.* Repacking of the transmembrane domains of P-glycoprotein during the transport ATPase cycle. *EMBO J.* **20**, 5615–5625 (2001).

49. Holland, I. B., Blight, M. A. ABC-ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans. *J. Mol. Biol.* **293**, 381–399 (1999).

50. Becker, J. *et al.* Molecular models of human P-glycoprotein in two diferent catalytic states. *BMC Struct. Biol.* **9**, 1–18 (2009).

51. Loo, T. W. & Clarke, D. M. Attachment of a "molecular spring" restores drug-stimulated ATPase activity to P-glycoprotein lacking both Q loop glutamines. *Biochem. Biophys. Res. Commun.* **483,** 366-370 (2017).

52. Loo, T. W. & Clarke, D. M. A short cross-linker activates human P-glycoprotein missing a catalytic carboxylate. *Biochem. Pharmacol.* **145**, 27–33 (2017).

53. Callaghan, R., Ford, R. C. & Kerr, I. D. The translocation mechanism of P-glycoprotein. *FEBS Lett*. **580**, 1056–1063 (2006).

54. Senior, A. E., Al-Shawi, M. K. & Urbatsch, I. L. The catalytic cycle of P-glycoprotein. *FEBS Lett.* **377**, 285–289 (1995).

55. Siarheyeva, A., Liu, R. & Sharom, F. J. Characterization of an asymmetric occluded state of P-glycoprotein with two bound nucleotides: Implications for catalysis. *J. Biol. Chem.* **285**, 7575–7586 (2010).

56. George, A. M. & Jones, P. M. Perspectives on the structure-function of ABC transporters: The switch and constant contact models. *Prog. Biophys. Mol. Biol.* **109**, 95–107 (2012).

57. Jones, P. M & George, A. M. Opening of the ADP-bound active site in the ABC transporter ATPase dimer: Evidence for a constant contact, alternating sites model for the catalytic cycle. *Proteins*. **75**, 387–396 (2009).

58. Zoghbi, M. E. et al. Substrate-induced conformational changes in the nucleotide-binding domains of lipid bilayer-associated P-glycoprotein during ATP hydrolysis. *J. Biol. Chem.* **292**, 20412-20424 (2017).

59. Dastvan, R., Mishra, S., Peskova, Y. B., Nakamoto, R. K. & Mchaourab, H. S. Mechanism of allosteric modulation of P-glycoprotein by transport substrates and inhibitors. *Science.* **364**, 689–692 (2019).

60. Dey, S., Ramachandra, M., Pastan, I., Gottesman, M. M. & Ambudkar, S. V. Evidence for two nonidentical drug-interaction sites in the human P-glycoprotein. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 10594–10599 (1997).

61. Shapiro, A. B., Corder, A. B. & Ling, V. P-glycoprotein-mediated Hoechst 33342 transport out of the lipid bilayer. *Eur. J. Biochem.* **250**, 115–121 (1997).

62. Shapiro, A. B., Fox, K., Lam, P. & Ling, V. Stimulation of P-glycoprotein-mediated drug transport by prazosin and progesterone. Evidence for a third drug-binding site. *Eur. J. Biochem.* **259**, 841–850 (1999).

63. Condren, M. E. & Bradshaw, M. D. Ivacaftor: a novel gene-based therapeutic approach for cystic fibrosis. *J. Pediatr. Pharmacol. Ther.* **18**, 8–13 (2013).

64. Loo, T. W. & Clarke, D. M. Identification of residues in the drug-binding domain of human P-glycoprotein. Analysis of transmembrane segment 11 by cysteine-scanning mutagenesis and inhibition by dibromobimane. *J. Biol. Chem.* **274**, 35388–35392 (1999).

65. Loo, T. W. & Clarke, D. M. Vanadate trapping of nucleotide at the ATP-binding sites of human multidrug resistance P-glycoprotein exposes diferent residues to the drug-binding site. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 3511–3516 (2002).

66. Loo, T. W., Bartlett, M. C. & Clarke, D. M. Substrate-induced conformational changes in the transmembrane segments of human P-glycoprotein. Direct evidence for the substrate-induced ft mechanism for drug binding. *J. Biol. Chem.* **278**, 13603–13606 (2003).

67. Loo, T. W. & Clarke, D. M. Mapping the binding site of the inhibitor tariquidar that stabilizes the first transmembrane domain of P-glycoprotein. *J. Biol. Chem.* **290**, 29389–29401 (2015).

68. Chufan, E. E. *et al.* Multiple transport-active binding sites are available for a single substrate on human P-glycoprotein (ABCB1). *Plos One.* **8**, e82463 (2013).

69. Subramanian, N., Schumann-Gillett, A., Mark, A. E. & O'Mara, M. L. Probing the pharmacological binding sites of P-glycoprotein using umbrella sampling simulations. *J. Chem. Inf. Model.* **59**, 2287–2298 (2019).

70. Hafkemeyer, P. *et al.* Contribution to substrate specifcity and transport of nonconserved residues in transmembrane domain 12 of human P-glycoprotein. Biochemistry. 37, 16400–16409 (1998).

71. Loo, T. W. & Clarke, D. M. The packing of the transmembrane segments of human multidrug resistance P-glycoprotein is revealed by disulfide cross-linking analysis. *J. Biol. Chem.* **275**, 5253–5256 (2000).

72. Loo, T. W. & Clarke, D. M. Location of the rhodamine-binding site in the human multidrug resistance P-glycoprotein. *J. Biol. Chem.* **277**, 44332–44338 (2002).

73. Tsuruo, T., Lida, H., Tsukagoshi, S. & Sakurai, Y. Overcoming of vincristine resistance in P388 leukemia in vivo and in vitro through enhanced cytotoxicity of vincristine and vinblastine by verapamil. *Cancer Res.* **41**, 1967–1972 (1981).

74. Tsuruo T. *et al.* Effects of quinidine and related compounds on cytotoxicity and cellular accumulation of vincristine and adriamycin in drug-resistant tumor cells. *Cancer Res.* 44, 4303–4307 (1984).

75. Twentyman, P. R., Fox, N. E. & White, D. J. Cyclosporin A and its analogues as modifiers of adriamycin and vincristine resistance in a multi-drug resistant human lung cancer cell line. *Br. J. Cancer.* **56**, 55–57 (1987).

76. Höllt, V., Kouba, M., Dietel, M. & Vogt, G. Stereoisomers of calcium antagonists which differ markedly in their potencies as calcium blockers are equally effective in modulating drug transport by P-glycoprotein. *Biochem. Pharmacol.* **43**, 2601–2608 (1992).

77. Hofmann, J. *et al.* Mechanism of action of dexniguldipine-HCl (B8509-035), a new potent modulator of multidrug resistance. *Biochem. Pharmacol.* **49**, 603–609 (1995).

78. Peck, R. A. *et al.* Phase I and pharmacokinetic study of the novel MDR1 and MRP1 inhibitor biricodar administered alone and in combination with doxorubicin. *J. Clin. Oncol.* **19**, 3130–3141 (2001).

79. Nüssler. V. *et al.* In vitro efficacy of known P-glycoprotein modulators compared to droloxifene E and Z: studies on a human T-cell leukemia cell line and their resistant variants. Leuk Lymphoma **31**, 589–597 (1998).

80. Martin, C. *et al.* The molecular interaction of the high affinity reversal agent XR9576 with P-glycoprotein. *Br. J. Pharmacol.* **128**, 403–411 (1999).

81. Hyafil, F., Vergely, C., Du Vignaud, P. & Grand-Perret, T. In vitro and in vivo reversal of multidrug resistance by GF120918, an acridonecarboxamide derivative. *Cancer Res.* **53**, 4595–4602 (1993).

82. Dantzig, A. H. *et al.* Reversal of P-glycoprotein-mediated multidrug resistance by a potent cyclopropyldibenzosuberane modulator, LY335979. *Cancer Res.* **56**, 4171–4179 (1996).

83. van Zuylen, L. *et al.* The orally administered P-glycoprotein inhibitor R101933 does not alter the plasma pharmacokinetics of docetaxel. *Clin. Cancer Res.* **6**, 1365–1371 (2000).

84. Mandal, D., Moitra, K., Ghosh, D., Xia, D. & Dey, S. Evidence for modulatory sites at the lipid-protein interface of the human multidrug transporter p-glycoprotein. *Biochemistry.* **51**, 2852–2866 (2012).

85. Maki, N., Hafkemeyer, P. & Dey, S. Allosteric modulation of human P-glycoprotein: inhibition of transport by preventing substrate translocation and dissociation. *J. Biol. Chem.* **278**, 18132–18139 (2003).

86. Mistry, P. *et al.* In vitro and in vivo reversal of P-glycoprotein-mediated multidrug resistance by a novel potent modulator, XR9576. *Cancer Res.* **61**, 749–758 (2001).

87. Wollmann, J., Richter, M., Molnár, J. & Hilgeroth, A. First insight into the symmetry and flexibility of membrane efflux pump P-glycoprotein by novel bifunctional modulators. *ChemBioChem.* **6**, 1353–1356 (2005).

88. Ma, J. & Biggin, P. C. Substrate versus inhibitor dynamics of P-glycoprotein. *Proteins.* **81**, 1653–1668 (2013).

89. Shao, Y., Ayesh, S. & Stein, W. D. Mutually co-operative interactions between modulators of P-glycoprotein. *Biochim. Biophys. Acta* **1360**, 30–38 (1997).

90. Weidner, L. D. *et al.* Tariquidar is an inhibitor and not a substrate of human and mouse P-glycoprotein. *Drug Metab. Dispos.* **44**, 275–282 (2016).

91. Spoelstra, E. C., Wwsterhoff, H. V., Pinedo, H. M., Dekker, H. & Lankelma, J. The multidrug-resistance-reverser verapamil interferes with cellular P-glycoprotein-mediated pumping of daunorubicin as a non-competing substrate. *Eur. J. Biochem.* **221**, 363–373 (1994).

92. Hung, C-C. *et al.* Functional impact of ABCB1 variants on interactions between P-glycoprotein and methadone. *Plos One.* **8**, e59419 (2013).

93. Caceres, G. *et al.* HG-829 is a potent noncompetitive inhibitor of the ATP-binding cassette multidrug resistance transporter ABCB1. *Cancer Res.* **72**, 4204–4213 (2012).

94. Conseil, G. *et al.* Flavonoids: a class of modulators with bifunctional interactions at vicinal ATP- and steroid-binding sites on mouse P-glycoprotein. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 9831–9836 (1998).

95. Jin, M. S., Oldham, M. L., Zhang, Q. & Chen, J. Crystal structure of the multidrug transporter P-glycoprotein from Caenorhabditis elegans. *Nature* **490**, 566–569 (2012).

96. Oldham, M. L., Khare, D., Quiocho, F.A., Davidson, A. L. & Chen, J. Crystal structure of a catalytic intermediate of the maltose transporter. *Nature.* **450**, 515–521(2007).

97. Loo, T. W., Bartlett, M. C. & Clarke, D. M. Human P-glycoprotein Contains a Greasy Ball-and-Socket Joint at the Second Transmission Interface. *J. Biol. Chem.* **288**, 20326–20333 (2013).

98. Wise, J. G. Catalytic transitions in the human mdr1 P-glycoprotein drug binding sites. *Biochemistry.* **51**, 5125–5141 (2012).

99. Chang, S. Y., Liu, F. F., Dong, X. Y. & Sun, Y. Molecular insight into conformational transmission of human P-glycoprotein. *J. Chem. Phys.* **139**, 225102 (2013).

100. Prajapati, R. & Sangamwar, A. T. Translocation mechanism of P-glycoprotein and conformational changes occurring at drugbinding site: Insights from multi-targeted molecular dynamics. *Biochim. Biophys. Acta.* **1838**, 2882–2898 (2014).

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

102. Currier, S. J. *et al.* Identification of residues in the first cytoplasmic loop of P-glycoprotein involved in the function of chimeric human MDR1-MDR2 transporters. *J. Biol. Chem.* **267**, 25153–25159 (1992).

103. Kwan, T. & Gros, P. Mutational analysis of the P-glycoprotein frst intracellular loop and fanking transmembrane domains. *Biochemistry.* **37**, 3337–3350 (1998).

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

105. Loo, T. W. & Clarke, D. M. Functional consequences of glycine mutations in the predicted cytoplasmic loops of P-glycoprotein. *J. Biol. Chem.* **269**, 7243–7249 (1994).

106. Loo, T. W. & Clarke, D. M. The transmembrane domains of the human multidrug resistance P-glycoprotein are sufficient to mediate drug binding and trafficking to the cell surface. *J. Biol. Chem.* **274**, 24759–24765 (1999).

107. Loo, T. W. & Clarke, D. M. Drug rescue distinguishes between different structural models of human P-glycoprotein. *Biochemistry.* **52**, 7167–7169 (2013).

108. Swartz, D. J., Weber, J. & Urbatsch, I. L. P-glycoprotein is fully active after multiple tryptophan substitutions. *Biochim. Biophys. Acta* 1**828**, 1159–1168 (2013).

109. Raymond, M. *et al.* A Ste6p/P-glycoprotein homologue from the asexual yeast Candida albicans transports the a-factor mating pheromone in Saccharomyces cerevisiae. *Mol. Microbiol.* **27**, 587–598 (1998).

110. Kapoor, K., Bhatnagar, J., Chufan, E. E. & Ambudkar, S. V. Mutations in intracellular loops 1 and 3 lead to misfolding of human P-glycoprotein (ABCB1) that can be rescued by cyclosporine A, which reduces its association with chaperone Hsp70. *J. Biol. Chem.* **288**, 32622–32636 (2013).

111. Loo, T. W. & Clarke, D. M. Locking intracellular helices 2 and 3 together inactivates human P-glycoprotein. *J. Biol. Chem.* **289**, 229–236 (2014).

112. Pajeva, I. K., Hanl, M. & Wiese, M. Protein contacts and ligand binding in the inward-facing model of human P-glycoprotein. *Chem Med Chem.* **8**, 748–762 (2013).

113. Siarheyeva, A., Lopez, J. J. & Glaubitz, C. Localization of multidrug transporter substrates within model membranes. *Biochemistry.* **45**, 6203–6211 (2006).

114. Bahadduri, P. M., Polli, J. E., Swaan, P. W. & Ekins, S. Membrane transporters in drug discovery and development. In Q. Yan (Ed.), *Methods Mol. Biol.* (Vol. **637**). Humana Press (2010).

115. Loo, T. W. & Clarke, D. M. The transmission interfaces contribute asymmetrically to the assembly and activity of human P-glycoprotein. *J. Biol. Chem.* **290**, 16954–16963 (2015).

116. Kaur, V., Garg, T., Rath, G. & Goyal, A. K. Therapeutic potential of nanocarrier for overcoming to P-glycoprotein. *J. Drug Target.* **22**, 859-870 (2014).

117. Ferreira, R. J., Bonito, C. A., Ferreira, M. J. U. & dos Santos, D. J. V. A. About P-glycoprotein: a new druggable domain is emerging from structural data. *WIREs Comput. Mol. Sci.* **7**, e1316 (2017).

118. Isenberg, B., Thole, H., Tümmler, B. & Demmer, A. Identification and localization of three photobinding sites of iodoarylazidoprazosin in hamster P-Glycoprotein. *Eur. J. Biochem.* **268**, 2629–2634 (2001).

119. Demmer, A. *et al*. Localization of the iodomycin binding site in hamster P-glycoprotein. *J. Biol. Chem.* **272**, 20913–20919 (1997).

120. Demmer, A., Andreae, S., Thole, H. & Tümmler, B. Iodomycin and iodipine, a structural analogue of azidopine, bind to a common domain in hamster P-glycoprotein. *Eur. J. Biochem.* **264**, 800–805 (1999).

121. Borchers, C. *et al.* Characterization of the dexniguldipine binding site in the multidrug resistance-related transport protein P-glycoprotein by photoaffinity labeling and mass spectrometry. *Mol. Pharmacol.* **61**, 1366–1376 (2002).

122. Kim, J. Y. *et al.* Identification of an ABCB/P-glycoprotein-specific inhibitor of auxin transport by chemical genomics. *J. Biol. Chem.* **285**, 23309–23317 (2010).

123. Zolnerciks, J. K., Wooding, C. & Linton, K. J. Evidence for a Sav1866-like architecture for the human multidrug transporter P-glycoprotein. *FASEB J.* **21**, 3937-3948 (2007).

124. Brewer, F. K., Follit, C. A., Vogel, P. D. & Wise, J. G. In silico screening for inhibitors of P-glycoprotein that target the nucleotide binding domains. *Mol. Pharmacol.* **86**, 716–726 (2014).

125. Di Pietro, A. *et al*. Modulation by flavonoids of cell multidrug resistance mediated by P-glycoprotein and related ABC transporters. *Cell Mol. Life Sci.* **59**, 307–322 (2002).

126. Ferreira, R. J. *et al.* Nitrogen-containing naringenin derivatives for reversing multidrug resistance in cancer. *Bioorg. Med. Chem.* **28**, 115798 (2020).

127. Palmeira, A. *et al.* Dual inhibitors of P-glycoprotein and tumor cell growth: (Re)discovering thioxanthones. *Biochem. Pharmacol.* **83**, 57–68 (2012).

128. Szabon-Watola, M. I. *et al.* Fluorescent probes of the isoxazole-dihydropyridine scaffold: MDR-1 binding and homology model. *Bioorg. Med. Chem. Lett.* **24**, 117–1121 (2014).

129. Shahraki, O. *et al.* Molecular dynamics simulation and molecular docking studies of 1,4-Dihydropyridines as P-glycoprotein's allosteric inhibitors. *J. Biomol. Struct. Dyn*. **36**, 112–125 (2018).

130. Ooko, E. *et al.* Modulation of P-glycoprotein activity by novel synthetic curcumin derivatives in sensitive and multidrug-resistant T-cell acute lymphoblastic leukemia cell lines. *Toxicol. Appl. Pharmacol.* **305**, 216–233 (2016).

131. Zhang, Z. & Chen, J. Atomic structure of the cystic fibrosis transmembrane conductance regulator. *Cell* **167**, 1586-1597 (2016).

132. Choi, K. H., Chen, C. J., Kriegler, M. & Roninson, I. B. An altered pattern of cross-resistance in multidrug-resistant human cells results from spontaneous mutations in the mdr1 (P-glycoprotein) gene. *Cell.* **53**, 519–229 (1988).

133. Loo, T. W. & Clarke, D. M. Functional consequences of phenylalanine mutations in the predicted transmembrane domain of P-glycoprotein. *J. Biol. Chem.* **268**, 19965–19972 (1993).

134. Harker, W. G., MacKintosh, F. R. & Sikic, B. I. Development and characterization of a human sarcoma cell line, MES-SA, sensitive to multiple drugs. *Cancer Res.* **43**, 4943–4950 (1983).

135. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).

136. Rose, P. W. *et al.* The RCSB protein data bank: integrative view of protein, gene and 3D structural information. *Nucleic Acids Res.* **45**, D271–D281 (2017).

137. Chemical Computing Group Inc., Molecular Operating Environment (MOE) 2016.08 (2016).

138. Labute, P. Protonate3D: Assignment of ionization states and hydrogen coordinates to macromolecular structures. *Proteins.* **75**, 187–205 (2009).

139. Pronk, S. *et al.* GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics.* **29**, 845–854 (2013).

140. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SofwareX* **1–2**, 19–25 (2015).

141. Páll, S., Abraham, M.J., Kutzner, C., Hess, B., Lindahl, E. Tackling Exascale Software Challenges in Molecular Dynamics Simulations with GROMACS. In: Markidis, S., Laure, E. (eds) Solving Software Challenges for Exascale. EASC 2014. Lecture Notes in Computer Science(), vol **8759**. Springer, Cham (2015).

142. Reif, M. M., Hünenberger, P. H. & Oostenbrink, C. New interaction parameters for charged amino acid side chains in the GROMOS Force Field. *J. Chem. Theory Comput.* **8**, 3705–3723 (2012).

143. Schmid, N. *et al.* Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur. Biophys. J.* **40**, 843–856 (2010).

144. Daura, X., Mark, A. E. & Van Gunsteren, W. F. Parametrization of aliphatic CHn united atoms of GROMOS96 force field. *J. Comput. Chem.* **19**, 535–547 (1998).

145. Bonvin, A. M., Mark, A. E. & Van Gunsteren, W. F. The GROMOS96 benchmarks for molecular simulation. *Comput. Phys. Commun.* **128**, 550–557 (2000).

146. Poger, D. & Mark, A. E. On the validation of molecular dynamics simulations of saturated and cis-monounsaturated phosphatidylcholine lipid bilayers: A comparison with experiment. *J. Chem. Teory Comput.* **6**, 325–336 (2010).

147. Poger, D., Van Gunsteren, W. F. & Mark, A. E. A new force feld for simulating phosphatidylcholine bilayers. *J. Comput. Chem.* **31**, 1117–1125 (2010).

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

149. Wolf, M. G., Hoefing, M., Aponte-Santamaría, C., Grubmüller, H. & Groenhof, G. g_membed: Efcient insertion of a membrane protein into an equilibrated lipid bilayer with minimal perturbation. *J. Comput. Chem.* **31**, 2169–2174 (2010).

150. Wolf, M. G., Hoefing, M., Aponte-Santamaría, C., Grubmüller, H. & Groenhof, G. Corrigendum: g_membed: Efcient insertion of a membrane protein into an equilibrated lipid bilayer with minimal perturbation. *J. Comput. Chem.* **37**, 2038–2038 (2016).

151. Van Gunsteren, W. F. & Berendsen, H. J. C. Computer simulation of molecular dynamics: methodology, applications, and perspectives in chemistry. *Angew. Chem. Int. Ed.* **29**, 992–1023 (1990).

152. Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., Hermans, J.. Interaction Models for Water in Relation to Protein Hydration. In: Pullman, B. (eds) Intermolecular Forces. The Jerusalem Symposia on Quantum Chemistry and Biochemistry, vol **14**. Springer, Dordrecht (1981).

153. Leekumjorn, S. & Sum, A. K. Molecular Characterization of Gel and Liquid-Crystalline Structures of Fully Hydrated POPC and POPE Bilayers. J. Phys. Chem. B 2007, **111**, 21, 6026–6033.

154. Davis, I. W. *et al.* MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383 (2007).

155. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).

156. Colovos, C. & Yeates, T. O. Verifcation of protein structures: Patterns of nonbonded atomic interactions. *Protein Sci.* **2**, 1511–1519 (1993).

157. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Tornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291 (1993).

158. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R. & Tornton, J. M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR.* **8**, 477–486 (1996).

159. Schwede, T. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.* **31**, 3381–3385 (2003).

160. Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics.* **22**, 195–201 (2006).

161. Bordoli, L. *et al.* Protein structure homology modeling using SWISS-MODEL workspace. *Nat. Protoc.* **4**, 1–13 (2008).

162. Biasini, M. *et al.* SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **42**, W252–W258 (2014).

163. Lovell, S. C. *et al.* Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins.* **50**, 437–450 (2003).

164. Safa, A. R. *et al.* Molecular basis of preferential resistance to colchicine in multidrug-resistant human cells conferred by Gly-185-Val-185 substitution in P-glycoprotein. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 7225–7229 (1990).

165. Rao, U. S. Mutation of glycine 185 to valine alters the ATPase function of the human P-glycoprotein expressed in Sf9 cells. *J. Biol. Chem.* **270**, 6686–6690 (1995).

166. Ramachandra, M., Ambudkar, S. V., Gottesman, M. M., Pastan, I. & Hrycyna, C. A. Functional characterization of a glycine 185-tovaline substitution in human P-glycoprotein by using a vaccinia-based transient expression system. *Mol. Biol. Cell.* **7**, 1485–1498 (1996).

167. Loo, T. W. & Clarke, D. M. Rapid purifcation of human P-glycoprotein mutants expressed transiently in HEK 293 cells by nickel-chelate chromatography and

characterization of their drug-stimulated ATPase activities. *J. Biol. Chem.* **270**, 21449–21452 (1995).

168. Omote, H., Figler, R. A., Polar, M. K. & Al-Shawi, M. K. Improved energy coupling of human P-glycoprotein by the glycine 185 to valine mutation. *Biochemistry.* **43**, 3917–3928 (2004).

169. Harker, W. G. & Sikic, B. I. Multidrug (pleiotropic) resistance in doxorubicin-selected variants of the human sarcoma cell line MES-SA. *Cancer Res.* **45**, 4091–4096 (1985).

170. Chen, G. *et al.* Multidrug-resistant human sarcoma cells with a mutant P-glycoprotein, altered phenotype, and resistance to cyclosporins. *J. Biol. Chem.* **272**, 5974–5982 (1997).

171. Chen, G. K., Lacayo, N. J., Durán, G. E., Cohen, D. & Sikic, B. I. Loss of cyclosporin and azidopine binding are associated with altered ATPase activity by a mutant P-glycoprotein with deleted phe(335). *Mol. Pharmacol.* **57**, 769–777 (2000).

172. Brady, G. P. & Stouten, P. F. Fast prediction and visualization of protein binding pockets with PASS. *J. Comput. Aided Mol. Des.* **14**, 383–401 (2000).

173. Kohlbacher, O. & Lenhof, H. P. BALL–rapid sofware prototyping in computational molecular biology. Biochemicals Algorithms Library. *Bioinformatics.* **16**, 815–824 (2000).

174. van der Spoel, D., van Maaren, P. J., Larsson, P. & Tîmneanu, N. Termodynamics of hydrogen bonding in hydrophilic and hydrophobic media. *J. Phys. Chem. B.* **110**, 4393–4398 (2006).

175. Blau, C. & Grubmuller, H. g_contacts: Fast contact search in bio-molecular ensemble data. *Comput. Phys. Commun.* **184**, 2856–2859 (2013).

176. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).

177. Tamai, I. & Safa, A. R. Azidopine noncompetitively interacts with vinblastine and cyclosporin A binding to P-glycoprotein in multidrug resistant cells. *J. Biol. Chem.* **266**, 16796–16800 (1991).

178. Loo, T. W., Bartlett, M. C. & Clarke, D. M. Simultaneous binding of two different drugs in the binding pocket of the human multidrug resistance P-glycoprotein. *J. Biol. Chem.* **278**, 39706–39710 (2003).

179. Loo, T. W., Bartlett, M. C. & Clarke, D. M. ATP hydrolysis promotes interactions between the extracellular ends of transmembrane segments 1 and 11 of human multidrug resistance P-glycoprotein. *Biochemistry.* **44**, 10250–10258 (2005).

180. Loo, T. W. & Clarke, D. M. Recent progress in understanding the mechanism of P-glycoprotein-mediated drug efflux. *J. Membr. Biol.* **206**, 173–185 (2005).

181. Schüttelkopf, A. W. & van Aalten, D. M. F. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 1355–1363 (2004).

182. Lemkul, J. A., Allen, W. J. & Bevan, D. R. Practical considerations for building GROMOS-compatible small-molecule topologies. *J. Chem. Inf. Model.* **50**, 2221–2235 (2010).

183. Wang, J., Wang, W., Kollman, P. A. & Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **25**, 247–260 (2006).

184. Kumari, R., Kumar, R. & Lynn, A. *g_mmpbsa* —A GROMACS tool for high-throughput MM-PBSA calculations. J. Chem. Inf. Model. **54**, 1951–1962 (2014).

185. Ferreira, R. J., Ferreira, M-J. U. & dos Santos, D. J. V. A. Do adsorbed drugs onto P-glycoprotein influence its efflux capability? Phys. Chem. Chem. Phys. **17**, 22023–22034 (2015).

186. Bruggemann, E. P., Currier, S. J., Gottesman, M. M. & Pastan I. Characterization of the azidopine and vinblastine binding site of P-glycoprotein. *J. Biol. Chem.* **267**, 21020-21026 (1992).

187. *MarvinSketch* was used for drawing, displaying and characterizing chemical structures, substructures and reactions, Marvin version 19.18.0, ChemAxon (https://www.chemaxon.com)

188. Chemical Computing Group Inc., Molecular Operating Environment (MOE) 2019.01 (2019).

189. Halgren, T. A. MMFF VI. MMFF94s option for energy minimization studies. *J. Comput. Chem.* **20**, 720–729 (1999).

190. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 14101–14107 (2007).

191. Nosé, S. & Klein, M. L. Constant pressure molecular dynamics for molecular systems. *Mol. Phys.* **50**, 1055–1076 (1983).

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

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

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

195. Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. GROMACS 4: Algorithms for Highly Efcient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Teory Comput.* **4**, 435–447 (2008).

196. Miyamoto, S. & Kollman, P. A. SETTLE: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **13**, 952–962 (1992).

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

198. Essmann, U. *et al.* A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).

199. Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efcient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).

200. Morris, G. *et al.* AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–91 (2009).

201. Ferreira, R.J. *et al.* Optimizing the flavanone core toward new selective nitrogen-containing modulators of ABC transporters. *Future Med. Chem.* **10**, 725–741 (2018).

202. Durães, F. *et al.* Antimicrobial activity of a library of thioxanthones and their potential as efflux pump inhibitors. *Pharmaceuticals.* **14**, 572 (2021).

203. Durrant, J. D. & McCammon, J. A. BINANA: a novel algorithm for ligand-binding characterization. *J. Mol. Graph. Model.* **29**, 888–893 (2011).

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

205. Dongping, M. & Cali J. J. Identify P-glycoprotein substrates and inhibitors with the rapid HTS P-gp-Glo™ Assay system. *Promega Notes* **96**, 11–14 (2007).

206. Ferreira R. J. *et al.* Terpenoids from Euphorbia pedroi as multidrug-resistance reversers. *J. Nat. Prod.* **81**, 2032–2040 (2018).

207. Litman, T., Skovsgaard, T. & Stein, W. D. Pumping of drugs by P-glycoprotein: a two-step process? *J. Pharmacol. Exp. Ther.* **307**, 846–853 (2003).

208. Borges-Walmsley, A. R., Mckeegan, M. I. & Walmsley, K. S. Structure and function of efux pumps that confer resistance to drugs. *Biochem. J.* **376**, 313–338 (2003).

209. Gottesman, M. M., Lavi, O., Hall, M. D. & Gillet, J.-P. Toward a Better Understanding of the Complexity of Cancer Drug Resistance. *Annu. Rev. Pharmacol. Toxicol.* **56**, 85–102 (2016).

210. Ferreira, R. J., Ferreira, M. J. U. & dos Santos, D. J. V. A. Reversing cancer multidrug resistance: Insights into the efux by ABC transports from *in silico* studies. *WIRES Comput. Mol. Sci.* **5**, 27–55 (2015).

211. Sharom, F. J. The P-glycoprotein efflux pump: does it transport drug hows? *J. Membr. Biol.* **160**, 161–175 (1997).

212. Al-Shawi, M. K. & Omote, H. The remarkable transport mechanism of P-glycoprotein: a multidrug transporter. *J. Bioenerg. Biomembr.* **37**, 489–496 (2005).

213. Julien, M. & Gros, P. Nucleotide-induced conformational changes in P-glycoprotein and in nucleotide binding site mutants monitored by trypsin sensitivity. *Biochemistry.* **39**, 4559–4568 (2000).

214. Loo, T. W. & Clarke, D. M. Determining the Dimensions of the Drug-binding Domain of Human P-glycoprotein Using Tiol Cross-linking Compounds as Molecular Rulers. *J. Biol. Chem.* **276**, 36877–36880 (2001).

215. Loo, T. W., Bartlett, M. C. & Clarke, D. M. Human P-glycoprotein is active when the two halves are clamped together in the closed conformation. *Biochem. Biophys. Res. Commun.* **395**, 436–440 (2010).

216. Loo, T. W. & Clarke, D. M. Identifcation of the distance between the homologous halves of P-glycoprotein tat triggers the high/low ATPase activity switch. *J. Biol. Chem.* **289**, 8484–8492 (2014).

217. Loo, T. W. & Clarke, D. M. Drugs Modulate Interactions between the First Nucleotide-Binding Domain and the Fourth Cytoplasmic Loop of Human P-Glycoprotein. *Biochemistry.* **55**, 2817–2820 (2016).

218. Loo, T. W. & Clarke, D. M. P-glycoprotein ATPase activity requires lipids to activate a switch at the frst transmission interface. *Biochem. Biophys. Res. Commun.* **472**, 379–383 (2016).

219. Loo, T. W., Bartlett, M. C., Detty, M. R. & Clarke, D. M. 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 (2012).

220. Shapiro, A. B. & Ling, V. Extraction of Hoechst 33342 from the cytoplasmic leafet of the plasma membrane by P-glycoprotein. *Eur. J. Biochem.* **250**, 122–129 (1997).

221. Shapiro, A. B. & Ling, V. Efect of quercetin on Hoechst 33342 transport by purifed and reconstituted P-glycoprotein. *Biochem. Pharmacol.* **53**, 587–596 (1997).

222. Dolghih, E., Bryant, C., Renslo, A. R. & Jacobson, M. P. Predicting Binding to P-Glycoprotein by Flexible Receptor Docking. *Plos Comput. Biol.* **7**, e1002083 (2011).

223. Ferreira, R. J., Ferreira, M. J. U. & dos Santos, D. J. V. A. Do drugs have access to the P-lycoprotein drug-binding pocket through gates? *J. Chem. Teory Comput.* **11**, 4525–4529 (2015).

224. Mukhametov, A. & Raevsky, O. A. On the mechanism of substrate/non-substrate recognition by P -glycoprotein. *J. Mol. Graph. Model.* **71**, 227–232 (2017).

225. Wang, Z. & Liao, J.-L. Probing Structural Determinants of ATP-Binding Cassette Exporter Conformational Transition Using Coarse-Grained Molecular Dynamics. *J. Phys. Chem. B.* **119**, 1295–1301 (2015).

226. Mittra, R. *et al.* Location of contact residues in pharmacologically distinct drug binding sites on P-glycoprotein. *Biochem. Pharmacol.* **123**, 19–28 (2017).

227. Wen, P.-C., Verhalen, B., Wilkens, S., Mchaourab, H. S. & Tajkhorshid, E. On the origin of large fexibility of P-glycoprotein in the inward-facing state. *J. Biol. Chem.* **288**, 19211–19220 (2013).

228. Clay, A. T., Lu, P. & Sharom, F. J. Interaction of the P-glycoprotein multidrug transporter with sterols. *Biochemistry.* **54**, 6586–6597 (2015).

229. Qin, L., Hiser, C., Mulichak, A., Garavito, R. M. & Ferguson-Miller, S. Identifcation of conserved lipid/detergent-binding sites in a high-resolution structure of the membrane protein cytochrome c oxidase. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 16117–16122 (2006).

230. Benkert, P., Künzli, M. & Schwede, T. QMEAN server for protein model quality estimation. *Nucleic Acids Res.* **37**, W510–W514 (2009).

231. Domicevica, L. & Biggin, P. C. Homology modelling of human P-glycoprotein. *Biochem. Soc. Trans.* **43**, 952–958 (2015).

232. Polli, J. W. *et al.* Rational use of in vitro P-glycoprotein assays in drug discovery. *J. Pharmacol. Exp. Ther.* **299**, 620–628 (2001).

233. Rautio, J. *et al.* In vitro p-glycoprotein inhibition assays for assessment of clinical drug interaction potential of new drug candidates: a recommendation for probe substrates. *Drug Metab. Dispos.* **34**, 786–792 (2006).

234. Studer, G., Biasini, M. & Schwede, T. Assessing the local structural quality of transmembrane protein models using statistical potentials (QMEANBrane). *Bioinformatics.* **30**, 505–511 (2014).

235. Li, J., Jaimes, K. F. & Aller, S. G. Refned structures of mouse P-glycoprotein. *Protein Sci.* **23**, 34–46 (2014).

236. Martin, C. *et al.* Communication between multiple drug binding sites on P-glycoprotein. *Mol. Pharmacol.* **58**, 624–632 (2000).

237. Waghray, D. & Zhang, Q. Inhibit or Evade Multidrug Resistance P-glycoprotein in Cancer Treatment. *J. Med. Chem.* **61**, 5108–5121 (2018).

238. Gannon, M. K. *et al.* Rhodamine Inhibitors of P-Glycoprotein: an amide/thioamide "switch" for ATPase activity. J. Med. Chem. 52, 3328–3341 (2009).

239. Sekhar, A., Vallurupalli, P. & Kay, L. E. Defning a length scale for millisecond-timescale protein conformational exchange. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 11391–11396 (2013).

240. Bonito C. A. *et al*. Theoretical insights on helix repacking as the origin of P-glycoprotein promiscuity. *Scientific Reports.* **10**, 9823 (2020).

241. Clouser, A. F. & Atkins, W. M. Long Range Communication between the drug-binding sites and nucleotide binding domains of the efflux transporter ABCB1. *Biochemistry.* **61**, 730-740 (2022).

242. Kapoor, K., Pant, S. & Tajkhorshid, E. Active participation of membrane lipids in inhibition of multidrug transporter P-glycoprotein. *Chem. Sci.* **12**, 6293–6306 (2021).

243. Altenberg, G. A. *et al.* Unidirectional fluxes of rhodamine 123 in multidrug-resistant cells: evidence against direct drug extrusion from the plasma membrane. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 4654–4657(1994).

244. Clouser, A. F., Alam, Y. H. & Atkins, W. M. Cholesterol asymmetrically modulates the conformational ensemble of the nucleotide-binding domains of P-glycoprotein in lipid nanodiscs. *Biochemistry.* **60**, 85–94 (2021).

245. Ferreira, R. J., Bonito, C. A., Cordeiro, M. N. D. S., Ferreira, M. J. U. & dos Santos, D. J. V. A. Structure-function relationships in ABCG2: Insights from molecular dynamics simulations and molecular docking studies. *Sci. Rep.* **7**, 1–17 (2017).

246. Shaffer, B. C. *et al.* Drug resistance: still a daunting challenge to the successful treatment of AML. *Drug Resist. Updat.* **15**, 62-69 (2012).

247. Boumendjel, A. *et al*. 4-Hydroxy-6-methoxyaurones with high-affinity binding to cytosolic domain of P-glycoprotein. *Chem. Pharm. Bull.* **50**, 854–856 (2002).

248. Zolnerciks, J. K. *et al.* The Q loops of the human multidrug resistance transporter ABCB1 are necessary to couple drug binding to the ATP catalytic cycle. *FASEB J.* **28**, 4335-4346 (2014).

APPENDIX

SCIENTIFIC REPORTS
natureresearch

Check for updates

OPEN

Theoretical insights on helix repacking as the origin of P-glycoprotein promiscuity

Cátia A. Bonito [1], Ricardo J. Ferreira[2], Maria-José. U. Ferreira[3], Jean-Pierre Gillet[4], M. Natália D. S. Cordeiro [1] & Daniel J. V. A. dos Santos[1,3] ✉

P-glycoprotein (P-gp, ABCB1) overexpression is, currently, one of the most important multidrug resistance (MDR) mechanisms in tumor cells. Thus, modulating drug efflux by P-gp has become one of the most promising approaches to overcome MDR in cancer. Yet, more insights on the molecular basis of drug specificity and efflux-related signal transmission mechanism between the transmembrane domains (TMDs) and the nucleotide binding domains (NBDs) are needed to develop molecules with higher selectivity and efficacy. Starting from a murine P-gp crystallographic structure at the inward-facing conformation (PDB ID: 4Q9H), we evaluated the structural quality of the herein generated human P-gp homology model. This initial human P-gp model, in the presence of the "linker" and inserted in a suitable lipid bilayer, was refined through molecular dynamics simulations and thoroughly validated. The best human P-gp model was further used to study the effect of four single-point mutations located at the TMDs, experimentally related with changes in substrate specificity and drug-stimulated ATPase activity. Remarkably, each P-gp mutation is able to induce transmembrane α-helices (TMHs) repacking, affecting the drug-binding pocket volume and the drug-binding sites properties (e.g. volume, shape and polarity) finally compromising drug binding at the substrate binding sites. Furthermore, intracellular coupling helices (ICH) also play an important role since changes in the TMHs rearrangement are shown to have an impact in residue interactions at the ICH-NBD interfaces, suggesting that identified TMHs repacking affect TMD-NBD contacts and interfere with signal transmission from the TMDs to the NBDs.

Multidrug resistance (MDR) to anticancer drugs is, at the moment, a major contributor to chemotherapy failure[1]. In cancer, one of the most significant MDR mechanisms results from the overexpression of P-glycoprotein, a membrane efflux pump (P-gp, ABCB1)[2]. Thus, a deeper understanding on P-gp substrate specificity and efflux-related signal transmission mechanism remains crucial for the development of more potent and selective compounds able to modulate drug efflux[3]. P-glycoprotein exports a broad range of structurally unrelated compounds through an ATP-dependent mechanism[4]. P-gp is organized in two homologous functional units (N- and C-terminal halves) with a pseudo-2-fold symmetry. Each halve comprises one transmembrane domain (TMD), formed by six transmembrane α-helices (TMHs), and one cytoplasmic nucleotide-binding domain (NBD). Both N- and C-terminal halves are connected by a small peptide sequence (the "linker"; residues 627–688)[5,6]. The TMHs are directly linked to the respective NBD by the intracellular loops, through the functional TM helices 6 (NBD1) and 12 (NBD2) and non-covalently by short intracellular coupling helices (ICHs), located between the structural TMHs 2/3 (ICH1-NBD1), 4/5 (ICH2-NBD2), 8/9 (ICH3-NBD2) and 10/11 (ICH4-NBD1)[7,8] (Fig. 1). These ICHs were found to be important for the maturation and folding of the P-gp transporter, being also involved in the signal transmission pathway between the TMDs and NBDs[9,10].

The drug-binding pocket (DBP) is a large cavity formed by the TMHs of both N- and C-terminal P-gp halves and is capable of recognizing and accommodating several structurally distinct substrates. Each NBD contains the

[1]LAQV@REQUIMTE, Department of Chemistry and Biochemistry, Faculty of Sciences, University of Porto, Rua do Campo Alegre, 4169-007, Porto, Portugal. [2]Science for Life Laboratory, Department of Cell and Molecular Biology, Uppsala University, 75124, Uppsala, Sweden. [3]Research Institute for Medicines (iMed.ULisboa), Faculty of Pharmacy, Universidade de Lisboa, Av. Prof. Gama Pinto, 1649-003, Lisboa, Portugal. [4]Laboratory of Molecular Cancer Biology, Molecular Physiology Research Unit-URPhyM, Namur Research Institute for Life Sciences (NARILIS), Faculty of Medicine, University of Namur, B-5000, Namur, Belgium. ✉e-mail: ddsantos@fc.up.pt
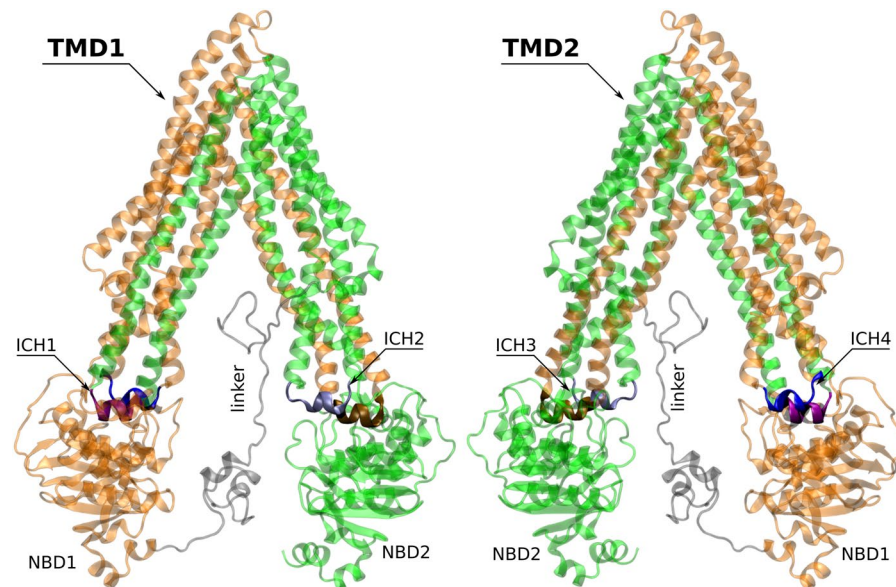
Figure 1. Structural representation of the human P-glycoprotein in an inward-facing conformation. The 12 transmembrane α-helices (TMHs) are divided into two transmembrane domains (TMD1; orange and TMD2; green), being physically linked to the respective nucleotide-binding domain (NBD) by coils bridging TMH6/NBD1 and TMH12/NBD2 and also by non-covalent interactions involving short intracellular coupling helices: ICH1 (purple)/ICH4 (blue) with NBD1 and ICH2 (silver)/ICH3 (brown) with NBD2. The DBP is a large cavity between both TMDs. Figures were created with MOE[133] from the final human P-gp homology model.

catalytic site for ATP binding and hydrolysis[11,12]. Several experimental[5,10,12–36] and computational[6,37–47] studies were performed to better understand the details of drug specificity and efflux. However, the information gathered from these experiments are often scarcely related and the mechanisms of drug recognition/specificity and efflux are still unsolved. Therefore, for additional insights on P-gp efflux mechanism, the study of P-gp variants that are experimentally related with altered drug-resistance phenotypes, and changes in the ATPase activity is an interesting approach.

The first mutation in the human P-gp was identified in the colchicine-selected multidrug-resistance cell line (KB-C1)[48] and comprises a substitution of a glycine by a valine residue in position 185 (G185V; TMH3). This mutation is reported to confer increased resistance to colchicine, etoposide, doxorubicin and puromycin while decreasing the resistance to vinblastine, vincristine, taxol and actinomycin D. The same study reported, unexpectedly, a 3.6-fold decrease in colchicine binding and a 3.8–5.5-fold increase in vinblastine binding when compared with the WT transporter, leading the authors to conclude that G185V mostly affects the dissociation of drugs from P-gp rather than its initial association[49]. Later studies additionally concluded that the basal ATPase activity and the degree of activation by substrates was also increased in G185V, with only minor changes on nucleotide binding[50–52]. As it was also reported that G185V mutant is able to transport colchicine and etoposide in an energetically more efficient way, the residue at position 185 was identified as pivotal for transmitting conformational changes between the catalytic sites and colchicine-binding site[53].

Similar alterations were observed in another glycine variant (G830V; TMH9), obtained by site-directed mutagenesis, in the presence of verapamil and colchicine[52,54]. Herein, increased resistance to colchicine (3.3-fold) and decreased resistance to actinomycin D (0.29-fold) was reported but no significant change on vinblastine resistance was observed when compared with the WT protein[52]. Its verapamil-stimulated ATPase activity was found to be only slightly increased (1.4–1.7-fold), but no change in the vinblastine-stimulated ATPase activity was reported. Nonetheless, the pattern of drug-stimulated ATPase correlated well with the relative drug-resistance profiles in transfected cells[54]. No data on basal ATPase activity was provided.

Likewise, mutations involving the phenylalanine residue at position 978 were also identified to play an important functional role in P-gp. By mutating F978 to an alanine (F978A) a significant alteration of the drug resistant pattern was observed, conferring little or no resistant to either colchicine or doxorubicin but not changing the resistance to vinblastine or actinomycin D. This was further suggested to be due to a defect in the mutant protein, being unable to transport colchicine and having a reduced capability to transport vinblastine, even with an increased expression of the transporter. Quite interestingly, when testing other substitutions (F978S, F978L or F978Y), only the latter restored similar drug-resistance profiles for all tested molecules[55], but nonetheless azidopine labeling was indistinguishable from WT protein. Interestingly, including the previously characterized G185V together with F978A/S was also unsuccessful in fully restoring the ability of conferring resistance to colchicine[55]. In addition, little or no drug-stimulated ATPase activity by colchicine or verapamil was reported, being suggested that F978A mutants have either a decreased affinity for substrates and/or an impairment in coupling drug binding to ATPase activity[54]. Recent studies, using a F978C mutant, additionally showed complete absence of stimulation of ATP hydrolysis by several drugs, proposing that F978 residue is part of a common translocation route, crucial

for the propagation of conformational changes following ATP hydrolysis and important for the translocation process of high-affinity substrates[56]. Again, no data was reported on basal ATPase activity for this mutant.

Finally, the deletion of a phenylalanine residue at position 335 (ΔF335; TMH6) was reported in a multidrug-resistant variant cell line (DxP), by co-selection with doxorubicin and the cyclosporine D analogue, valspodar (PSC833), a potent P-gp modulator[57–59]. The transfected cells harboring the ΔF335 P-gp variant showed to be resistant to many chemotherapeutic drugs such as doxorubicin, daunorubicin, etoposide and paclitaxel. On the other hand, this variant conferred little resistance to vincristine, vinblastine and actinomycin D as well as a decreased ability to bind or transport cyclosporine, valspodar, vinblastine, actinomycin D and rhodamine-123, suggesting that the region surrounding the F335 residue is an important binding site for these compounds[59,60]. An enhancement of photoaffinity binding by [^{125}I]-iodoarylazidoprazosin in the presence of verapamil or PSC833[59] and a decrease of [^{3}H]-cyclosporine binding[60] was also referred. Moreover, by deleting the F335 residue a 2-fold increase in the basal ATPase was observed but, when in the presence of drugs, only verapamil-stimulated ATPase activity occured[60]. Interestingly, this mutant also presented a substantial decrease of 8-[α-^{32}P]azido-ATP labeling when compared with the WT P-gp, which lead the authors to suggest that the ATPase activity of the mutant protein may depend on the helix conformation defined by F335[60].

In this work, a human P-gp homology model in the apo inward-facing state conformation was generated based on the most recent murine P-gp crystallographic structure available in the beginning of our study (PDB IDs: 4Q9H)[61]. The P-gp model was refined through molecular dynamics (MD) simulations and validated using several approaches. The final human P-gp model was further used to understand the possible structural impact of the mutations described above on P-gp architecture/organization.

Material and Methods

Human P-gp homology modeling. The FASTA sequence of human P-gp was obtained from the Universal Protein Resource (UNIPROT) under the code P08183 (www.uniprot.org). The murine P-gp crystallographic structure (PDB IDs: 4Q9H), used as template, was retrieved from the Protein Data Bank (PDB; www.rcsb.org)[62,63]. Both the murine P-gp crystallographic structure and the FASTA sequence human P-gp were loaded into the MOE software and aligned by their sequence. The "linker" secondary structure, missing in all P-gp crystallographic structures so far, was obtained from a previously equilibrated murine P-gp[39] and used as template for modeling the respective sequence (A627-A688). Herein, the insertion of the "linker" in the novel structures was achieved by aligning the human P-gp homology model with the crystallographic template, followed by an "override" of the gap in the considered structure in which the murine P-gp linker was used as template to obtain the human homologue. Afterwards, 25 mainchain models sampling 25 sidechain orientations were performed, producing 625 models for the human P-gp structure in MOE (force field parameters used by default). The model with lowest total potential energy was selected, protonated using the Protonate 3D module[64] and exported as PDB file to be used by GROMACS version 5.0.7[65–67]. This way, an initial human P-gp homology model in an inward-facing conformation incorporating the "linker" obtained from a previously equilibrated murine P-gp[39], was obtained.

Construction of the protein membrane system. The topology of the human P-gp model was generated according with the GROMOS96 54a7 force field. A previous equilibrated 1-palmitoyl-2-oleoyl-phosphatidylcholine (POPC) membrane patch[39] was used with the lipid parameterization by Poger *et al.*[68,69]. The human P-gp homology model was inserted into a lipid bilayer (longer P-gp axis perpendicular to the *xy* membrane plane) to match the hydrophobic thickness of TMDs and membrane. The relative position of the membrane was obtained from the Orientations of Proteins in Membranes (OPM) database[70] (http://opm.phar.umich.edu) and protein insertion was achieved through the *g_membed*[71,72] module in GROMACS.

The protein embedded into the lipid bilayer was centered in a simulation box with dimensions *xyz* of 12.76 × 12.76 × 16.50 nm^3 and periodic boundary conditions (PBC). Finally, the system was solvated and neutralized with an adequate number of water molecules and counterions using other GROMACS' modules.

Molecular Dynamics: equilibration and production run. Firstly, an energy minimization run comprising the whole system was applied using the steepest descent method. Then, the temperature of the membrane system (303 K) was equilibrated for 10 ps in the *NVT* ensemble, spatially restraining all protein's heavy atoms. Following, the POPC lipid bilayer was allowed to correctly adjust to the protein interface through a 20 ns *NpT* run, still keeping the protein's heavy atoms restrained. Finally, three sequential 500 ps *NpT* runs were performed to progressively remove the protein's heavy atoms spatial restriction (mainchain, backbone and alpha-carbons, respectively). This system was the starting point for a 200 ns fully unrestrained *NpT* production run (Fig. S1, Supporting Information).

Model quality assessment. The stability of the P-gp model was monitored along the MD run through the evolution of the root mean square deviation (RMSD) of the Cα atoms, visual inspection and the MolProbity[73,74] evaluation server. After 200 ns of simulation time, more exhaustive evaluations were performed, through additional servers namely ERRAT[75], PROCHECK[76,77] and SwissModel Structure assessment tool[78–81]. Moreover, the stability and quality of the human P-gp model were also assessed considering the Ramachandran plot[82] and by checking for correlations between molecular docking and experimental data. The evaluation of the recently published human cryo-EM P-gp structure was also performed, for comparison purposes.

Construction of the human P-gp mutated structures and systems. From the final refined human P-gp homology model, four human P-gp variants (G185V, G830V, F978A and ΔF335) experimentally linked with changes in efflux and substrate specificity were built using MOE. Each P-gp variant was then embedded into a POPC membrane, water solvated and charge neutralized as described above. Energy minimization runs comprising the whole system were applied followed by a 10 ps *NVT* run at 303 K by spatially restraining all protein's

heavy atoms. Fully unrestrained *NpT* runs followed for 100 ns (Fig. S1). After 50 ns of simulation time, two system replicates were obtained for all P-gp variant systems, each one simulated for another 50 ns by randomly generating initial velocities, assigned from the correct temperature dependent Maxwell-Boltzman distribution, and starting with the final configuration obtained at the end of the first 50 ns. This way, for each P-gp variant, three replica systems were therefore simulated in a total of 200 ns of simulation time.

Simulation parameters.   All *NVT* equilibration runs were performed at 303 K using the Velocity-rescale (V-rescale)[83] thermostat. The Nosé-Hoover[84,85] thermostat and the Parrinello-Rahman[86] barostat for temperature (303 K) and pressure (1 bar), respectively, were applied in all *NpT* runs. Due to the presence of membranes, pressure equilibration was achieved through a semi-isotropic pressure coupling, with the systems' compressibility set to 4.5×10^{-5} bar^{-1}. All bond lengths were constrained using the LINCS[87,88] and SETTLE[89] (for water molecules) algorithms. The Particle Mesh Ewald (PME) with cubic interpolation[90,91] was employed, with a cut-off radius of 12 Å for both electrostatic and van der Waals interactions and an FFT grid spacing of 0.16 for long range electrostatics. Group-based and Verlet[92] cut-off schemes were applied for the calculation of non-bonded interactions on CPU or GPU, respectively.

Structural analysis of the human P-gp variants.   To evaluate the impact of each mutation in the volume of the internal DBP (only considering the transmembrane helical bundle buried within the membrane) and to allow a fast comparison between WT and variants, the DBP was estimated as the sum of the volumes of all water molecules found inside this cavity through in-house python scripts. To assess the effect of the mutations on TMDs rearrangement, the TM bundle was analyzed through the *g_bundle* module available in GROMACS. The total number of contacts between the ICHs residues and the respective NBD were calculated using the *g_hbond*[93] module and the contact frequencies were estimated by the *g_contacts*[94] module. Furthermore, the EPOS^{BP}[95,96] software (default parameters) was used to characterize the DBSs found within the DBP of the human WT P-gp model and variants. The top-ranked docking poses of each molecule in each zone were then overlapped with the cavity search results identifying, this way, lining atoms (within a distance of 5 Å from the pocket probes) and calculating mean pocket volumes and polarities (ratio of the sum of N, O, and S atoms to the sum of N, O, S, and C atoms). Visual inspections were performed with VMD[97] and MOE software. All analysis described above were performed using the last 50 ns of each simulation.

Docking studies.   Molecular docking was performed using the final human WT P-gp model and the generated variants (G185V, G830V, F987A and ΔF335). The chosen databases comprised P-gp substrates ($N = 33$), probes ($N = 7$) and modulators ($N = 19$), previously used in docking studies with our refined murine P-gp structure[39]. The ligands binding location was defined by a docking box comprising the whole internal cavity identified by Aller *et al.*[5], with dimensions *xyz* of $32.25 \times 26.25 \times 37.50$ Å^3, and centered at the DBP (*xy* corresponds to the membrane plane). Due to the large search volume (over 30.000 Å^3), Vina's 'exhaustiveness' parameter was manually set to 50 and twenty docking poses were generated. Visual inspection of the best ranked docking poses was made in MOE to allow the identification of individual docking zones.

Results and Discussion

Human P-gp homology model development.   Considering the high sequence identity and similarity with the human P-gp efflux pump (87% and 94%, respectively)[98], a human P-gp homology model was obtained, using the murine P-gp crystallographic structure of 2015 (PDB ID: 4Q9H) as template.

The murine P-gp crystallographic structure was chosen based on: i) it was the most recently published murine P-gp crystallographic structure in the beginning of our study, ii) it shows improvements in the resolution of several TMHs, ICH1 and some extracellular loops in respect to previous structures, stressing the quality of the starting template to the development of a reliable human P-gp model, and iii) oppositely to the recently published human P-gp structure[99], it was obtained without any ligands or antibody complexes that could have any influence on the native arrangement of the transmembrane helical domains.

After 14 ns of simulation time, and unlike the refined murine P-gp structure[39], a shift of the linker's upper loop downwards was observed in our human P-gp model (v1 model). Thus, to assess the correct position of the "linker" in the human P-pg model, two snapshots retrieved at 8 ns (v2 model) and 14 ns (v3 model) of simulation time were the starting points for two additional MD runs (Fig. S1).

At the end of 200 ns MD simulations, a shift of the linker's upper loop downwards was also observed in both v2 and v3 models confirming the possibility that this new position acquired by the "linker" structure is favored in our model. This result also demonstrates the high flexibility of the "linker" region offering a plausible explanation for its absence in all crystallographic structures so far. Although an adjustment of the NBDs was observed in all P-gp models, both v1 and v2 models also revealed distortions in the "linker" secondary structure, namely through the formation of an α-helix in its middle coil. Additionally, a kink in the TMHs 6, 10 and 12 was found in both models compromising the DBP bottom and the portals. Together with the low scores obtained in the evaluation servers (Table S1), the identification of structural alterations were the underlying reasons leading to the rejection of both v1 and v2 models. However, a stable human P-gp model was obtained after 200 ns of simulation time, although with a POPC lipid molecule located in the portal 4/6 (v3a model), in agreement with the computational findings of Tajkhorshid and colleagues[100]. Interestingly, no secondary structure in the middle coil of the linker structure was found in this model. As it is experimentally known that some lipids play an important role in the stability and function of protein membranes including P-gp[101–103], it was important to assess the impact of the lipid molecule in the protein stability. Therefore, two new system snapshots were retrieved from the v3a model at 200 ns of simulation time as starting points for two additional 200 ns MD runs (in a total of 400 ns of each system), where the POPC molecule was removed from the system (v3b) or moved into the lipid bilayer (v3c) (Fig. S1).
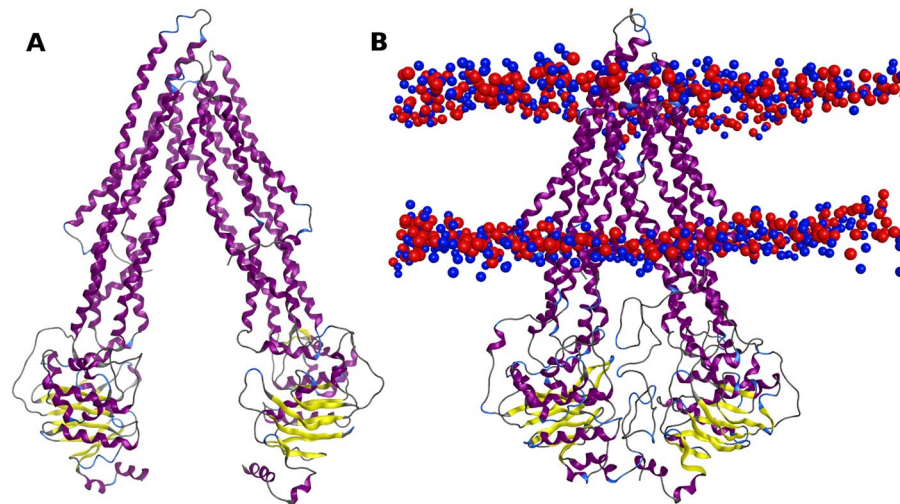
Figure 2. (**A**) Representation of the murine P-gp crystallographic structure (PDB ID: 4Q9H) used as template vs (**B**) human WT P-gp model (v3c) in the presence of the "linker" and a POPC membrane. The lipid bilayer boundaries are represented through the phosphate (red) and nitrogen (blue) atoms of the lipid headgroups.

Human P-gp homology model validation. Although lower scores were observed in the initial human P-gp model obtained with MOE, significant structural improvements were achieved in the refined models, after molecular dynamics (MD) simulations (Table S1). Although the SwissModel online server still utilizes an algorithm that is not adequate for evaluating membrane proteins[104,105], all our structures scored above −5.0, a solid indicator on the good quality of the herein developed models[104].

Another important tool to assess the structural quality of a given structure is through the analysis of the Ramachandran plots of each v3 P-gp model (Fig. S2). Substantial improvements were observed in the refined P-gp models with only 16 (v3a), and 19 (v3b) and 18 (v3c) outliers, respectively versus 34 outliers found in the initial P-gp homology model. The outliers were found mainly in coils, being the only exceptions A348 residue, located at the TMH6 (v3a and v3b), and the F359 (TMH11) and N81 (TMH1) residues in the v3c model.

The RMSD evolution of the P-gp models during MD simulations (Fig. S3) showed that any considerable changes in protein's conformation occurred during the first 200 ns of simulation time. Moreover, a comparison between the refined v3 models showed that all models were stable during the last 200 ns of MD runs with the POPC molecule initially found in the portal 4/6 (v3a model) not affecting the protein stability. However, visual inspection of the three P-gp models revealed differences in the TMHs spatial positions when compared to the starting MD structures (200 ns). While no changes were observed in both v3b and v3c models, the POPC molecule found in the portal 4/6 in the v3a model produced deviations in the spatial arrangement of the TMHs 4 and 6. Moreover, in the v3b model (obtained by removing the POPC molecule from the system's topology), changes in the positions of most TMHs were observed, including helices 4/6 and 10/12 forming both DBP portals. When the POPC was moved into the lipid bilayer in the v3c model, only slight differences in the spatial position of the TM helices 1, 3 and 4 were observed (*Supporting PDB file*).

All v3 models were also validated using docking studies to assess the number of DBSs within the DBP. The molecules tested were clustered in substrates, probes and modulators, in agreement with experimental[106,107] and *in silico* data[40]. Three DBSs were found inside the DBP in the v3c model, in agreement with the previously refined murine P-gp structure[39] and experimental data[5,20,25,108], while in both v3a and v3b models the three DBSs were found to be less defined (data not shown).

Altogether, our results showed significant improvements in the quality of human P-gp models after MD simulations indicating that the presence of the "linker" and lipid bilayer are undoubtedly important for the stability of the transporter, in agreement with other studies[39,40,104].

The analysis of the online evaluation servers, Ramachandran plot and RMSD did not reveal significant changes among the refined human P-gp models. However, as differences in the TMHs spatial position and number of DBSs within the DBP were found in both v3a and v3b models, the v3c model was considered the most stable and suitable human P-gp homology model to study the human P-gp variants. This model will be named as wild-type (WT) (Fig. 2). To compare the data gathered from the analysis of P-gp variants with the human WT model, v3c MD run was extended from 400 ns to 500 ns. This way, after a total of 700 ns of simulation time, a stable human P-gp model with high quality and robustness was obtained.

During the development of this work, the first human P-gp cryo-EM structure in a nucleotide-free inward-facing conformation was made available (PDB ID: 6QEX)[99]. According to the online servers (Table S1), our human WT P-gp model (v3c) presents better scores than the recently published 6QEX structure, indicating that the quality of our P-gp model is not lower than the quality of the 6QEX structure. Additionally, in the recently available QMEANBrane[109] server our human WT model also displayed similar local scores (0.746; WT vs 0.783; 6QEX) but more favorable membrane insertion energy (Fig. S4). Structural alignment between the human WT P-gp model and 6QEX structure showed similar spatial positions of the transmembrane and cytoplasmic domains

when compared with the initial template (Figs. S5 and S6). Nevertheless, significant differences in the secondary structure of the TMH4 (S237–A248) and TMH10 (S880-K885) were found between our P-gp model and 6QEX structure. When comparing the RMSD of each TMH (Table S2), TMH4 and TMH10 are the helices that present higher RMSD values between the WT model and 6QEX. Interestingly, a chimeric human-mouse P-gp cryo-EM structure (PDB ID: 6FN4)[110] presents similar alterations at the same regions of TMHs 4 and 10 and high RMSD values when compared to the human WT P-gp model.

In order to understand the nature of the TMH4 and TMH10 alterations in the human P-gp cryo-EM structure, four mouse P-gp crystallographic structures in the *apo* state inward-facing conformation (PDB IDs: 4M1M, 4Q9H, 5KPI[111] and 6GDI[112]) and in the presence of ligands at the DBP (PDB IDs: 4M2S[113] and 4Q9L[61]) were used for comparison. Visual inspection showed that both TMH4 and 10 are continuous α-helices in all *apo* and *holo* P-gp crystallographic structures considered, as observed in our human WT P-gp model. Moreover, Kim *et al.*[114] recently published a human P-gp cryo-EM structure in the ATP-bound, outward-facing conformation (PDB ID: 6C0V), demonstrating that the TMHs 4 and 10 must be continuous α-helices to ensure the correct closure of the cytoplasmic pore when shifting from the inward to outward-facing conformation upon NBD dimerization. Furthermore, structural alignment between each P-gp crystallographic structure and the human WT P-gp model also revealed lower RMSD values for both TM helices 4 and 10, in contrast with what was observed in the 6EQX structure (Table S2). As P-gp is a highly flexible protein, we cannot exclude that the recently cryo-EM structure captured some degree of dynamic transitions between several conformations of TM helices 4 and 10, but the above results also imply that v3c is a robust and reliable human WT P-gp model that can be further used in mechanistic and drug discovery studies. The comparison with the experimental data retrieved from the P-gp crystallographic structures indicate that our human WT model, although preserving the secondary structure of the TMDs, retains enough flexibility in the TM helices 4 and 10, thus being comparable to that observed in the human 6QEX structure.

Structural analysis of the human P-gp variants.

To gain additional insights on drug specificity and efflux-related signal transmission mechanism, the structural impact of four P-gp mutations (G185V[48–54], G830V[53,55], F978A[53,55,56] and ΔF335[57–60]), experimentally linked with changes in efflux and substrate specificity were analyzed. The mutations were selected according to their location within the DBP namely, at the substrate binding sites (SBSs) H and R sites (G185V and G830V, respectively) and at the modulator binding site (M-site) (F978A and ΔF335) (Fig. S7). Herein, while both H and R sites were initially characterized by their interaction with Hoechst 33342[108] or Rhodamine-123[35], respectively, the modulator M site was identified from the localization of the co-crystallized ligands QZ-SSS and QZ-RRR in the first crystallographic structure of murine P-gp[5]. Regarding SBSs, both were later characterized by molecular docking[40] and experimentally confirmed by electron microscopy[112].

One of the properties that can be altered by mutations in the transmembrane region of P-gp pump is the volume of the internal cavity. An estimation of the probability distribution function of the DBP volumes P (V) (Fig. S8) in the human WT P-gp model and variants showed that all mutations induced a reduction of the DBP volume, more pronounced in the F978A variant. Nevertheless, while G830V and F978A variants still sampled a wide range of DBP volumes during MD simulations, as observed in WT, G185V and ΔF335 variants showed narrower volume distributions suggesting a different structural cohesion of the transmembrane domains.

Thus, to investigate this hypothesis, the TMDs arrangement was analyzed through *g_bundle* and compared to the human WT P-gp model. Accordingly, this tool gives information about the bundle of axes (e.g. TMHs) such as the distance, length, and z-shift of the axis mid-points with respect to the average center of all axes as well as the total, lateral and radial tilt with respect to the average axis (*xy* corresponds to the membrane plane and *z* to the longest protein axis). The statistically significant changes in the bundle parameters are summarized in the Supporting Information (Tables S3–S8 and Figs. S9–S20).

Overall, the results show distinguishable changes in the TMHs repacking in all P-gp variants, including the helices where the respective mutation is located (Supporting PDB file). Moreover, all mutations showed significant changes in the bundle parameters of the TMHs 4/6 and 10/12 that form the DBP portals and in the "crossing helices" 4/5 and 10/11, that directly linked the TMD1 to NBD2 and TMD2 to NBD1. However, significant differences in the TMHs repacking between the mutations lying at the SBSs (G185V and G830V) and the mutations located at the M-site (F978A and ΔF335) were found.

Although being located at opposite helices (Fig. S7), the G185V and G830V mutations surprisingly showed similar structural changes in the TMHs rearrangement involving both TMDs. However, G830V mutation located at the TMH9 in the C-terminal P-gp halve seemed to have a stronger impact in the TMHs reorientation when compared to the G185V mutation in the TMH3 (Tables S5–S7). On the other hand, mutations at the M-site showed a completely distinct behavior. While the F978A mutation (TMH12) seemed to preferentially affect the TMHs rearrangement of the C-terminal P-gp halve where the mutation is located (Tables S4–S8), the deletion of F335 residue (ΔF335, TMH6) apparently did not have a severe impact in the TMHs rearrangement as could be initially expected (Supporting PDB file).

Altogether, the analysis of the helical bundle indicates that (i) all P-gp mutations affect the TMHs repacking including the TMHs that form the DBP portals, which may compromise the access of drugs to the internal cavity, (ii) all P-gp mutations affect the bundle of the "crossing helices" 4/5 and 10/11, described as important helices in the NBD dimerization process upon ATP binding[114] and (iii) mutations located at the SBSs have a different impact on helical repacking than those located at the M-site. Finally, as a result of the TMHs reorientation, all the selected mutations induce a reduction of the DBP volume in relation to WT (Fig. S6), but while both glycine and ΔF335 mutations induce a slight decreased in the pocket volume, the partial repacking caused by the F978A mutation may explain the severe reduction of the DBP volume found in this variant.

Interactions between coupling helices and nucleotide-binding domains.   Since the selected P-gp variants are experimentally linked with altered drug-resistance profiles and changes in either the basal or drug-stimulated ATPase activity, the residue interactions at the ICH-NBD interfaces thought to be involved in signal transmission and efflux-related conformational changes[9,10,26,45,115] were evaluated and compared to the human WT P-gp model. The total number of contacts was estimated using the *g_hbond* module and is depicted in *Supporting Information* (Figs. S21 and S22).

For both glycine variants, only at ICH3-NBD2 interface a significant decrease in the total number of contacts was observed. Oppositely, both M-site variants displayed a distinct behavior. While the F978A variant showed a decrease in the total number of contacts at the ICH3-NBD2 interface (similar to the glycine mutants), ΔF335 showed significant changes in the total number of contacts in three of the four interfaces, namely ICH2-NBD2, ICH3-NBD2 and ICH4-NBD1. As all mutants seemed to induce changes in the ICH-NBD total number of contacts, we further identified which residue pairs were involved. Mean residue-residue contact frequencies ≥ 0.5 and variations above 10% were considered significant and are summarized at *Supporting Information* (Table S10). For clarity purposes, each ICH-NBD will be analyzed separately in the following section.

Concerning the ICH1-NBD1 interface, all P-gp variants showed an increase of the mean contact frequencies between I160 (ICH1) and L443 (Walker A, NBD1), both located in regions identified to be involved in ATP binding[26,116,117]. Oppositely, only in F978A and ΔF335 a decrease in contact frequencies involving D164 (ICH1) and R404 (NBD1) was observed. Quite interestingly, D164 was also reported to be part of an extensive interaction surface between the TMDs and NBDs, with D164C mutants additionally revealing lower cell surface expression[118]. Regarding the ICH1-NBD1 hydrogen bond network, a significant decrease in the hydrogen-bond lifetimes (life, Table S10) was also observed for all mutants.

The other interface at this nucleotide-binding domain is the ICH4-NBD1. Herein, the most affected residue pairs were R905–S434/Q438/Q441, with a greater decrease of contact frequencies in both glycine mutants; V908/R467 (increased in all mutants); and S905/Y401 and S909–Q441/R467/V472, mainly increased in the F978A and ΔF335 mutants. Again, mutational studies implied both S905 and S909 in the activation and ATPase stimulation when in the presence of drugs and/or lipids[26,30]. Interestingly, and specifically concerning the ΔF335, new contacts between V907/F480, L910/R547, E913/R464 and Q914/R464 explain the increase in the total number of contacts reported above. Regarding the hydrogen bond network between the ICH4 and NBD1 residues, a significant decrease in the average HB number, lifetime and energy of HB formation was observed only in both glycine mutants (Table S10).

At the opposite NBD, all mutations seem to induce a general decrease in the overall contacts frequencies, with the ICH2 residues I265 and F267, involved in P-gp maturation and activity[10,32] and the NBD2 residue R1110, being the most affected ones. However, this decrease is partially mitigated by a reinforcement of the HB network in all mutations and, specifically for the ΔF335, through new contacts between F267 (ICH2) and G1134/R1188/A1189/R1192 (NBD2) or G269 (ICH2) and N1136 (NBD2) (Table S10). Finally, and regarding residue F1086, identified through *in vitro* studies as important for coupling of ATP binding to conformational changes in the TMDs[10], almost all contacts frequencies decrease except when paired with R262 (G185V and F978A) or I265 (ΔF335).

Finally, for the ICH3-NBD2 interface, all variants showed, in general, a negative variation in the contact frequencies, most particularly between V801/S802 (ICH3) and Y1087 (NBD2), thought to play an important role in P-gp activity and assembly[115]. In the same way, all mutations induced a decrease of the contact frequencies between the D805 (ICH3), an important residue thought to be involved in the TMD-NBD communication[119], and Y1044, located at the A-loop of NBD2.

Altogether, these results indicate that although mutations in the TMDs of the human P-gp affect directly the transmembrane region with changes in the DBP volume and DBSs features, they also induce changes in the residue interactions at the ICH-NBD signal transmission interfaces involved in the TMD-NBD communication. However, a comparison among P-gp variants showed that both mutations at the SBSs (G185V and G830V) induce identical changes in the total number of contacts between the ICH and NBD residues, while the impact of the M-site mutations (F978A and ΔF335) seems to be dependent of the TMH where the mutation is located, as observed in the helical bundle. Finally, the analysis of the mean contact frequencies in the human WT P-gp model and its variants, identified some residue pairs potentially involved in TMD-NBD communication, indicating an interaction network between the ICHs and NBDs residues, including those that directly interact with ATP, in agreement with several other studies[10,26,45,119,120].

Docking Results
Identification of Drug-Binding Sites in the human P-gp model.
Based on the docking poses of known P-gp substrates, probes and modulators, Ferreira and co-workers[40] defined the location of the DBSs within the DBP (firstly identified by Shapiro et al.[20,24,25] and Aller et al.[5]) using a refined murine P-gp structure[39]. To acquire a deeper knowledge on P-gp substrate binding, docking studies targeting the DBP of the human P-g WT model were undertaken. As twenty docking poses were generated per molecule, to simplify the results only the top-ranked binding energies (ΔG) at each DBS will be compared to the data obtained from previous studies[40].

Overall, most of the molecules tested bound at the three DBSs found within the DBP of the human WT model (Table S9), as observed in the murine P-gp structure. Nevertheless, higher number of molecules interacted with the H-site in respect to the refined murine P-gp. However, considering the standard error reported for VINA (2.85 kcal·mol^{-1})[121], no conclusions could be made regarding the affinity of these compounds since they have similar ΔG among the DBSs within the DBP of the WT model and similar ΔG than those reported in the refined murine P-gp structure.
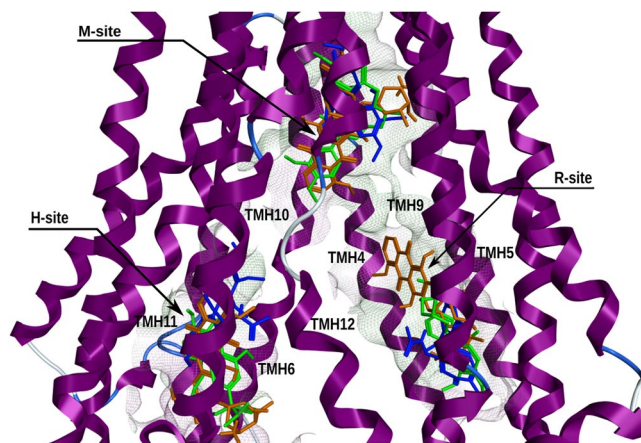
Figure 3. Representation of the DBSs found within the DBP of the human WT P-gp model. The three DBSs are defined by the best-ranked docking poses at each binding cavity of well-known P-gp substrates and modulators e.g. verapamil (green), doxorubicin (dark orange) and colchicine (blue).

Altogether, these results show that most P-gp substrates, probes and modulators interacted with the three DBSs (Fig. 3) suggesting that the DBP of the human WT P-gp model discriminate ligands differently than the internal cavity of the refined murine P-gp structure.

Identification of Drug-Binding Sites in the human P-gp variants. To assess if the TMHs repacking found in the P-gp variants affect drug binding, docking studies using known P-gp substrates and probes were performed as previously described and compared to the WT model. The molecules top-ranked binding energies (ΔG) at each DBS are depicted in Table S9 of the Supporting Information.

The total repacking of the transmembrane α-helices (TMHs) observed in both substrate binding site (SBS) variants (G185V and G830V) led to changes on substrate binding mostly affecting the R and H sites. Although more dramatic in the G830V variant, most of the compounds tested that showed to interact with the three DBSs in the WT model, did not bound at the H-site or did not interact with both SBSs. Additionally, and in contrast with what was observed for the G185V variant, some of the substrates that docked at the three DBSs in WT, did not interact with both M and H sites in the G830V variant. Although, no clear conclusions could be achieved about the possible changes in the molecule's affinity upon the TMHs rearrangement observed in these variants, nonetheless it becomes clear that the total helical repacking induced by the glycine mutations had a dramatic influence on the availability of each drug-binding site (DBS) to the evaluated set of molecules.

On the other hand, although the modulator binding site (M-site) variants (F978A and ΔF335) also showed changes on substrate binding mostly affecting the H or both SBSs, these mutations did not have a severe impact on drug binding as observed in the glycine variants. Nevertheless, due to the standard error reported for VINA and although no clear correlation could be obtained about the possible changes in drug affinity upon the TMHs repacking observed in these variants, it is quite interesting to note that specific mutations at the M-site have such a large influence on the SBSs rather than what would be initially expected, at the M-site.

Altogether, the analysis of the docking results demonstrates that all P-gp variants present changes on drug binding as a result of the total or partial TMHs repacking induced by these mutations, mostly affecting the R and H sites.

Characterization of Drug-Binding sites. To better understand the changes on substrate binding observed in the P-gp variants, the pocket volume, residues distribution and mean polarity of each DBS were assessed for the human WT model and variants using the EPOS[BP] software. The results were further compared to the human WT model and the refined murine P-gp structure.

Overall, both M and R sites showed similar pocket volumes among the human WT model (M-site, 1284 Å^3 and R-site, 1902 Å^3) and the refined murine P-gp structure (M-site, 1300 Å^3 and R-site, 1900 Å^3). In contrast, the H-site is considerable smaller (1232 Å^3) in human WT than in the refined murine structure (2200 Å^3). Additionally, and much like the refined murine structure, the M-site of the human WT showed to be the most hydrophobic DBS with higher number of aromatic residues in contrast with the H and R sites, that presented a higher percentage of polar residues. When comparing the site's residues distribution between species, no significant changes were found in the M and R sites, although the human WT H-site is more hydrophilic than the previously reported for the murine P-gp structure. However, despite the differences in volume and residues distribution found in the human WT H-site, this SBS was capable of binding more compounds than those observed in the refined murine P-gp structure.

Regarding P-gp variants, the total repacking induced by both glycine mutations led to changes in the residues' side-chain facing the H-site, with a remarkable increase in the content of hydrophobic (G185V, +53%; G830V, +75%) and aromatic side-chains (G830V, +63%) in respect to WT. Nevertheless, as a result of the stronger impact of G830V mutation on TMHs repacking, dramatic changes in the DBSs volume were also found. While a reduction around 50% was observed in the volume of R-site, the volumes of both M and H sites increased about

37% and 58%, respectively, in respect to WT. Therefore, for these mutations the data suggest that a slight increase in the hydrophilicity of the substrates may hamper its binding to the more hydrophobic sites.

Although the F978A mutation preferentially affected the TMHs repacking of the C-terminal P-gp halve, alterations in the residues' side-chain facing the H-site were also observed with significant variations in the percentage of hydrophobic (+89%) and aromatic side-chains (+26%). When compared to the other P-gp variants, the ΔF335 variant seemed to show minor variations in the volume and residues' side-chain facing all DBSs, although the H-site remained the most affected showing a decrease of its volume around 20% and an increase in the percentage of hydrophobic (+12%) and aromatic (+20%) side-chains. Similar to the glycine variants, lowering the hydrophobicity and reducing the number or aromatic rings is expected to decrease the binding to the mutated binding sites.

Altogether, these results indicate that the changes on drug binding found in all P-gp variants are related with alterations in the DBSs properties upon the rearrangement of the TMHs that delimited each DBSs. The visual inspection of each DBS showed that all mutations also induced changes in the DBSs structure, being more severe in the H-site. While in the WT model, the pocket entrance of the H-site is completely opened to the DBP with the formation of a cleft capable of accommodating substrates, all mutations induced distortions that affected either the pocket entrance and/or the cleft, compromising the binding of substrates (Fig. S23).

In sum, the characterization of the DBSs indicates that the total or partial TMHs repacking in response to mutations in the transmembrane domains (TMDs) seems to affect the shape, volume and residues distribution of the three DBSs within the drug-binding pocket (DBP). Interestingly, although changes on drug binding were also observed in the M and R sites, the H-site was the most affected by all mutations becoming unable to interact with most of the P-gp substrates and probes tested.

Final Discussion

Modulating drug efflux by P-gp pump is one of the promising strategies to reverse MDR in cancer cells. Nevertheless, the lack of information about the molecular basis underlying drug specificity and efflux-related signal transmission mechanism between the TMD-NBD domains impairs the development of more potent and selective compounds able to overcome MDR. Therefore, to provide additional insights on drug specificity and efflux mechanism, the impact of four P-gp mutations (G185V, G830V, F978A and ΔF335), experimentally linked with changes in efflux, basal and drug-stimulated ATPase activity, were comprehensively assessed.

In this work, a human P-gp homology model was developed based on murine P-gp crystallographic structure in the *apo* state inward-facing conformation (PDB ID: 4Q9H). The initial homology model obtained was refined through MD simulations in the presence of a "linker" retrieved from a previously equilibrated murine P-gp[39] structure, inserted in a POPC membrane also used in previous studies[122] and thereafter validated. By comparison with a recently published cryo-EM P-gp structure[99], our homology model was found to maintain an adequate reliability and robustness crucial for the herein proposed analysis.

The final human P-gp model was then used to thoroughly characterize the effect of the P-gp mutations mentioned above in the structure of the transporter. The mutations are located at the transmembrane region surrounding the M (F978A and ΔF335), H (G185V) and R (G830V) sites. Taken together, these mutations seem to induce TMHs repacking affecting the DBP portals by changing the "crossing helices" 4/5 and 10/11 important to NBD dimerization, and also compromising the access of drugs to the internal cavity, and by reducing the DBP volume. Additionally, as a result of the TMHs repacking significant changes in the volume, shape and polarity of the DBSs within the DBP were also observed in all P-gp variants, mostly affecting the binding of substrates at the H and R sites. It is noteworthy that although F978A and ΔF335 mutations lie at the M-site, they affect in a similar manner the SBSs properties and drug binding as observed in both glycine variants, suggesting a communication pathway between the M-site and the SBSs through the functional TMH6 and TMH12, in agreement with some experimental studies[20,22,123]. Therefore, we hypothesize that changes in the structure and polarity of the DBSs induced by the TMHs repacking i) provide a possible explanation for P-gp promiscuity, reported in the literature[13] and ii) suggest that small variations in both substrates and modulators may be enough to impair substrate binding and/or to enhance the modulators' activities. Interestingly, several examples on the latter are already described in literature as suitable approaches in enabling molecules to evade efflux[124] or even to switch the activity of known substrates into high-affinity compounds able to inhibit P-gp's ATPase activity[125].

Even though the mutations described above directly affected the transmembrane region, they also induced changes in the total number of contacts at the ICH-NBD interfaces, suggesting that the TMHs rearrangement is involved in the TMD-NBD communication, in agreement with several experimental[10,126,127] and *in silico* studies[8,9,45,46]. Furthermore, all P-gp variants showed significant changes in the mean contact frequencies of specific residue pairs, mainly located at the ICH2/ICH3-NBD2, an important transmission interface to couple drug binding to ATPase activity, but also being critical for P-gp folding[9,45,127,128]. Thus, another interesting approach to the modification of substrates/inhibitors is the development of allosteric modulators able to specifically interact at the ICH-NBD interfaces and impair the signal transmission between the TMD and NBDs[116]. Currently, only two scaffolds are currently known to interact in such domains, namely dihydropyridines (TMD-NBD1)[128] and flavonoids (NBD2)[129], and additional efforts must be taken in the future to explore this hypothesis.

Nevertheless, although all P-gp mutations induced similar structural effects on the transporters' architecture, it seems clear that mutations at the M-site (F978A and ΔF335) have a completely different impact on P-gp structure than the mutations located at the SBSs (G185V and G830V). Both glycine mutations induced a total TMHs repacking affecting drug binding at the SBSs. Additionally, the glycine mutations do not have a significant impact in the total number of contacts as expected, suggesting that G185 (TMH3) and G830V (TMH9) residues have equivalent roles in P-gp function and possibly more involved in drug binding.

On the other hand, the effects of the M-site mutations show to be dependent of the affected TMH. The F978A mutation (TMH12) preferentially induced a partial TMH repacking, affecting drug binding at the SBSs while

ΔF335 mutation (TMH6) dramatically change the residue interactions at the ICH-NBD interfaces. These results support the hypothesis that while the F978 residue is likely involved in drug binding as reported in some experimental studies[53,55,130,131], F335 residue is involved in TMD-NBD communication[59,60].

Conclusion

Overall, this work provides clear evidence that mutations at these specific TMHs (i) are responsible for inducing a repacking of the TMHs, changing the DBP volume and drug binding sites, mostly affecting drug binding at the SBSs and (ii) also impact the ICH-NBD signal transmission interfaces, suggesting that a perturbation in the TMDs (e.g. mutations or binding of substrates) induce a TMHs rearrangement that are transmitted to the NBDs through changes in the residue interactions between the ICHs and the respective NBD, in agreement with experimental studies[61,130,131].

Nevertheless, as protein conformational changes may occur at least on the timescale of microseconds[132] and the computational power available is limited, it should be clear that our findings are based in theoretical models aiming to predict the possible structural impact of single-point mutations on P-gp architecture/organization. Additionally, as these P-gp mutations are experimentally related with changes in drug-stimulated ATPase activity upon binding of specific substrates, more studies are needed to assess their direct impact on drug and ATP binding. Therefore, further studies including molecules with altered efflux properties are undergoing, in an attempt to clarify these issues.

Data availability

The final configurations of the MD refined human P-gp homology structure and its variants (G185V, G830V, F978A and ΔF335) are available for download at our website (http://chemistrybits.com/).

References

1. Gottesman, M. M., Lavi, O., Hall, M. D. & Gillet, J.-P. Toward a Better Understanding of the Complexity of Cancer Drug Resistance. *Annu. Rev. Pharmacol. Toxicol.* **56**, 85–102 (2016).
2. Juliano, R. L. & Ling, V. A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants. *Biochim. Biophys. Acta.* **455**, 152–162 (1976).
3. Ferreira, R. J., Ferreira, M. J. U. & dos Santos, D. J. V. A. Reversing cancer multidrug resistance: Insights into the efflux by ABC transports from *in silico* studies. *WIRES Comput. Mol. Sci.* **5**, 27–55 (2015).
4. Sharom, F. J. The P-glycoprotein efflux pump: how does it transport drugs? *J. Membr. Biol.* **160**, 161–175 (1997).
5. Aller, S. G. *et al*. Structure of P-Glycoprotein Reveals a Molecular Basis for Poly-Specific Drug Binding. *Science.* **323**, 1718–1722 (2009).
6. Ferreira, R. J., Ferreira, M. J. U. & dos Santos, D. J. V. A. Assessing the stabilization of P-glycoprotein's nucleotide-binding domains by the linker, using molecular dynamics. *Mol. Inf.* **32**, 529–540 (2013).
7. Borges-Walmsley, A. R., Mckeegan, M. I. & Walmsley, K. S. Structure and function of efflux pumps that confer resistance to drugs. *Biochem. J* **376**, 313–338 (2003).
8. Ferreira, R. J., Bonito, C. A., Cordeiro, M. N. D. S., Ferreira, M. J. U. & dos Santos, D. J. V. A. Structure-function relationships in ABCG2: Insights from molecular dynamics simulations and molecular docking studies. *Sci. Rep.* **7**, 1–17, https://doi.org/10.1038/s41598-017-15452-z (2017).
9. Pajeva, I. K., Hanl, M. & Wiese, M. Protein contacts and ligand binding in the inward-facing model of human P-glycoprotein. *Chem Med Chem.* **8**, 748–762 (2013).
10. Loo, T. W., Bartlett, M. C. & Clarke, D. M. Human P-glycoprotein Contains a Greasy Ball-and-Socket Joint at the Second Transmission Interface. *J. Biol. Chem.* **288**, 20326–20333 (2013).
11. Loo, T. W. & Clarke, D. M. 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 (2001).
12. Loo, T. W. & Clarke, D. M. Determining the Dimensions of the Drug-binding Domain of Human P-glycoprotein Using Thiol Cross-linking Compounds as Molecular Rulers. *J. Biol. Chem.* **276**, 36877–36880 (2001).
13. Al-Shawi, M. K. & Omote, H. The remarkable transport mechanism of P-glycoprotein: a multidrug transporter. *J. Bioenerg. Biomembr.* **37**, 489–496 (2005).
14. Rosenberg, M. F. *et al*. Repacking of the transmembrane domains of P-glycoprotein during the transport ATPase cycle. *EMBO J* **20**, 5615–5625 (2001).
15. Julien, M. & Gros, P. Nucleotide-induced conformational changes in P-glycoprotein and in nucleotide binding site mutants monitored by trypsin sensitivity. *Biochemistry.* **39**, 4559–4568 (2000).
16. Callaghan, R., Ford, R. C. & Kerr, I. D. The translocation mechanism of P-glycoprotein. *FEBS Lett* **580**, 1056–1063 (2006).
17. Siarheyeva, A., Liu, R. & Sharom, F. J. Characterization of an asymmetric occluded state of P-glycoprotein with two bound nucleotides: Implications for catalysis. *J. Biol. Chem.* **285**, 7575–7586 (2010).
18. Verhalen, B. *et al*. Energy transduction and alternating access of the mammalian ABC transporter P-glycoprotein. *Nature.* **543**, 738–741 (2017).
19. Dey, S., Ramachandra, M., Pastan, I., Gottesman, M. M. & Ambudkar, S. V. Evidence for two nonidentical drug-interaction sites in the human P-glycoprotein. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 10594–10599 (1997).
20. Shapiro, A. B., Fox, K., Lam, P. & Ling, V. Stimulation of P-glycoprotein-mediated drug transport by prazosin and progesterone. Evidence for a third drug-binding site. *Eur. J. Biochem.* **259**, 841–850 (1999).
21. Shapiro, A. B. & Ling, V. Positively cooperative sites for drug transport by P-glycoprotein with distinct drug specificities. *Eur. J. Biochem.* **250**, 130–137 (1997).
22. Shapiro, A. B. & Ling, V. Effect of quercetin on Hoechst 33342 transport by purified and reconstituted P-glycoprotein. *Biochem. Pharmacol.* **53**, 587–596 (1997).
23. Shapiro, A. B. & Ling, V. Reconstitution of drug transport by purified P-glycoprotein. *J. Biol. Chem.* **270**, 16167–16175 (1995).
24. Shapiro, A. B., Corder, A. B. & Ling, V. P-glycoprotein-mediated Hoechst 33342 transport out of the lipid bilayer. *Eur. J. Biochem.* **250**, 115–121 (1997).
25. Shapiro, A. B. & Ling, V. Transport of LDS-751 from the cytoplasmic leaflet of the plasma membrane by the rhodamine-123-selective site of P-glycoprotein. *Eur. J. Biochem.* **254**, 181–188 (1998).
26. Loo, T. W. & Clarke, D. M. Drugs Modulate Interactions between the First Nucleotide-Binding Domain and the Fourth Cytoplasmic Loop of Human P-Glycoprotein. *Biochemistry.* **55**, 2817–2820 (2016).

27. Loo, T. W., Bartlett, M. C. & Clarke, D. M. ATP hydrolysis promotes interactions between the extracellular ends of transmembrane segments 1 and 11 of human multidrug resistance P-glycoprotein. *Biochemistry.* **44**, 10250–10258 (2005).
28. Loo, T. W. & Clarke, D. M. Attachment of a "molecular spring" restores drug-stimulated ATPase activity to P-glycoprotein lacking both Q loop glutamines. *Biochem. Biophys. Res. Commun.* 1–5, https://doi.org/10.1016/j.bbrc.2016.12.137 (2016).
29. Loo, T. W., Bartlett, M. C. & Clarke, D. M. Human P-glycoprotein is active when the two halves are clamped together in the closed conformation. *Biochem. Biophys. Res. Commun.* **395**, 436–440 (2010).
30. Loo, T. W. & Clarke, D. M. P-glycoprotein ATPase activity requires lipids to activate a switch at the first transmission interface. *Biochem. Biophys. Res. Commun.* **472**, 379–383 (2016).
31. Loo, T. W. & Clarke, D. M. Identification of the Distance between the Homologous Halves of P-glycoprotein That Triggers the High/Low ATPase Activity Switch. *J. Biol. Chem.* **289**, 8484–8492 (2014).
32. Loo, T. W. & Clarke, D. M. The Transmission Interfaces Contribute Asymmetrically to the Assembly and Activity of Human P-glycoprotein. *J. Biol. Chem.* **290**, 16954–16963 (2015).
33. Loo, T. W. & Clarke, D. M. 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 (2000).
34. Loo, T. W. & Clarke, D. M. Vanadate trapping of nucleotide at the ATP-binding sites of human multidrug resistance P-glycoprotein exposes different residues to the drug-binding site. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 3511–3516 (2002).
35. Loo, T. W. & Clarke, D. M. Location of the rhodamine-binding site in the human multidrug resistance P-glycoprotein. *J. Biol. Chem.* **277**, 44332–44338 (2002).
36. Loo, T. W., Bartlett, M. C., Detty, M. R. & Clarke, D. M. 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 (2012).
37. Dolghih, E., Bryant, C., Renslo, A. R. & Jacobson, M. P. Predicting Binding to P-Glycoprotein by Flexible Receptor Docking. *Plos Comput. Biol.* **7**, e1002083, https://doi.org/10.1371/journal.pcbi.1002083 (2011).
38. Ferreira, R. J., Ferreira, M. J. U. & dos Santos, D. J. V. A. Do Drugs Have Access to the P-Glycoprotein Drug-Binding Pocket through Gates? *J. Chem. Theory Comput.* **11**, 4525–4529 (2015).
39. Ferreira, R. J., Ferreira, M. J. U. & dos Santos, D. J. V. A. Insights on P-Glycoprotein's Efflux Mechanism Obtained by Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **8**, 1853–1864 (2012).
40. Ferreira, R. J., Ferreira, M. J. U. & dos Santos, D. J. V. A. Molecular docking characterizes substrate-binding sites and efflux modulation mechanisms within P-glycoprotein. *J. Chem. Inf. Model* **53**, 1747–1760 (2013).
41. Hrycyna, C. A. *et al.* Structural flexibility of the linker region of human P-glycoprotein permits ATP hydrolysis and drug transport. *Biochemistry.* **37**, 13660–13673 (1998).
42. Sato, T. *et al.* Functional role of the linker region in purified human P-glycoprotein. *FEBS J* **276**, 3504–3516 (2009).
43. Ma, J. & Biggin, P. C. Substrate versus inhibitor dynamics of P-glycoprotein. *Proteins.* **81**, 1653–1668 (2013).
44. Mukhametov, A. & Raevsky, O. A. On the mechanism of substrate/non-substrate recognition by P -glycoprotein. *J. Mol. Graph. Model* **71**, 227–232 (2017).
45. Wise, J. G. Catalytic transitions in the human mdr1 P-glycoprotein drug binding sites. *Biochemistry.* **51**, 5125–5141 (2012).
46. Prajapati, R. & Sangamwar, A. T. Translocation mechanism of P-glycoprotein and conformational changes occurring at drug-binding site: Insights from multi-targeted molecular dynamics. *Biochim. Biophys. Acta.* **1838**, 2882–2898 (2014).
47. Wang, Z. & Liao, J.-L. Probing Structural Determinants of ATP-Binding Cassette Exporter Conformational Transition Using Coarse-Grained Molecular Dynamics. *J. Phys. Chem. B* **119**, 1295–1301 (2015).
48. Choi, K. H., Chen, C. J., Kriegler, M. & Roninson, I. B. An altered pattern of cross-resistance in multidrug-resistant human cells results from spontaneous mutations in the mdr1 (P-glycoprotein) gene. *Cell.* **53**, 519–229 (1988).
49. Safa, A. R. *et al.* Molecular basis of preferential resistance to colchicine in multidrug-resistant human cells conferred by Gly-185-Val-185 substitution in P-glycoprotein. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 7225–7229 (1990).
50. Rao, U. S. Mutation of glycine 185 to valine alters the ATPase function of the human P-glycoprotein expressed in Sf9 cells. *J. Biol. Chem.* **270**, 6686–6690 (1995).
51. Ramachandra, M., Ambudkar, S. V., Gottesman, M. M., Pastan, I. & Hrycyna, C. A. Functional characterization of a glycine 185-to-valine substitution in human P-glycoprotein by using a vaccinia-based transient expression system. *Mol. Biol. Cell.* **7**, 1485–1498 (1996).
52. Loo, T. W. & Clarke, D. M. Rapid Purification of Human P-glycoprotein Mutants Expressed Transiently in HEK 293 Cells by Nickel-Chelate Chromatography and Characterization of their Drug-stimulated ATPase Activities. *J. Biol. Chem.* **270**, 21449–21452 (1995).
53. Omote, H., Figler, R. A., Polar, M. K. & Al-Shawi, M. K. Improved energy coupling of human P-glycoprotein by the glycine 185 to valine mutation. *Biochemistry.* **43**, 3917–3928 (2004).
54. Loo, T. W. & Clarke, D. M. Functional Consequences of Glycine Mutations in the Predicted Cytoplasmic Loops of P-glycoprotein. *J. Biol. Chem.* **269**, 7243–7249 (1994).
55. Loo, T. W. & Clarke, D. M. Functional consequences of phenylalanine mutations in the predicted transmembrane domain of P-glycoprotein. *J. Biol. Chem.* **268**, 19965–19972 (1993).
56. Mittra, R. *et al.* Location of contact residues in pharmacologically distinct drug binding sites on P-glycoprotein. *Biochem. Pharmacol.* **123**, 19–28 (2017).
57. Harker, W. G. & Sikic, B. I. Multidrug (pleiotropic) resistance in doxorubicin-selected variants of the human sarcoma cell line MES-SA. *Cancer Res.* **45**, 4091–4096 (1985).
58. Harker, W. G., MacKintosh, F. R. & Sikic, B. I. Development and characterization of a human sarcoma cell line, MES-SA, sensitive to multiple drugs. *Cancer Res.* **43**, 4943–4950 (1983).
59. Chen, G. *et al.* Multidrug-resistant human sarcoma cells with a mutant P-glycoprotein, altered phenotype, and resistance to cyclosporins. *J. Biol. Chem.* **272**, 5974–5982 (1997).
60. Chen, G. K., Lacayo, N. J., Durán, G. E., Cohen, D. & Sikic, B. I. Loss of cyclosporin and azidopine binding are associated with altered ATPase activity by a mutant P-glycoprotein with deleted phe(335). *Mol. Pharmacol.* **57**, 769–777 (2000).
61. Szewczyk, P. *et al.* Snapshots of ligand entry, malleable binding and induced helical movement in P-glycoprotein. *Acta Crystallogr. D Biol. Crystallogr* **71**, 732–741 (2015).
62. Rose, P. W. *et al.* The RCSB protein data bank: integrative view of protein, gene and 3D structural information. *Nucleic Acids Res* **45**, D271–D281 (2017).
63. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res* **28**, 235–242 (2000).
64. Labute, P. Protonate3D: Assignment of ionization states and hydrogen coordinates to macromolecular structures. *Proteins.* **75**, 187–205 (2009).
65. Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput* **4**, 435–447 (2008).
66. van Der Spoel, D. *et al.* GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).
67. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).

68. Poger, D. & Mark, A. E. On the Validation of Molecular Dynamics Simulations of Saturated and cis-Monounsaturated Phosphatidylcholine Lipid Bilayers: A Comparison with Experiment. *J. Chem. Theory Comput.* **6**, 325–336 (2010).
69. Poger, D., van Gunsteren, W. F. & Mark, A. E. A new force field for simulating phosphatidylcholine bilayers. *J. Comput. Chem.* **31**, 1117–1125 (2010).
70. Lomize, M. A., Pogozheva, I. D., Joo, H., Mosberg, H. I. & Lomize, A. L. OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Res* **40**, D370–D376 (2012).
71. Wolf, M. G., Hoefling, M., Aponte-Santamaría, C., Grubmüller, H. & Groenhof, G. g_membed: Efficient insertion of a membrane protein into an equilibrated lipid bilayer with minimal perturbation. *J. Comput. Chem.* **31**, 2169–2174 (2010).
72. Wolf, M. G., Hoefling, M., Aponte-Santamaría, C., Grubmüller, H. & Groenhof, G. Corrigendum: g_membed: Efficient insertion of a membrane protein into an equilibrated lipid bilayer with minimal perturbation. *J. Comput. Chem.* **37**, 2038–2038 (2016).
73. Davis, I. W. *et al*. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35**, W375–W383 (2007).
74. Chen, V. B. *et al*. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr* **66**, 12–21 (2010).
75. Colovos, C. & Yeates, T. O. Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Sci.* **2**, 1511–1519 (1993).
76. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R. & Thornton, J. M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR.* **8**, 477–486 (1996).
77. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291 (1993).
78. Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics.* **22**, 195–201 (2006).
79. Biasini, M. *et al*. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* **42**, W252–W258 (2014).
80. Bordoli, L. *et al*. Protein structure homology modeling using SWISS-MODEL workspace. *Nat. Protoc.* **4**, 1–13 (2008).
81. Schwede, T. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res* **31**, 3381–3385 (2003).
82. Lovell, S. C. *et al*. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins.* **50**, 437–450 (2003).
83. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 14101–14107 (2007).
84. Nosé, S. & Klein, M. L. Constant pressure molecular dynamics for molecular systems. *Mol. Phys.* **50**, 1055–1076 (1983).
85. Hoover, H. Canonical dynamics: Equilibrium phase-space distributions. *Phys Rev A* **31**, 1695–1697 (1985).
86. Parrinello, A. & Rahman, M. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
87. Hess, B., Bekker, H., Berendsen, H. J. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).
88. Hess, B. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *J. Chem. Theory Comput.* **4**, 116–122 (2008).
89. Miyamoto, S. & Kollman, P. A. SETTLE: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **13**, 952–962 (1992).
90. Essmann, U. *et al*. A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).
91. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
92. Pronk, S. *et al*. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics.* **29**, 845–854 (2013).
93. van der Spoel, D., van Maaren, P. J., Larsson, P. & Timneanu, N. Thermodynamics of Hydrogen Bonding in Hydrophilic and Hydrophobic Media. *J. Phys. Chem. B* **110**, 4393–4398 (2006).
94. Blau, C. & Grubmuller, H. g_contacts: Fast contact search in bio-molecular ensemble data. *Comput. Phys. Commun.* **184**, 2856–2859 (2013).
95. Brady, G. P. & Stouten, P. F. Fast prediction and visualization of protein binding pockets with PASS. *J. Comput. Aided Mol. Des* **14**, 383–401 (2000).
96. Kohlbacher, O. & Lenhof, H. P. BALL–rapid software prototyping in computational molecular biology. Biochemicals Algorithms Library. *Bioinformatics.* **16**, 815–824 (2000).
97. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
98. Chufan, E. E. *et al*. Multiple Transport-Active Binding Sites Are Available for a Single Substrate on Human P-Glycoprotein (ABCB1). *Plos One.* **8**, e82463, https://doi.org/10.1371/journal.pone.0082463, (2013).
99. Alam, A., Kowal, J., Broude, E., Roninson, I. & Locher, K. P. Structural insight into substrate and inhibitor discrimination by human P-glycoprotein. *Science.* **363**, 753–756 (2019).
100. Wen, P.-C., Verhalen, B., Wilkens, S., Mchaourab, H. S. & Tajkhorshid, E. On the origin of large flexibility of P-glycoprotein in the inward-facing state. *J. Biol. Chem.* **288**, 19211–19220 (2013).
101. Clay, A. T., Lu, P. & Sharom, F. J. Interaction of the P-Glycoprotein Multidrug Transporter with Sterols. *Biochemistry.* **54**, 6586–6597 (2015).
102. Qin, L., Hiser, C., Mulichak, A., Garavito, R. M. & Ferguson-Miller, S. Identification of conserved lipid/detergent-binding sites in a high-resolution structure of the membrane protein cytochrome c oxidase. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 16117–16122 (2006).
103. Ferreira, R. J., dos Santos, D. J. V. A. & Ferreira, M.-J. U. P-glycoprotein and membrane roles in multidrug resistance. *Future Med. Chem.* **7**, 929–946 (2015).
104. Domicevica, L. & Biggin, P. C. Homology modelling of human P-glycoprotein. *Biochem. Soc. Trans* **43**, 952–958 (2015).
105. Benkert, P., Künzli, M. & Schwede, T. QMEAN server for protein model quality estimation. *Nucleic Acids Res* **37**, W510–W514 (2009).
106. Polli, J. W. *et al*. Rational use of *in vitro* P-glycoprotein assays in drug discovery. *J. Pharmacol. Exp. Ther.* **299**, 620–628 (2001).
107. Rautio, J. *et al*. *In vitro* p-glycoprotein inhibition assays for assessment of clinical drug interaction potential of new drug candidates: a recommendation for probe substrates. *Drug Metab. Dispos.* **34**, 786–792 (2006).
108. Shapiro, A. B. & Ling, V. Extraction of Hoechst 33342 from the cytoplasmic leaflet of the plasma membrane by P-glycoprotein. *Eur. J. Biochem.* **250**, 122–129 (1997).
109. Studer, G., Biasini, M. & Schwede, T. Assessing the local structural quality of transmembrane protein models using statistical potentials (QMEANBrane). *Bioinformatics.* **30**, 505–511 (2014).
110. Alam, A. *et al*. Structure of a zosuquidar and UIC2-bound human-mouse chimeric ABCB1. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E1973–E1982 (2018).
111. Esser, L. *et al*. Structures of the Multidrug Transporter P-glycoprotein Reveal Asymmetric ATP Binding and the Mechanism of Polyspecificity. *J. Biol. Chem.* **292**, 446–461 (2017).
112. Thonghin, N. *et al*. Novel features in the structure of P-glycoprotein (ABCB1) in the post-hydrolytic state as determined at 7.9 Å resolution. *BMC Struct. Biol.* **18**, 17 (2018).
113. Li, J., Jaimes, K. F. & Aller, S. G. Refined structures of mouse P-glycoprotein. *Protein Sci.* **23**, 34–46 (2014).

114. Kim, Y. & Chen, J. Molecular structure of human P-glycoprotein in the ATP-bound, outward-facing conformation. *Science.* **359**, 915–919 (2018).
115. Loo, T. W., Bartlett, M. C. & Clarke, D. M. Permanent activation of the human P-glycoprotein by covalent modification of a residue in the drug-binding site. *J. Biol. Chem.* **278**, 20449–20452 (2003).
116. Ferreira, R. J., Bonito, C. A., Ferreira, M. J. U. & dos Santos, D. J. V. A. About P-glycoprotein: a new druggable domain is emerging from structural data. *WIREs Comput. Mol. Sci.* 1–22, https://doi.org/10.1002/wcms.1316 (2017).
117. Kwan, T. & Gros, P. Mutational analysis of the P-glycoprotein first intracellular loop and flanking transmembrane domains. *Biochemistry.* **37**, 3337–3350 (1998).
118. Kapoor, K., Bhatnagar, J., Chufan, E. E. & Ambudkar, S. V. Mutations in intracellular loops 1 and 3 lead to misfolding of human P-glycoprotein (ABCB1) that can be rescued by cyclosporine A, which reduces its association with chaperone Hsp70. *J. Biol. Chem.* **288**, 32622–32636 (2013).
119. Becker, J. *et al.* Molecular models of human P-glycoprotein in two different catalytic states. *BMC Struct. Biol.* **9**, 1–18 (2009).
120. Zhang, Y., Gong, W., Wang, Y., Liu, Y. & Li, C. Exploring movement and energy in human P-glycoprotein conformational rearrangement. *J. Biomol. Struct. Dyn.* **37**, 1104–1119 (2019).
121. Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).
122. Baptista, R., Ferreira, R. J., dos Santos, D. J. V. A., Fernandes, M. X. & Ferreira, M.-J. U. Optimizing the macrocyclic diterpenic core toward the reversal of multidrug resistance in cancer. *Future Med. Chem.* **8**, 629–645 (2016).
123. Martin, C. *et al.* Communication between multiple drug binding sites on P-glycoprotein. *Mol. Pharmacol.* **58**, 624–632 (2000).
124. Waghray, D. & Zhang, Q. Inhibit or Evade Multidrug Resistance P-glycoprotein in Cancer Treatment. *J. Med. Chem.* **61**, 5108–5121 (2018).
125. Gannon, M. K. *et al.* Rhodamine Inhibitors of P-Glycoprotein: an amide/thioamide "switch" for ATPase activity. *J. Med. Chem.* **52**, 3328–3341 (2009).
126. Loo, T. W. & Clarke, D. M. A salt bridge in intracellular loop 2 is essential for folding of human P-glycoprotein. *Biochemistry.* **52**, 3194–3196 (2013).
127. Loo, T. W. & Clarke, D. M. Locking Intracellular Helices 2 and 3 Together Inactivates Human P-glycoprotein. *J. Biol. Chem.* **289**, 229–236 (2014).
128. Isenberg, B., Thole, H., Tümmler, B. & Demmer, A. Identification and localization of three photobinding sites of iodoarylazidoprazosin in hamster P-Glycoprotein. *Eur. J. Biochem* **268**, 2629–2634 (2001).
129. Conseil, G. *et al.* Flavonoids: a class of Modulators with bifunctional interactions at vicinal ATP- and steroid-binding sites on mouse P-glycoprotein. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 9831–9836 (1998).
130. Hafkemeyer, P. *et al.* Contribution to substrate specificity and transport of nonconserved residues in transmembrane domain 12 of human P-glycoprotein. *Biochemistry.* **37**, 16400–16409 (1998).
131. Loo, T. W., Bartlett, M. C. & Clarke, D. M. Substrate-induced conformational changes in the transmembrane segments of human P-glycoprotein. Direct evidence for the substrate-induced fit mechanism for drug binding. *J. Biol. Chem* **278**, 13603–13606 (2003).
132. Sekhar, A., Vallurupalli, P. & Kay, L. E. Defining a length scale for millisecond-timescale protein conformational exchange. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 11391–11396 (2013).
133. Chemical Computing Group Inc., Molecular Operating Environment (MOE) 2016.08 and 2019.01 (2019).

Acknowledgements

Author contributions

C.A.B., R.J.F. and D.J.V.A.S. conceived the experiment(s), C.A.B. conducted the experiment(s), R.J.F., C.A.B. and D.J.V.A.S. analyzed the results and wrote the paper. C.A.B., R.J.F., D.J.V.A.S., M.J.U.F., J.-P.G. and M.N.D.S.C. reviewed the manuscript. All authors agreed with the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-66587-5.

Correspondence and requests for materials should be addressed to D.J.V.A.d.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

ANNEX S2

"Theoretical insights on helix repacking as the origin of

P-glycoprotein promiscuity"

TABLE OF CONTENTS

Figure S1. Schematic representation of the fully unrestrained *NpT* runs performed to obtain a stable human P-gp homology model and variants. The murine template used are colored as red (PDB ID:4Q9H), and the mutated structures in black.

Figure S2. Ramachandran plots of the initial and refined human P-gp models after 400 ns MD simulations. The outliers are represented as red crosses, allowed positions in yellow and core angles in green.
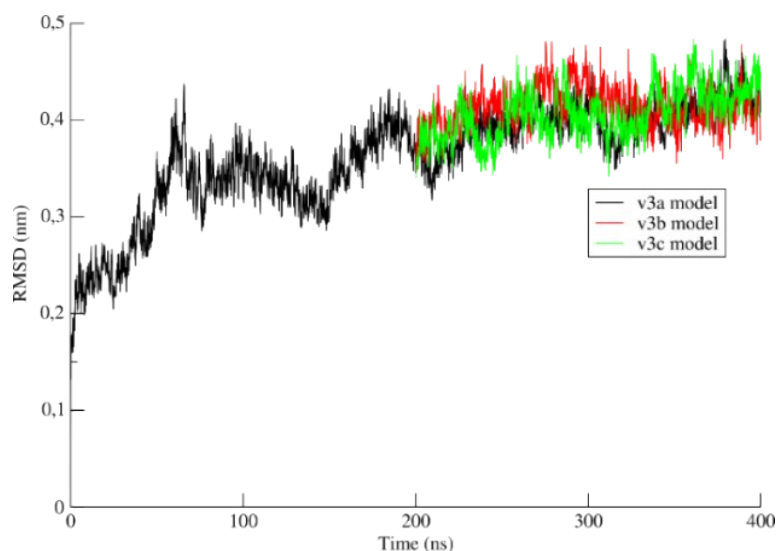
Figure S3. Root mean square deviation (RMSD) of the human P-gp v3 models during MD simulations.

Table S1. Structural quality assessment of the human P-pg homology models

			MOLPROBITY		SwissModel		PROCHEK	
		Assessment software						
		ERRAT	Score	Percentile	QMEAN6	Z-Score	Morris	G-factors
Crystal (PDB ID: 4Q9H)	w/o linker	90.05	1.86	83	0.61	-1.75	1-2-2	0.18
Homology model	w/o linker	74.40	2.22	63	0.59	-1.89	1-2-3	-0.41
	w/ linker	75.08	2.21	64	0.58	-2.05	1-2-3	-0.41
MD Refined P-gp models	V3a	93.42	1.71	89	0.46	-3.39	1-2-2	-0.61
	V3b	96.65	1.64	91	0.44	-3.52	1-2-2	-0.58
	V3c	97.79	1.65	91	0.43	-3.68	1-2-2	-0.57
Cryo-EM (PDB ID: 6QEX)	w/o linker	85.54	1.78	86	0.70	-1.65	1-2-2	0.06
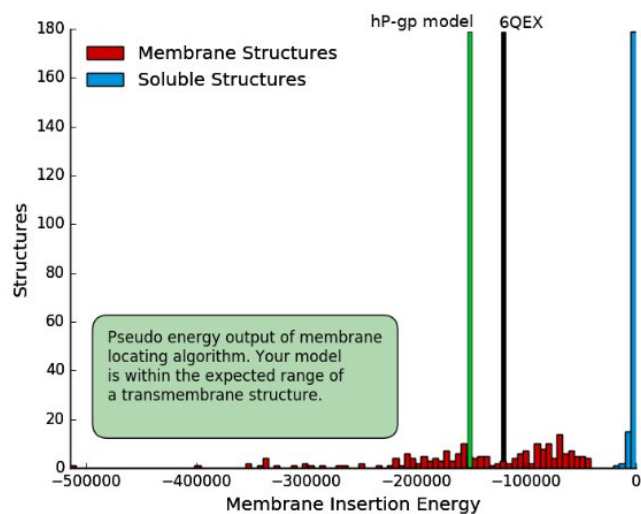
Figure S4. Evaluation of membrane insertion energies for the *in-house* P-gp model and cryo-EM human P-gp model (PDB ID: 6QEX) obtained from QMEMBrane server.
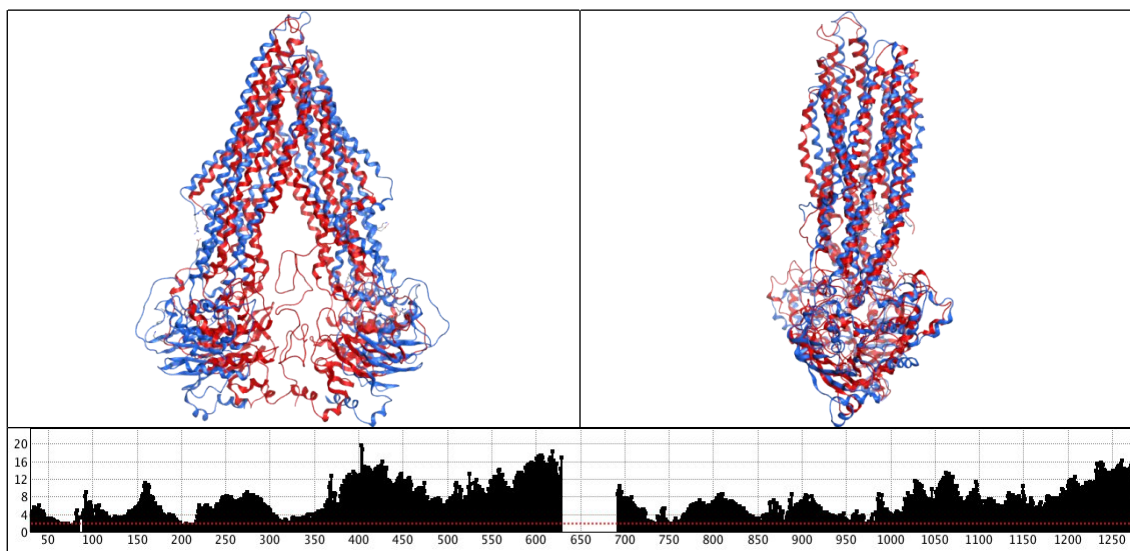
Figure S5. Root-mean square deviation (total RMSD: 8.003) between the template (4Q9H, blue) and the final homology model (v3c, red) structures (top) and residue-by-residue (bottom). RMSD values are depicted in angstroms.
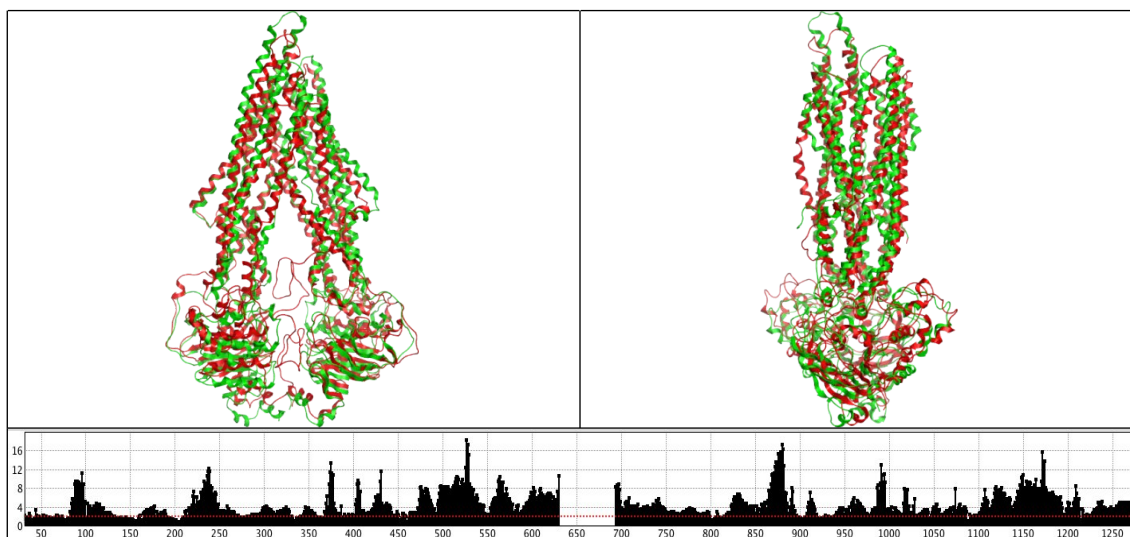
Figure S6. Root-mean square deviation (total RMSD = 4.931) between the cryo-EM (6QEX, green) and the final homology model (v3c, red) structures (top) and residue-by-residue (bottom). RMSD values are depicted in angstroms.

Table S2. RMSD values for each TMH between the human WT P-gp model (v3c) and 6QEX human P-gp structure. The RMSD values of the TMHs 4 and 10 between the WT model and human/mouse P-gp structures.

PDB	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
6QEX	1.71	1.57	1.59	6.07	1.13	2.08	1.43	1.62	1.61	4.94	1.73	3.34
6FN4				4.87						4.89		
6GDI				1.07						2.86		
5KPI				1.51						2.61		
4Q9H				1.58						2.72		
4M1M				1.42						2.72		
4M2S				1.42						2.66		
4Q9L				1.64						2.76		

4
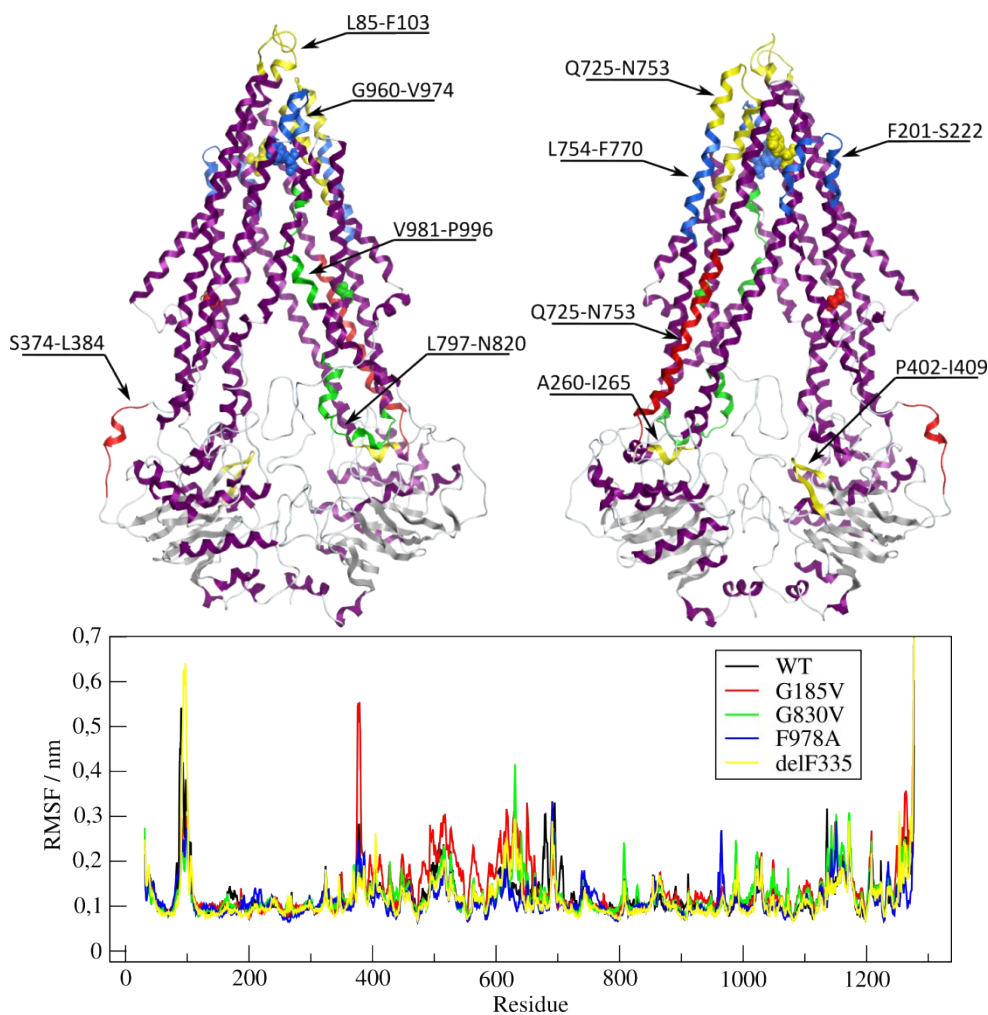
Figure S7. (top) Location of the selected mutations (in van der Waals, G185V, red; G830V, green; F978A, blue; ΔF335, yellow) in the human WT P-gp model. Aminoacid sequences with larger RMSF differences from WT model are depicted with the same color code; (bottom), comparison of root-mean square fluctuations (RMSF) from mutated P-gp models against WT (black).
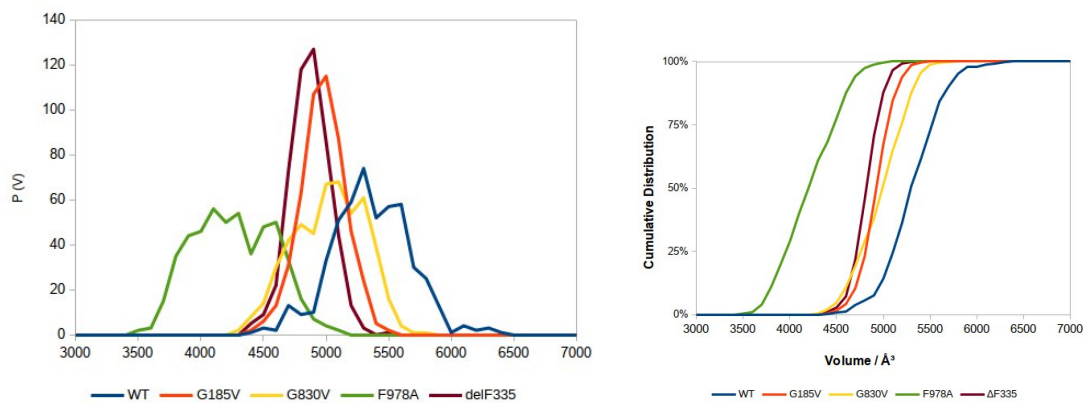
Figure S8. Probability distribution function P(V) and cumulative distribution of the DBP volumes in the human WT P-gp model and variants.

Table S3. Variation in the axis distance (*bun_dist*) observed in the P-gp variants transmembrane helices (compared to the WT; +, positive mean changes; −, negative changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
G185V						+			−			
G830V								−	−			
F978A									−			
ΔF335									−			
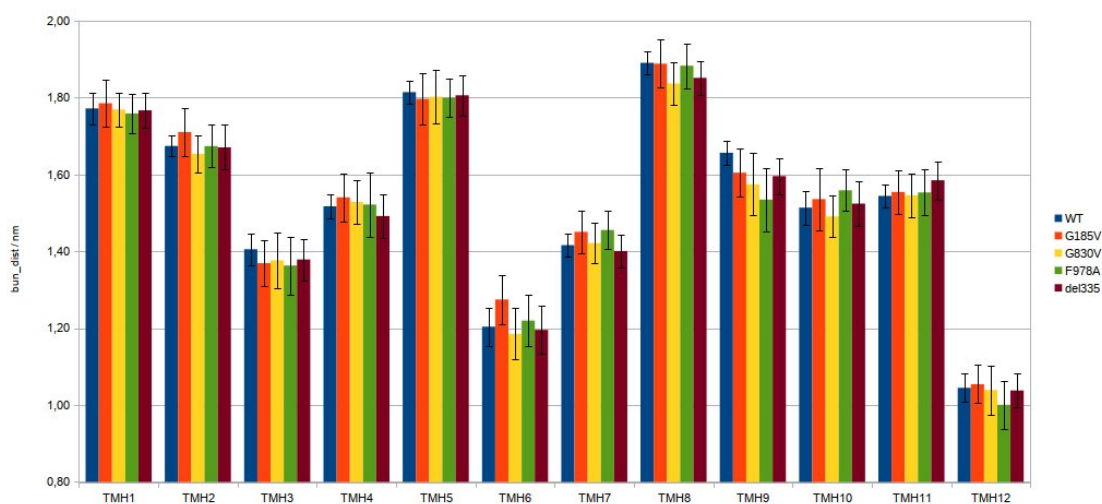
Figure S9. Comparison of individual transmembrane axis distance from the bundle center for all P-gp variants.
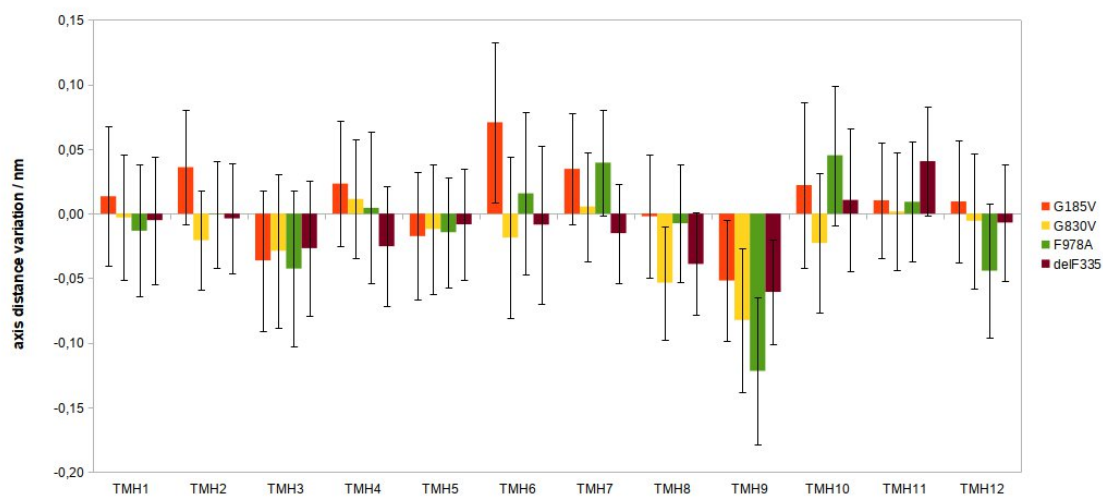
Figure S10. Variation in the transmembrane axis distance from the bundle center for all P-gp variants (difference from WT).

Table S4. Variation in the transmembrane axis length (*bun_len*) observed in all P-gp variants (compared to the WT; +, positive mean changes; –, negative changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
G185V										–	–	–
G830V									+	–		+
F978A						–	+				–	
ΔF335											–	–
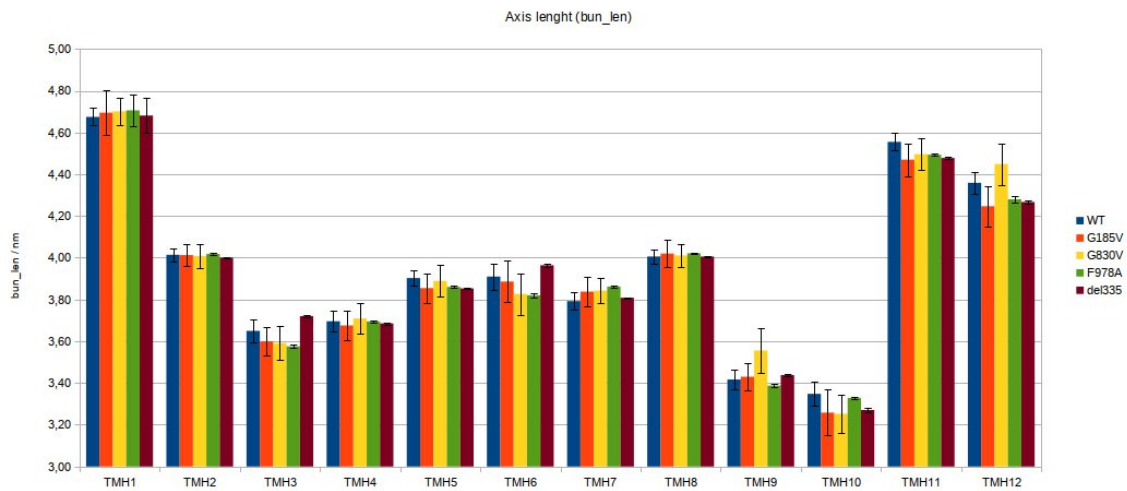
Figure S11. Comparison of individual transmembrane axis length for all P-gp variants.
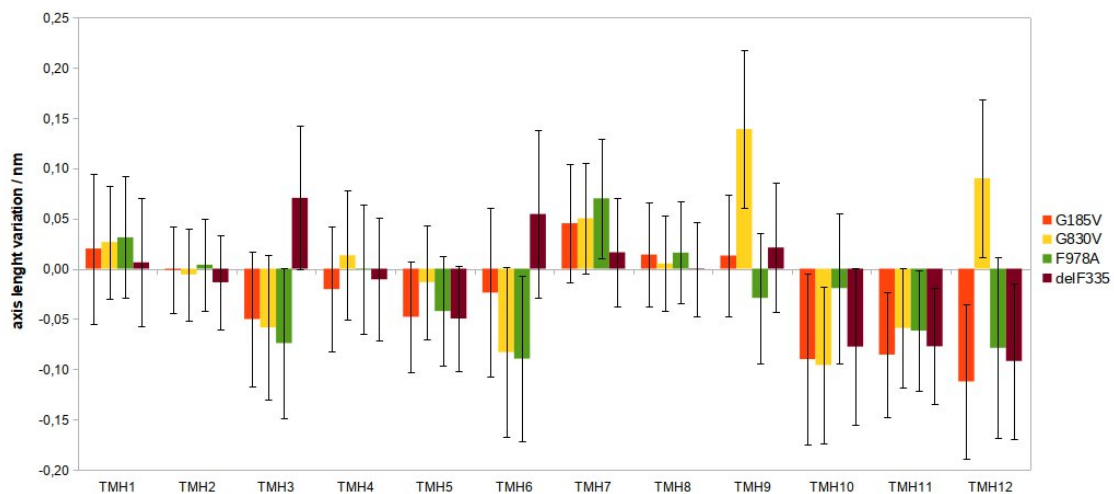
Figure S12. Variation in the transmembrane axis length for all P-gp variants (difference from WT).

Table S5. Variation in the *z*-shift of the axis mid-points (*bun_z*) observed in the P-gp variants transmembrane helices (compared to the WT; +, positive mean changes; −, negative changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
G185V							+			+		
G830V			−	−	−	−	+			+		+
F978A				−			+		−	+		−
ΔF335											−	
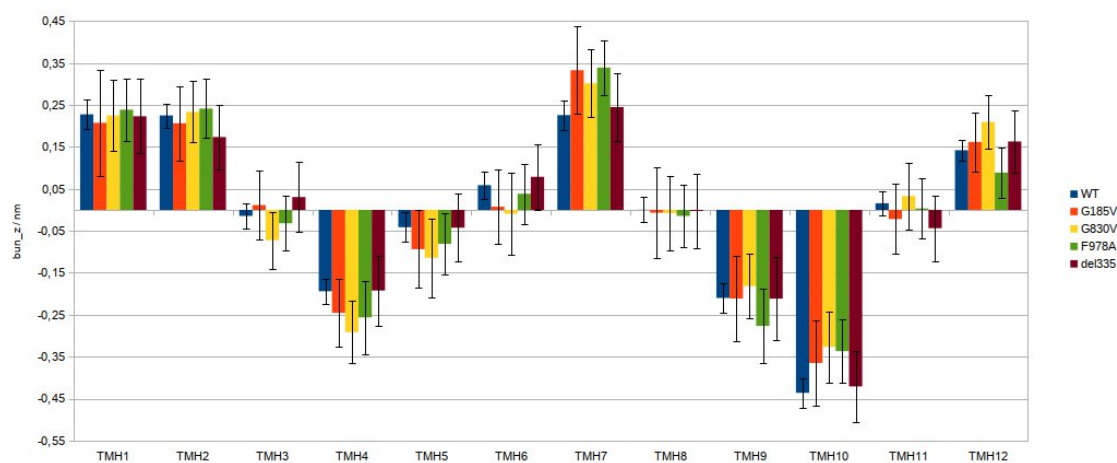
Figure S13. Comparison of individual transmembrane *z*-shifts, from the axis mid-points, for all P-gp variants.
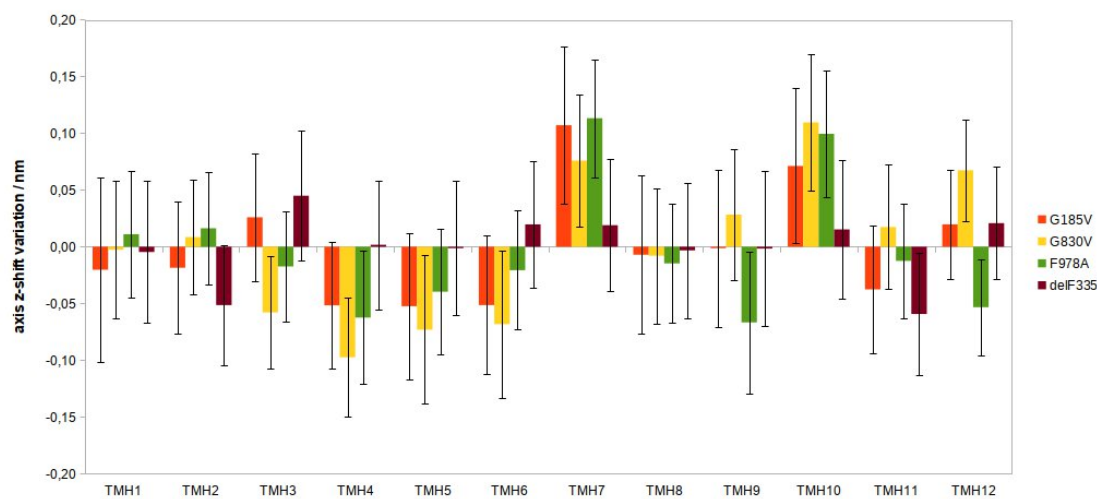
Figure S14. Variation in the transmembrane *z*-shift of the axis mid-points for all P-gp variants (difference from WT).

Table S6. Variation in the axis total tilt (*bun_tilt*) observed in the P-gp variants transmembrane helices (compared to the WT; +, positive mean changes; −, negative changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
G185V			+					+				
G830V		−	+		−			+			−	+
F978A								+				
ΔF335												
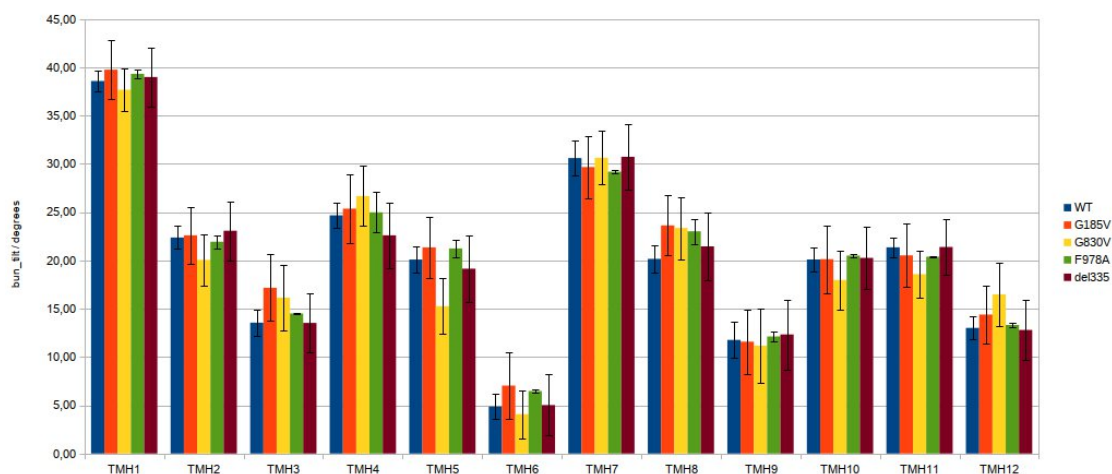
Figure S15. Comparison of individual transmembrane axis total tilt, against *z* axis, for all P-gp variants.
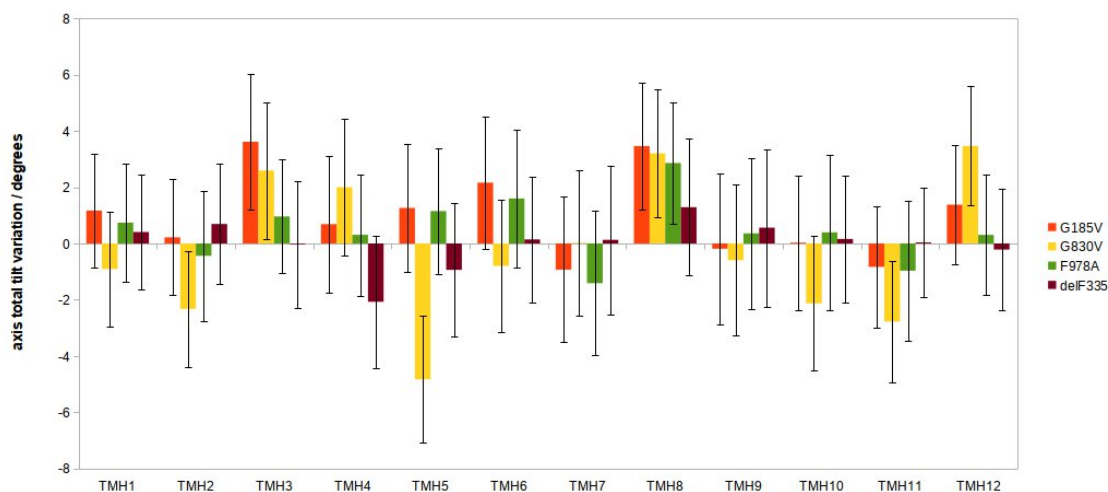
Figure S16. Variation in the transmembrane axis total tilt, against *z* axis, for all P-gp variants (difference from WT).

Table S7. Variation in the axis lateral tilt (*bun_tiltl*) observed in the P-gp variants transmembrane helices (compared to the WT; +, positive mean changes; −, negative changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
G185V	+		+			+		−				
G830V	+	+			−		−	−			+	
F978A	+					+	−	−				
ΔF335				+		+	+					
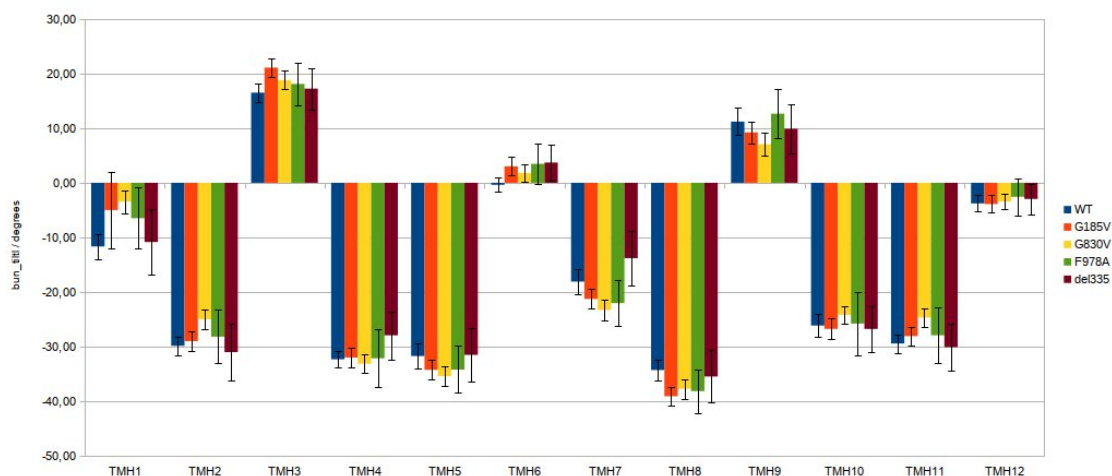
Figure S17. Comparison of individual transmembrane axis lateral tilt, against *z* axis, for all P-gp variants.
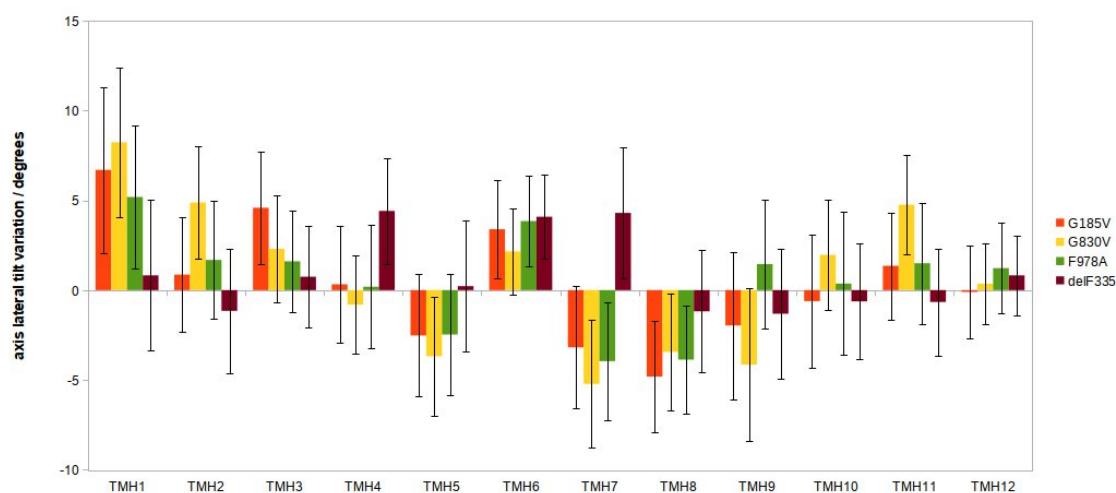
Figure S18. Variation in the transmembrane axis lateral tilt, against *z* axis, for all P-gp variants (difference from WT).

Table S8. Variation in the axis radial tilt (*bun_tiltr*) observed in the P-gp variants transmembrane helices (compared to the WT; +, positive mean changes; −, negative changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
G185V					+							
G830V				+	+	+				−		
F978A					+		+			+		
ΔF335					+							
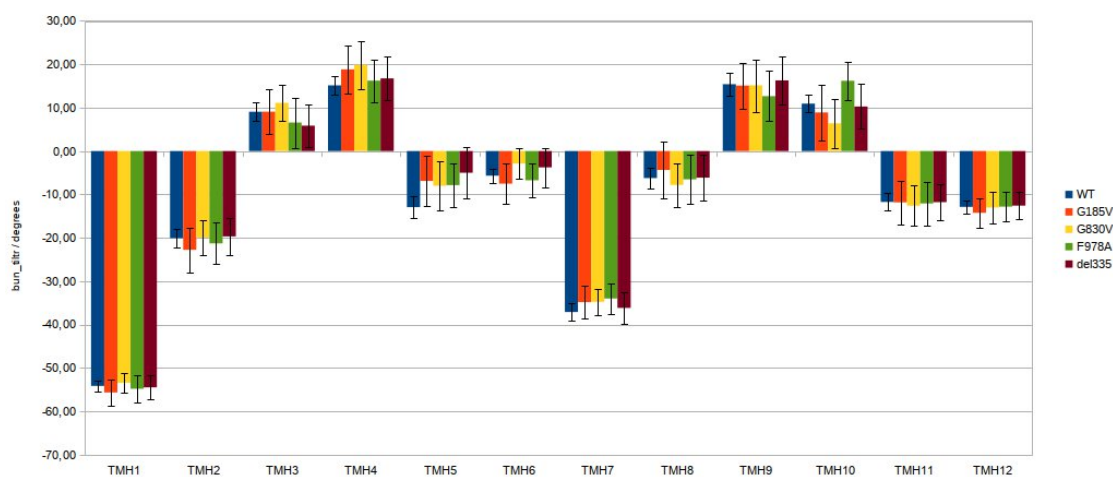
Figure S19. Comparison of individual transmembrane axis lateral tilt, against the bundle axis, for all P-gp variants.
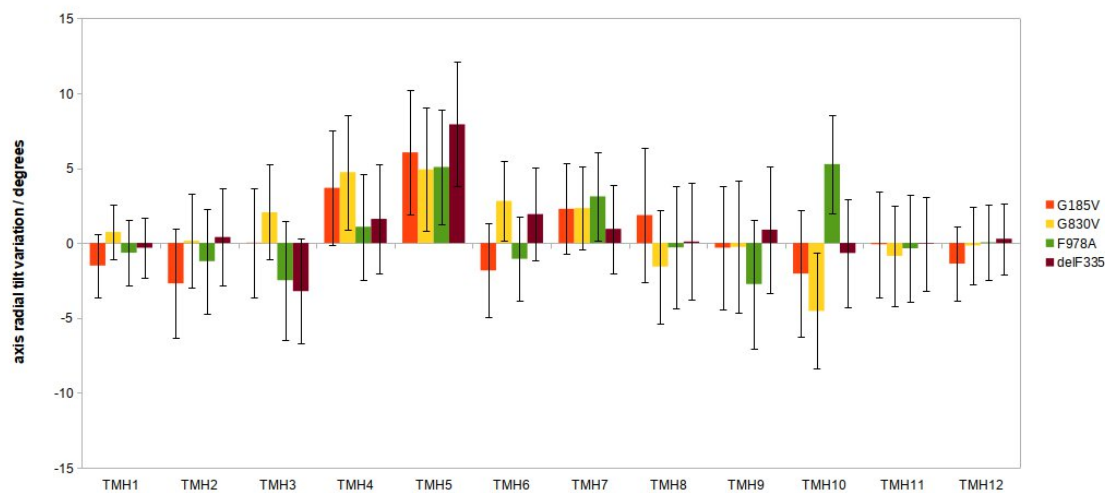
Figure S20. Variation in the transmembrane axis radial tilt, against the bundle axis, for all P-gp variants (difference from WT).

Table S9. Top-ranked binding energies (ΔG) in kcal.mol^{-1} obtained for each site within the DBP of the human WT P-gp model and variants.

SUBSTRATES	P-gp WT (ΔG; kcal.mol-1)			G185V (ΔG; kcal.mol-1)			G830V (ΔG; kcal.mol-1)			F978A (ΔG; kcal.mol-1)			ΔF335 (ΔG; kcal.mol-1)		
	M-site	R-site	H-site	M-site	R-site	H-site	M-site	R-site	H-site	M-site	R-site	H-site	M-site	R-site	H-site
Actinomycin D		-11,2			-11,5			-11,3			-12,4			-11,7	
Amprenavir	-8,6			-8,5			-9,0			-8,2			-8,2		
Bromocriptine	-10,0	-8,9	-8,4	-10,4	-9,5		-11,0			-10,0	-9,6		-10,5	-8,9	-8,9
Calcein-AM	-7,9	-7,8		-8,2	-8,4		-8,7			-8,0	-8,0		-7,7		
Calcein	-6,9	-7,7	-7,1	-7,6	-7,2		-9,4			-7,5	-7,3		-7,7	-7,9	
Chloroquine	-7,2	-6,8		-7,2			-7,7			-6,8			-7,3		
Colchicine	-8,7	-7,4	-6,8	-10,8	-7,0		-8,7	-7,8		-7,7			-8,1	-6,8	
Cyclosporine		-7,9			-9,0			-9,1			-9,6			-8,9	
Daunorubicin	-8,5	-8,3	-8,2	-9,1	-8,8		-8,0	-8,2		-9,2			-9,2	-7,9	
Dexamethasone	-8,5	-8,1	-8,2	-9,3	-8,4			-9,2		-8,9			-8,7		-7,5
Digoxigenin	-8,3	-8,0	-9,0	-8,4	-7,9			-10,3		-9,4			-8,8	-7,8	-7,7
Monotoxoside	-9,8	-8,9	-8,6	-10,1	-9,5			-10,6		-10,3	-9,2		-9,8	-8,6	-8,9
Digitoxoside	-9,1	-10,4	-9,1	-10,3	-9,6			-11,3		-11,1	-10,3		-10,2	-10,0	
Digoxin	-9,9	-9,9	-9,6		-10,7			-11,0		-11,6	-12,0			-11,6	
Diphenhydramine	-7,5			-6,7			-7,3			-7,0			-6,8		-6,1
Doxorubicin	-8,4	-8,3	-8,7	-8,6	-8,0		-10,1			-9,2	-9,2	-8,0	-9,2	-8,5	
Erythromycin		-7,7	-7,9		-7,9		-7,4			-8,3	-8,4		-7,6		
Etoposide	-8,4	-8,2	-8,0	-8,9	-9,3		-9,9			-8,8			-9,0		
Hoechst 33258	-9,9	-9,4	-8,4	-9,5	-9,3			-10,4		-9,5	-10,2		-9,2	-10,9	
Hoechst 33342	-9,7	-8,6		-9,4	-9,1		-10,6	-8,7		-9,7	-9,2		-9,9	-8,4	
Indinavir	-9,4	-8,9		-10,2				-10,4		-9,2			-9,6		
Itraconazole	-10,1	-9,5			-10,1			-10,5			-10,5			-11,1	
Ketoconazole	-9,5	-9,2		-9,5	-8,7		-10,7			-8,8	-9,2		-9,3	-9,0	
LDS-751	-7,6	-7,5		-7,9	-7,0		-8,8			-8,1	-8,1		-7,7	-7,9	-7,8
Methotrexate	-7,8	-7,7	-7,4	-8,3	-8,0		-8,9			-8,2	-8,7		-8,4	-7,9	
Midazolan	-9,1	-7,6	-7,0	-9,6			-9,1			-8,6	-7,5		-8,8		
Prazosin	-7,7	-7,6	-6,8	-8,0	-7,3		-8,8			-7,4	-7,6		-7,5	-8,1	-7,1
Progesterone	-8,5	-7,9	-8,4	-9,5			-10,3			-9,3			-9,3		
Quercetin	-7,9	-7,7	-7,2	-8,1			-7,6			-7,8	-8,1	-7,1	-7,7	-8,3	-7,5
Quinidine	-8,3	-8,1	-7,8	-8,1			-9,1			-7,9			-8,8		
R-propranolol	-7,1			-6,5			-7,8			-6,9			-7,2		-7,0
R-verapamil	-7,3	-6,9		-7,8			-8,7			-7,0			-7,5		
Reserpine	-8,7	-9,2		-9,0	-8,2		-10,2			-8,9	-8,8		-9,3	-8,6	
Rhodamine-123	-8,8		-7,1	-8,4	-7,5		-9,1				-8,1		-7,6		
Ritonavir	-8,3	-8,3	-7,7	-9,5	-8,7		-9,2			-9,1			-8,8		
S-propranolol	-6,8			-6,9			-8,1			-6,7	-6,5		-7,5	-7,2	
S-verapamil	-7,5	-6,9	-6,9	-7,7			-8,4			-7,7			-7,6		
Taxol (Paclitaxel)	-9,1	-8,9	-8,9	-9,0	-8,4		-9,8			-9,8	-9,6		-10,6		-9,8
Trimethoprim	-7,4	-6,5		-6,8			-6,9			-6,5	-6,5		-6,9	-6,3	
Vinblastine		-7,2	-6,8	-8,9	-8,6		-7,4	-7,3	-7,3		-9,0		-9,2		-8,8
Yohimbine	-8,9	-8,3	-7,8	-9,0					-9,4	-8,9	-8,2		-9,0	-8,0	-7,6
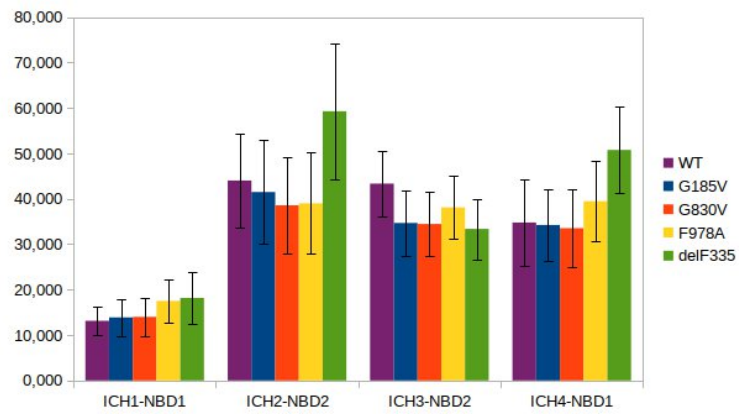
Figure S21. Comparison of total number of contacts at each ICH-NBD interface for the WT model and all P-gp variants (obtained with *gmx hbond*).
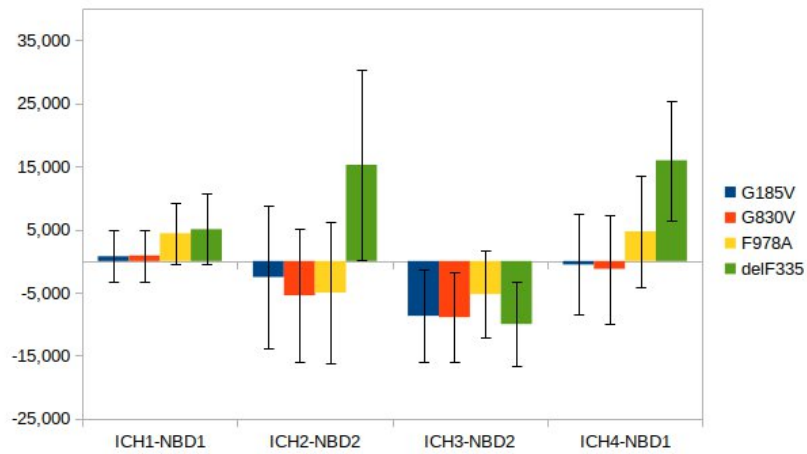
Figure S22. Variation in the total number of contacts found at each ICH-NBD interface, for all P-gp variants, when compared to the WT.

13

Table S10. Contact frequencies between ICH-NBD residues in the human WT P-gp model and variants.

ICH1	NBD1	WT Frequency	G185V Frequency	G185V Variation	G830V Frequency	G830V Variation	F978A Frequency	F978A Variation	delF335 Frequency	delF335 Variation
160 ILE	401 TYR	0,133733	0,069195	-6%	0,100466	-3%	0,065203	-7%	0,089820	-4%
160 ILE	**443 LEU**	0,125749	0,531604	41%	0,345309	22%	0,781770	66%	0,533600	41%
160 ILE	**444 TYR**	0,107784	0,188290	8%	0,157685	5%	0,159016	5%	0,125083	2%
164 ASP	403 SER	0,023952	0,033267	1%	0,041916	2%	0,067864	4%	0,228210	20%
164 ASP	404 ARG	0,990020	0,917498	-7%	0,931470	-6%	0,974717	-2%	0,956753	-3%
164 ASP	405 LYS	0,946108	0,889554	-6%	0,892881	-5%	0,717232	-23%	0,165003	-78%
Hydrogen bonds	life	1218,345	762,167	-37%	687,587	-44%	659,417	-46%	521,357	-94%
Hydrogen bonds	<N>	1,960	2,109	8%	1,998	2%	2,415	23%	2,094	7%
Hydrogen bonds	DG	-22,542	-21,105	-6%	-6,904	-69%	-20,994	-7%	-20,002	-11%

ICH4	NBD1	WT Frequency	G185V Frequency	G185V Variation	G830V Frequency	G830V Variation	F978A Frequency	F978A Variation	delF335 Frequency	delF335 Variation
904 PHE	443 LEU	0,259481	0,326680	7%	0,341317	8,18%	0,354624	10%	0,109115	-15%
905 ARG	**401 TYR**	0,117764	0,230872	11%	0,213573	9,58%	0,395875	28%	0,643380	53%
905 ARG	**434 SER**	0,255489	0,051231	-20%	0,128410	-12,71%	0,041916	-21%	0,021291	-23%
905 ARG	**438 GLN**	0,365269	0,151031	-21%	0,190286	-17,50%	0,358616	-1%	0,727212	36%
905 ARG	**441 GLN**	1,000000	0,699268	-30%	1,000000	0,00%	0,999335	0%	0,996008	0%
906 THR	441 GLN	0,159681	0,160346	0%	0,121757	-3,79%	0,149035	-1%	0,214904	6%
906 THR	474 SER	0,203593	0,179641	-2%	0,038589	-16,50%	0,384564	18%	0,136394	-7%
907 VAL	480 PHE								0,306720	new
908 VAL	**467 ARG**	0,548902	0,783766	23%	0,715902	16,70%	0,684631	14%	0,767132	22%
909 SER	**441 GLN**	0,986028	0,928144	-6%	0,931471	-5,46%	0,992016	1%	0,995343	1%
909 SER	**467 ARG**	0,069860	0,218230	15%	0,105788	3,59%	0,089820	2%	0,114438	4%
909 SER	**472 VAL**	0,151697	0,143047	-1%	0,139721	-1,20%	0,159015	1%	0,282768	13%
910 LEU	467 ARG	0,233533	0,420492	19%	0,542914	30,94%	0,000665	-23%	0,003992	-23%
910 LEU	547 ARG								0,413174	new
911 THR	467 ARG	0,293413	0,000665	-29%	0,060545	-23,29%	0,781770	49%	0,893546	60%
912 GLN	464 ARG	0,311377	0,460413	15%	0,393879	8,25%	0,352628	4%	0,321357	1%
912 GLN	467 ARG	0,231537	0,598137	37%	0,139721	-9,18%	0,195609	-4%	0,083832	-15%
913 GLU	464 ARG								0,038589	new
914 GLN	464 ARG								0,212242	new
Hydrogen bonds	life	682,765	311,115	-54%	392,277	-42,55%	689,296	1%	366,428	-46%
Hydrogen bonds	<N>	2,760	2,084	-24%	2,371	-14,08%	2,805	2%	3,622	31%
Hydrogen bonds	DG	-21,084	-19,012	-10%	-19,661	-6,75%	-20,876	-1%	-19,374	-8%

ICH2	NBD2	WT Frequency	G185V Frequency	G185V Variation	G830V Frequency	G830V Variation	F978A Frequency	F978A Variation	delF335 Frequency	delF335 Variation
262 ARG	1044 TYR	0,331337	0,258150	-7%	0,223553	-11%	0,242182	-9%	0,202262	-13%
262 ARG	1077 SER	0,522954	0,456420	-7%	0,427146	-10%	0,280106	-24%	0,384564	-14%
262 ARG	1081 GLN	0,526946	0,417831	-11%	0,533599	1%	0,394544	-13%	0,519627	-1%
262 ARG	1086 PHE	0,399202	0,538257	14%	0,522954	12%	0,507651	11%	0,326015	-7%
262 ARG	1200 ASP	0,081836	0,067864	-1%	0,075848	-1%	0,039255	-4%	0,222222	14%
263 THR	1117 SER	0,642715	0,604125	-4%	0,475050	-17%	0,705922	6%	0,473054	-17%
263 THR	1118 GLN	0,023952	0,162342	14%	0,031936	1%	0,053892	3%	0,015968	-1%
263 THR	1200 ASP	0,033932	0,224218	19%	0,030605	0%	0,057884	2%	0,014637	-2%
265 ILE	1086 PHE	0,279441	0,175649	-10%	0,244178	-4%	0,206920	-7%	0,292748	1%
265 ILE	1110 ARG	0,754491	0,592149	-16%	0,477711	-28%	0,592149	-16%	0,350632	-40%
266 ALA	1086 PHE	0,570858	0,500998	-7%	0,431138	-14%	0,522954	-5%	0,383899	-19%
267 PHE	1110 ARG	0,560878	0,419162	-14%	0,299401	-26%	0,465070	-10%	0,288091	-27%
267 PHE	1113 LEU		0,033932	new	0,230872	23%	0,023287	2%	0,629408	63%
267 PHE	1114 GLY				0,055888	6%	0,017964	2%	0,113107	11%
267 PHE	1115 ILE	0,932136	0,888889	-4%	0,836993	-10%	0,898869	-3%	0,819028	-11%
267 PHE	1134 GLY								0,226879	new
267 PHE	1188 ARG								0,474385	new
267 PHE	1189 ALA								0,271457	new
267 PHE	1192 ARG								0,640053	new
268 GLY	1110 ARG	0,315369	0,255489	-6%	0,133067	-18%	0,146374	-17%	0,044578	-27%
269 GLY	1136 ASN								0,622754	62%
Hydrogen bonds	life	391,839	534,687	36%	482,332	23%	679,153	73%	465,678	19%
Hydrogen bonds	<N>	1,417	1,340	-5%	1,154	-19%	1,304	-8%	2,215	56%
Hydrogen bonds	DG	-19,685	-20,864	6%	-20,177	3%	-21,062	7%	-19,888	1%

ICH3	NBD2	WT Frequency	G185V Frequency	G185V Variation	G830V Frequency	G830V Variation	F978A Frequency	F978A Variation	delF335 Frequency	delF335 Variation
799 GLN	1086 PHE	0,281437	0,292082	1,06%	0,329341	4,79%	0,371257	8,98%	0,125749	-15,57%
800 ASP	1086 PHE	0,734531	0,762475	2,79%	0,774451	3,99%	0,805057	7,05%	0,810379	7,58%
800 ASP	1087 TYR	0,992016	0,990685	-0,13%	0,957419	-3,46%	0,997339	0,53%	0,968729	-2,33%
801 VAL	1044 TYR		0,015968	new	0,124418	12,44%	0,036593	3,66%	0,066533	6,65%
801 VAL	1086 PHE	0,920160	0,868929	-5,12%	0,860279	-5,99%	0,850965	-6,92%	0,887559	-3,26%
801 VAL	1087 TYR	0,560878	0,319361	-24,15%	0,155689	-40,52%	0,214903	-34,60%	0,090486	-47,04%
802 SER	1087 TYR	0,704591	0,473719	-23,09%	0,360612	-34,40%	0,357286	-34,73%	0,234864	-46,97%
805 ASP	1044 TYR	0,908184	0,692615	-21,56%	0,576846	-33,13%	0,629408	-27,88%	0,471723	-43,65%
805 ASP	1045 PRO	0,069860	0,025283	-4,46%	0,019960	-4,99%	0,049900	-2,00%	0,031271	-3,86%
805 ASP	1046 THR	1,000000	0,899534	-10,05%	0,962741	-3,73%	0,996008	-0,40%	0,997339	-0,27%
805 ASP	1047 ARG	0,097804			0,051231	-4,66%	0,441118	34,33%	0,472388	37,46%
Hydrogen bonds	life	533,367	954,057	78,87%	835,569	56,66%	512,837	-3,85%	389,749	-26,93%
Hydrogen bonds	<N>	2,625	2,308	-12,08%	2,673	1,83%	2,623	-0,08%	2,650	0,94%
Hydrogen bonds	DG	-20,461	-21,634	5,73%	-21,322	4,21%	-6,755	-66,99%	-19,667	-3,88%

The variation percentage is estimated based on the mean frequencies of the three replica systems for each P-gp variant against the WT.
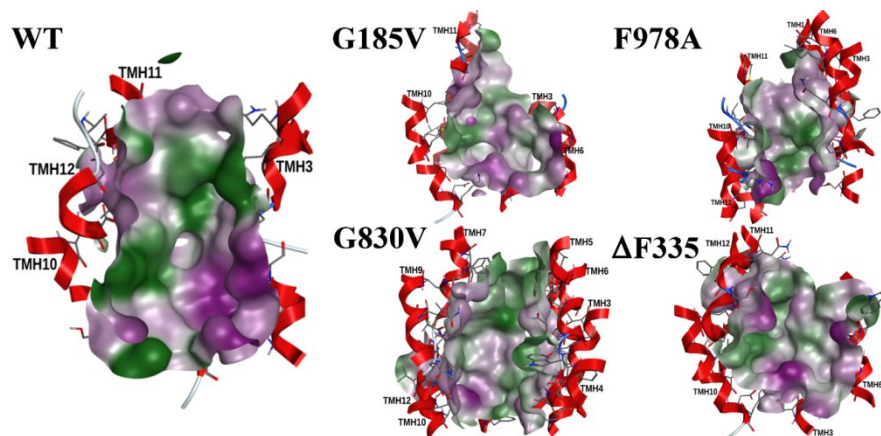
Figure S23. Graphical representation of the molecular surface for the H-site in the human WT P-gp model and variants. Polar regions are defined as pink, hydrophobic regions are colored green.
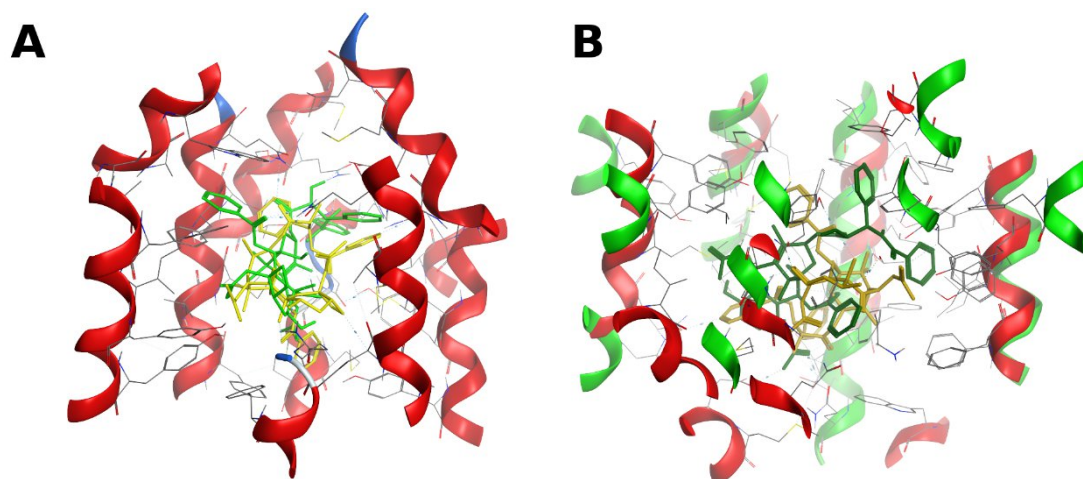
Figure S24. Docking results for taxol molecule in the 6QEX (green, top-ranked pose; yellow, cryo-EM taxol molecule) and in 6QEX (red) and homology (green) models (green, top-ranked pose; yellow, cryo-EM taxol molecule) models.

ANNEX S3

"Long-range communication between transmembrane- and nucleotide-binding

domains does not depend on drug binding to mutant P-glycoprotein."

TABLE OF CONTENTS

Table S1. Variation in the axis distance observed in the *holo* G185V variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative mean changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
G185V-H_COL-H	ns*	ns	+	+	-	ns	+	-	-	+	-	+
G185V-H_VIN-H	ns	ns	+	+	-	ns	+	-	-	+	-	+

* ns, not significant
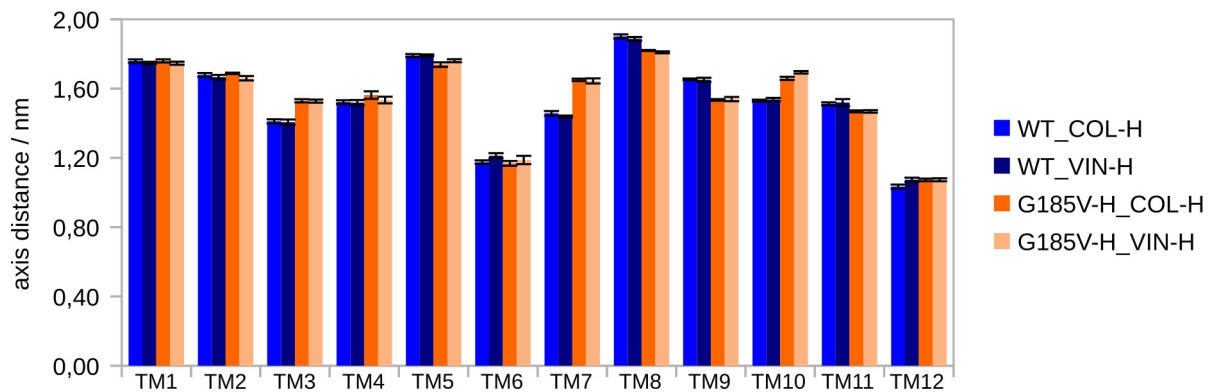
Figure S1: Comparison of individual transmembrane axis distance for the WT and G185V variant (*holo* systems).
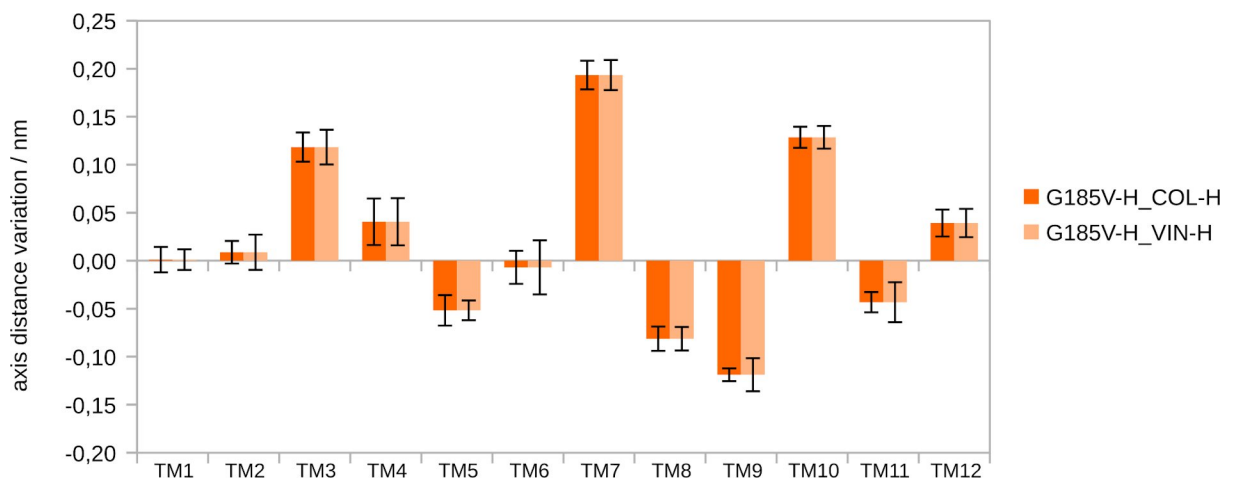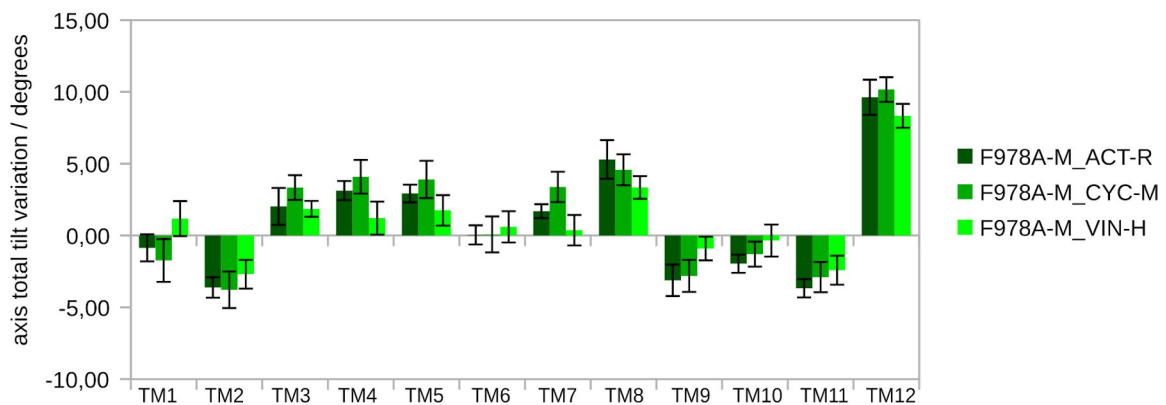

Figure S2: Variation in the transmembrane axis distance for the *holo* G185V variant (difference from the *holo* WT).

Table S2. Variation in the axis length observed in the *holo* G185V variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative mean changes).

|	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
G185V-H_COL-H	+	-	ns	+	-	-	-	+	+	+	-	ns
G185V-H_VIN-H	ns*	-	+	+	-	ns	-	ns	+	+	-	ns

* ns, not significant
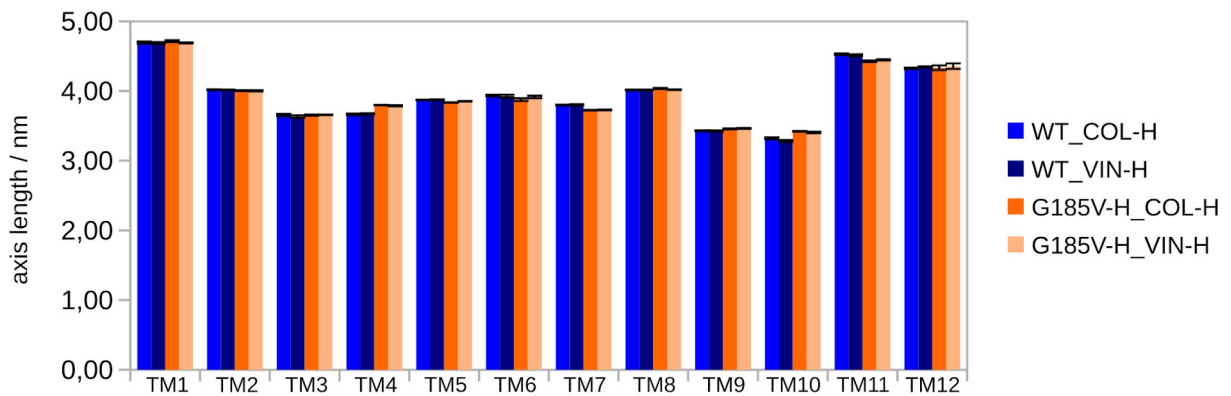
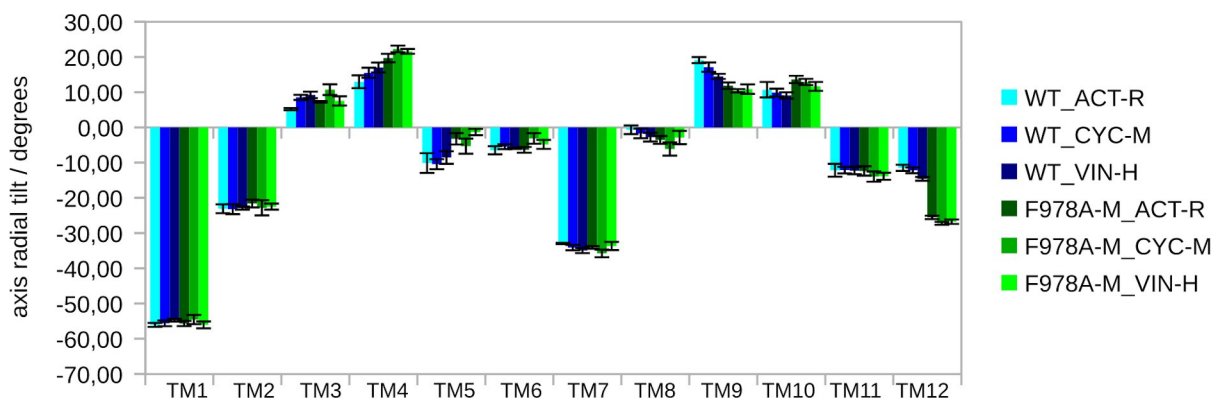**Figure S3:** Comparison of individual transmembrane axis length for the WT and G185V variant (*holo* systems).
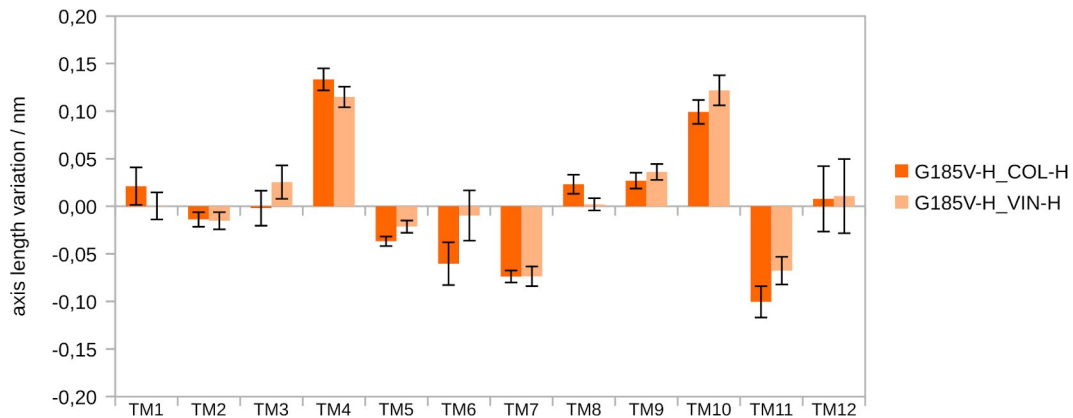
Figure S4: Variation in the transmembrane axis length for the *holo* G185V variant (difference from the *holo* WT).

¶
¶
¶

Table S3. Variation in the *z*-shift of the axis mid-points observed in the *holo* G185V variant transmembrane helices (compared to *holo* WT; +, positive mean changes; –, negative mean changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
G185V-H_COL-H	–	+	ns*	–	ns	–	+	–	+	+	+	–
G185V-H_VIN-H	–	+	ns	–	ns	–	+	–	+	+	+	–

* ns, not significant
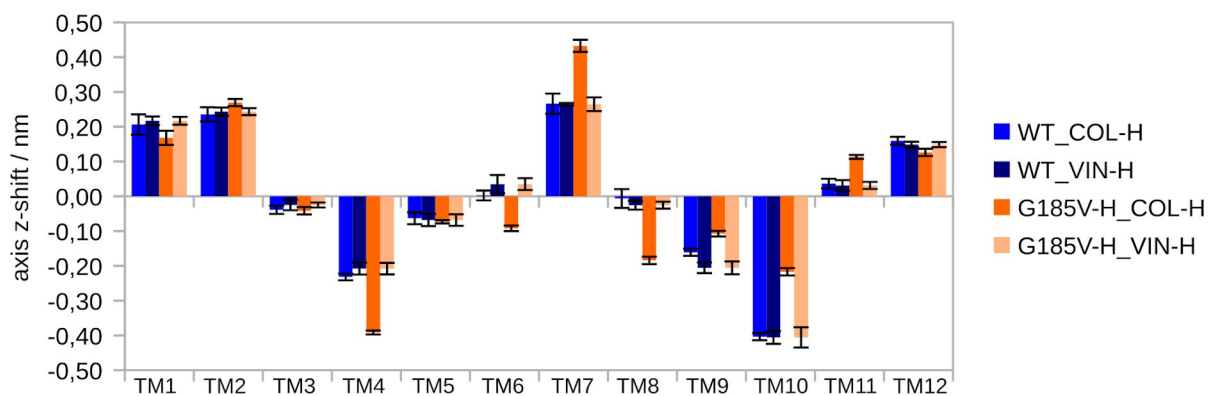
Figure S5: Comparison of individual transmembrane z-shift of the axis mid-points for the WT and G185V variant (*holo* systems).
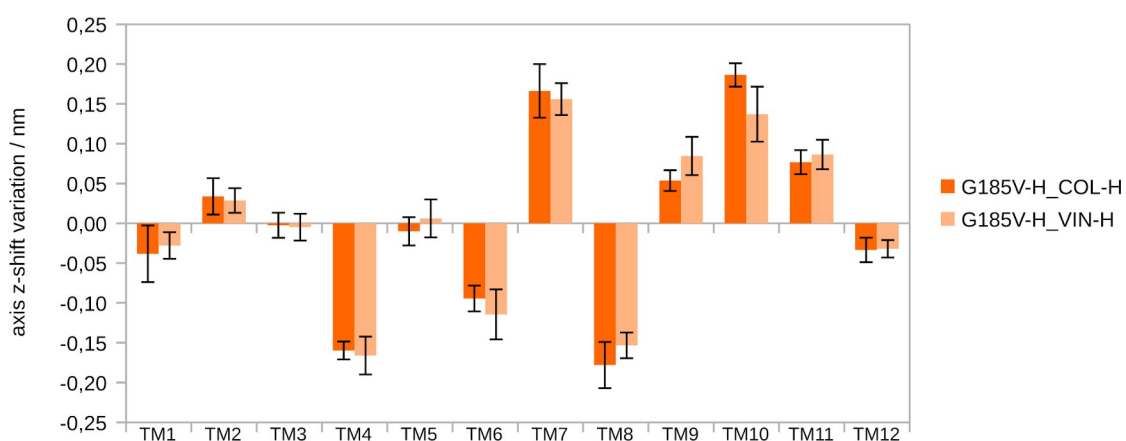
Figure S6: Variation in the transmembrane *z*-shift of the axis mid-points for the *holo* G185V variant (difference from the *holo* WT).

Table S4. Variation in the the axis total tilt observed in the *holo* G185V variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative mean changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
G185V-H_COL-H	ns*	–	+	+	+	+	ns	+	+	+	ns	–
G185V-H_VIN-H	ns	–	+	+	+	+	ns	+	+	+	ns	–

* ns, not significant
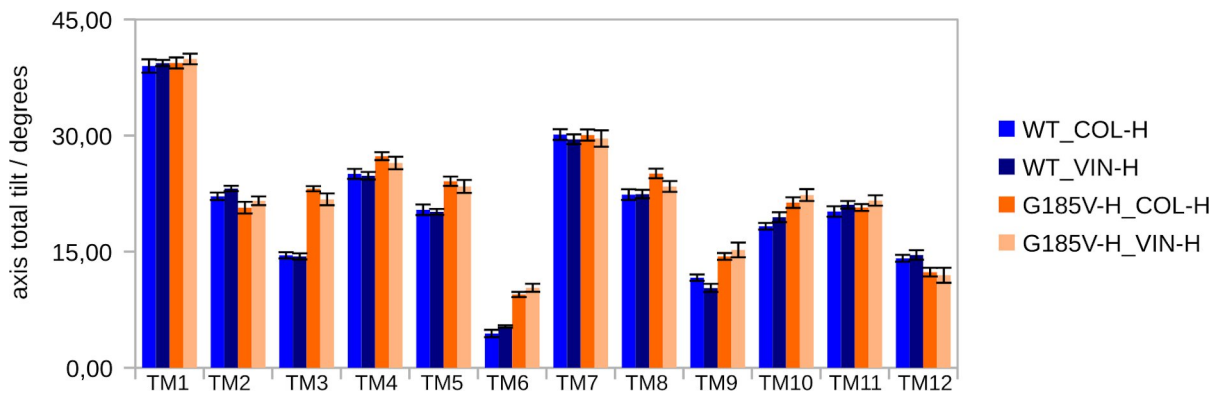
Figure S7: Comparison of individual transmembrane axis total tilt for the WT and G185V variant (*holo* systems).
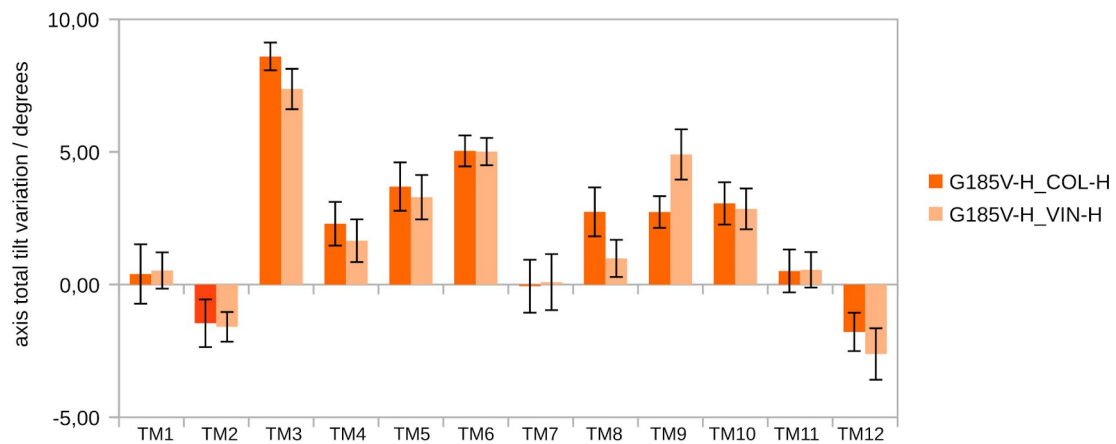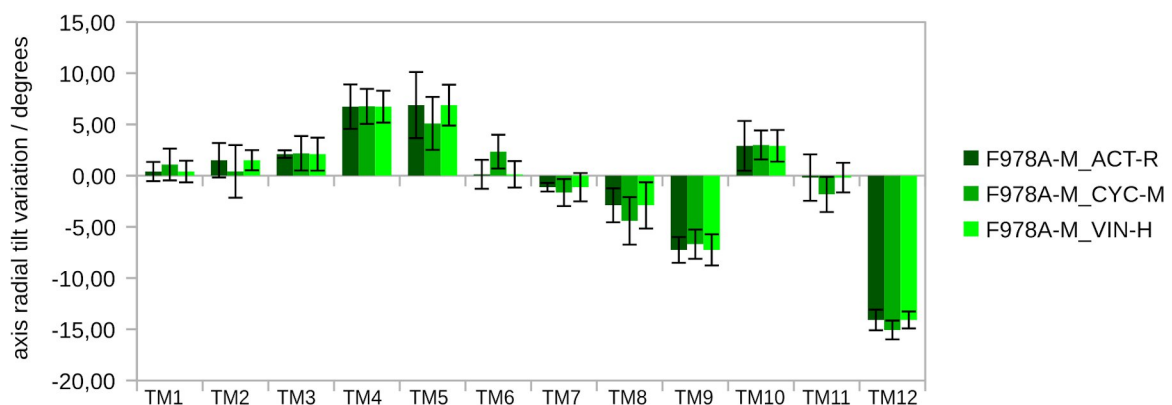

Figure S8: Variation in the transmembrane axis total tilt for the *holo* G185V variant (difference from the *holo* WT).

Table S5. Variation in the axis lateral tilt observed in the *holo* G185V variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative mean changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
G185V-H_COL-H	+	+	+	–	–	+	–	–	+	–	+	ns
G185V-H_VIN-H	+	+	+	ns*	+	+	–	ns	+	–	+	ns

* ns, not significant
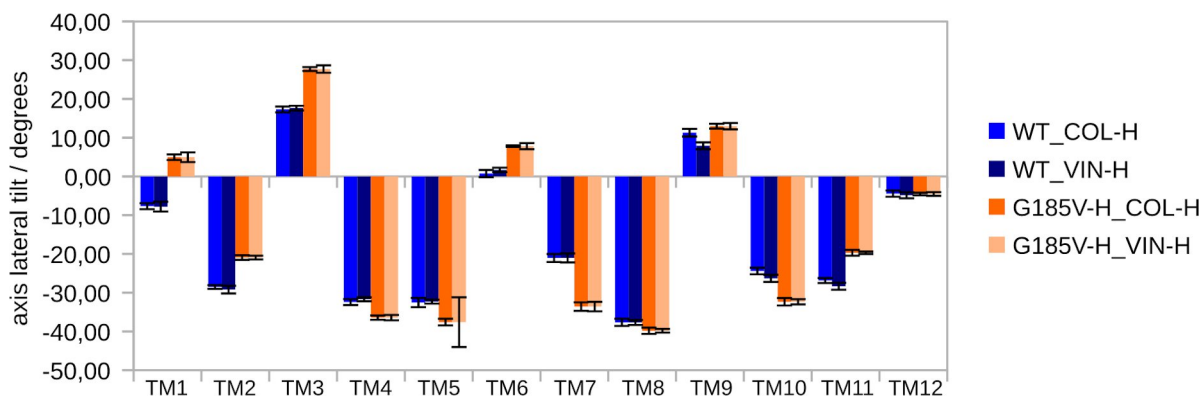
Figure S9: Comparison of individual transmembrane axis lateral tilt for the WT and G185V variant (*holo* systems).
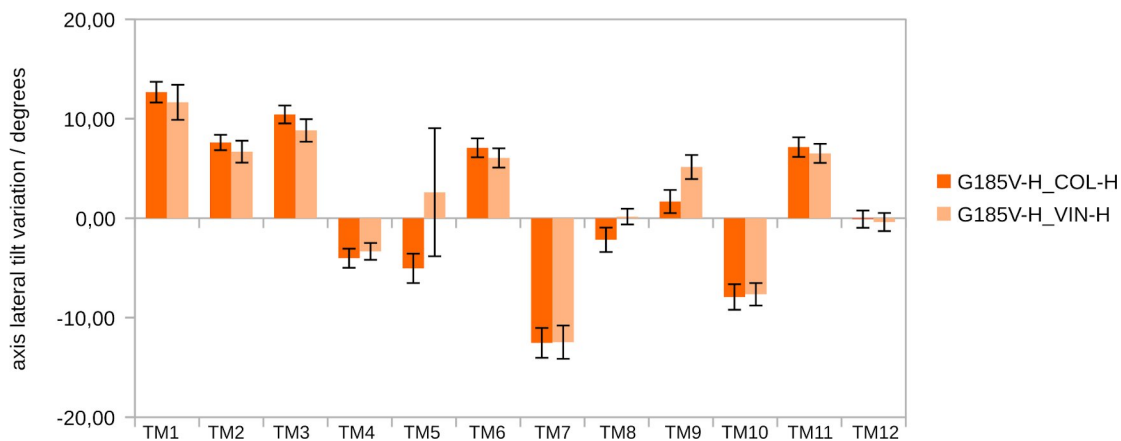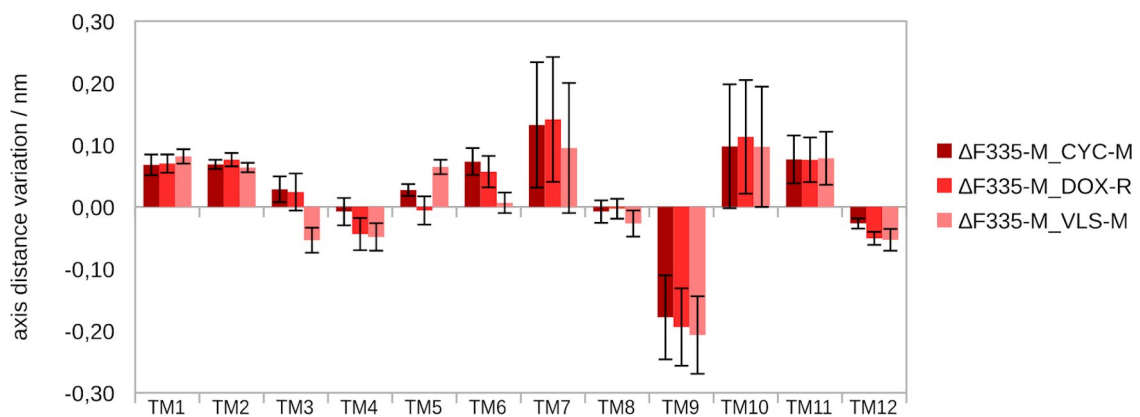

Figure S10: Variation in the transmembrane axis lateral tilt for the *holo* G185V variant (difference from the *holo* WT).

Table S6. Variation in the axis radial tilt observed in the *holo* G185V variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative changes).

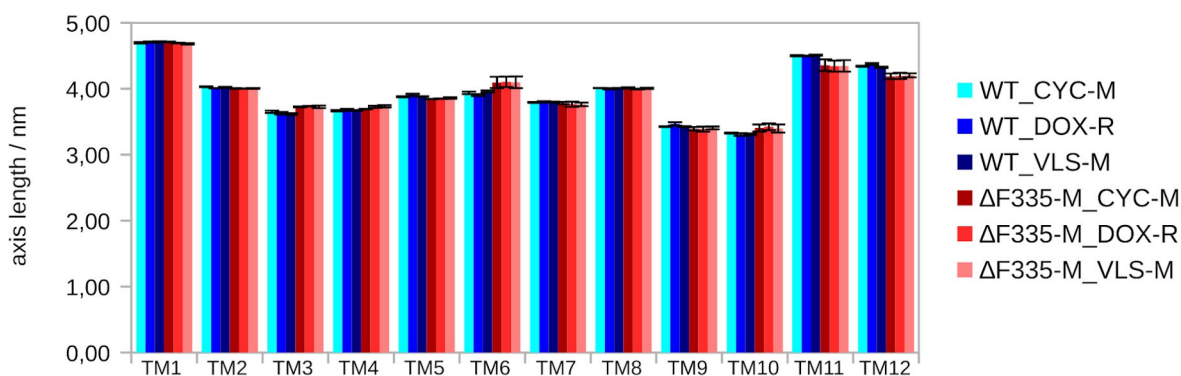|	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
G185V-H_COL-H	ns*	–	+	–	+	–	+	–	–	–	–	+
G185V-H_VIN-H	ns	–	+	–	+	–	+	–	–	–	–	+

* ns, not significant
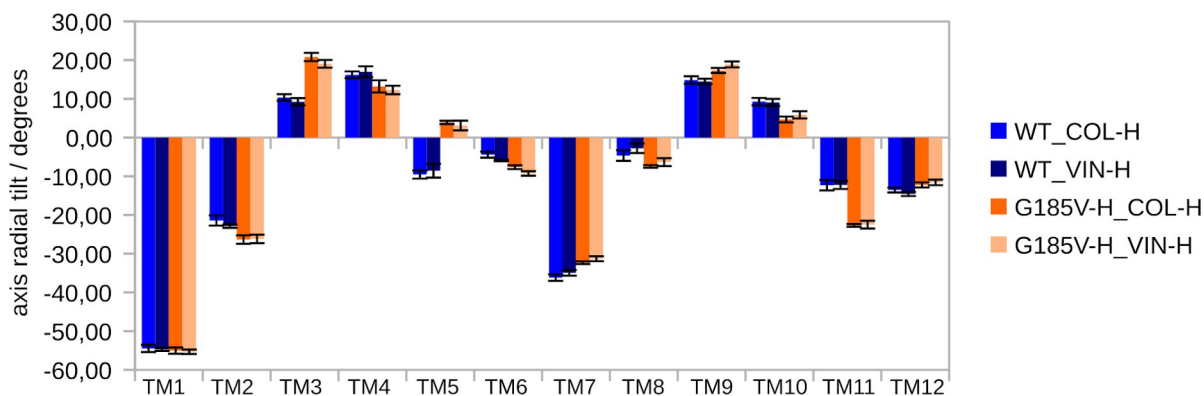
Figure S11: Comparison of individual transmembrane axis radial tilt for the WT and G185V variant (*holo* systems).
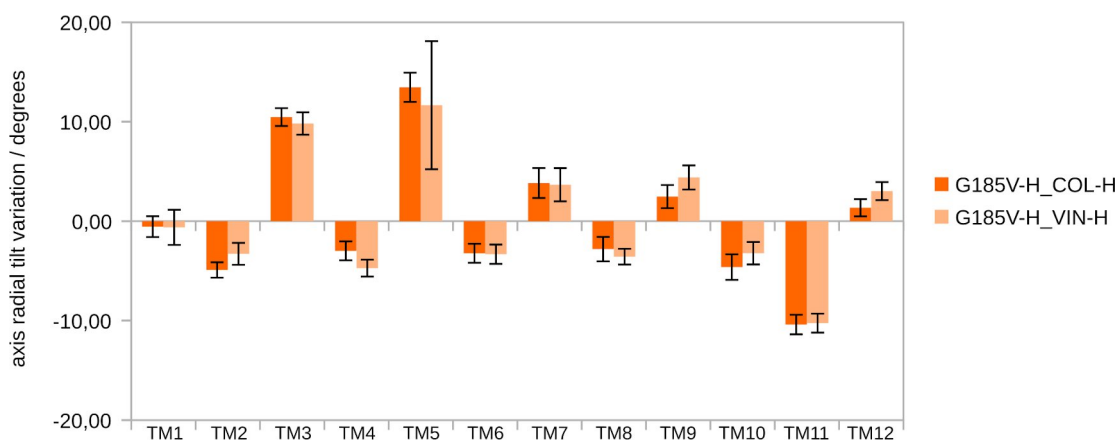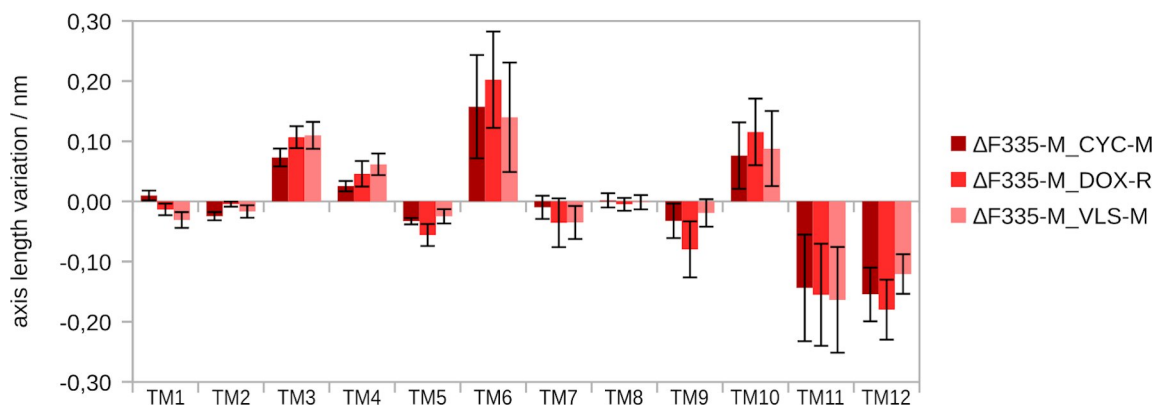

Figure S12: Variation in the transmembrane axis radial tilt for the *holo* G185V variant (difference from the *holo* WT).

Table S7. Variation in the axis distance observed in the *holo* G830V variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative mean changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
G830V-R_ACT-R	ns*	–	ns	ns	ns	+	ns	–	+	+	–	+
G830V-R_DOX-R	ns	–	ns	ns	–	+	+	–	ns	+	–	–

* ns, not significant
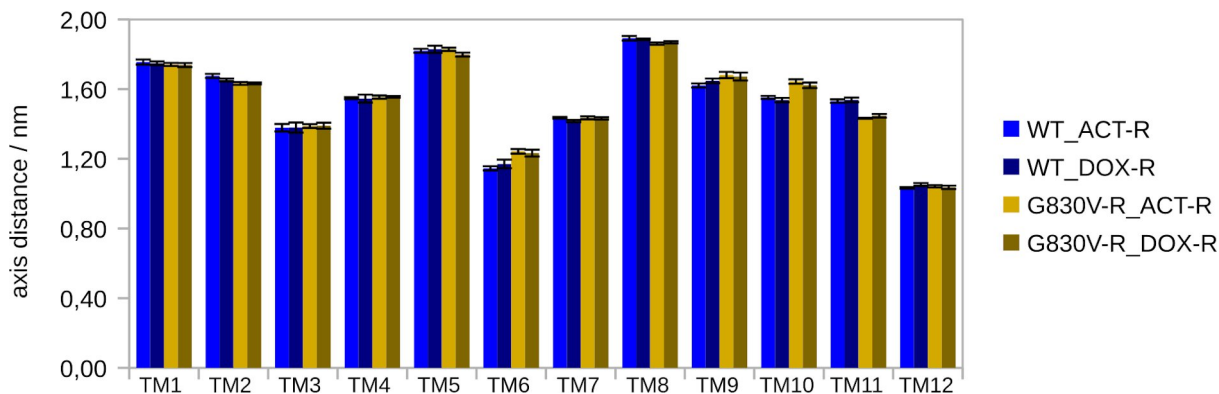
Figure S13: Comparison of individual transmembrane axis distance for the WT and G830V variant (*holo* systems).
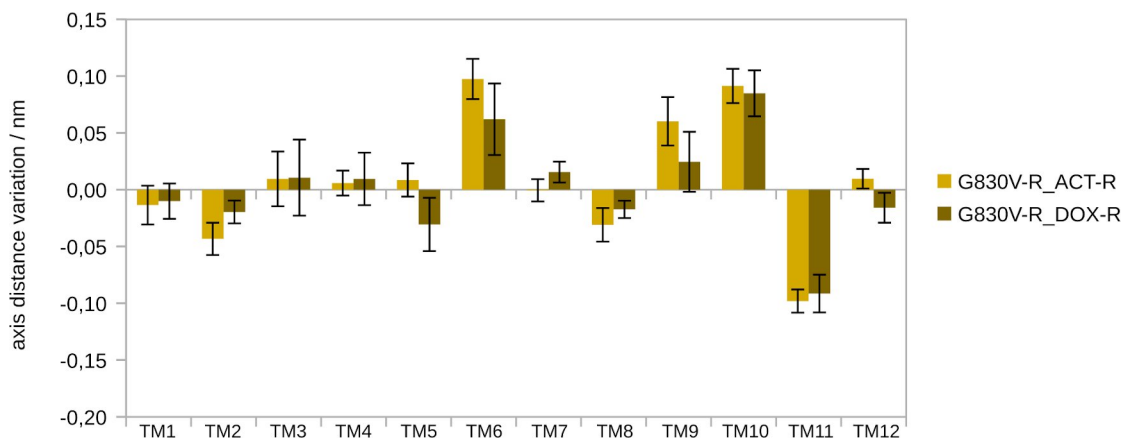
Figure S14: Variation in the transmembrane axis distance for the *holo* G830V variant (difference from the *holo* WT).

Table S8. Variation in the axis length observed in the *holo* G830V variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative mean changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
G830V-R_ACT-R	+	–	–	+	ns	–	+	+	+	+	–	+
G830V-R_DOX-R	ns*	ns	–	ns	–	–	ns	ns	+	+	–	+

* ns, not significant
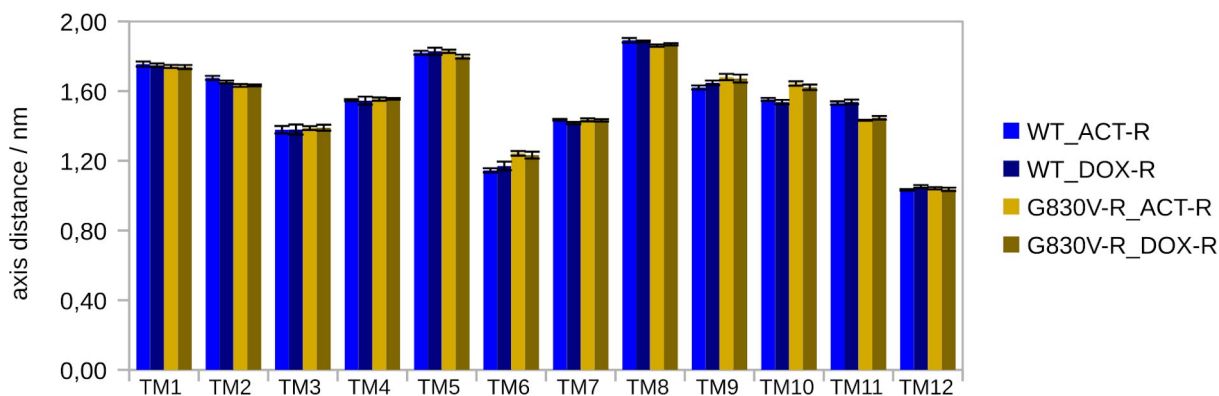
Figure S15: Comparison of individual transmembrane axis length for the WT and G830V variant (*holo* systems).
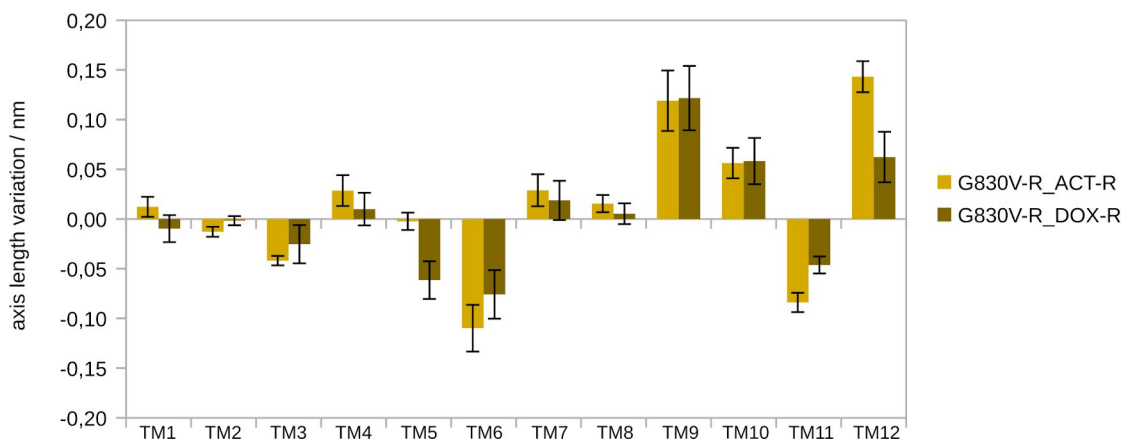
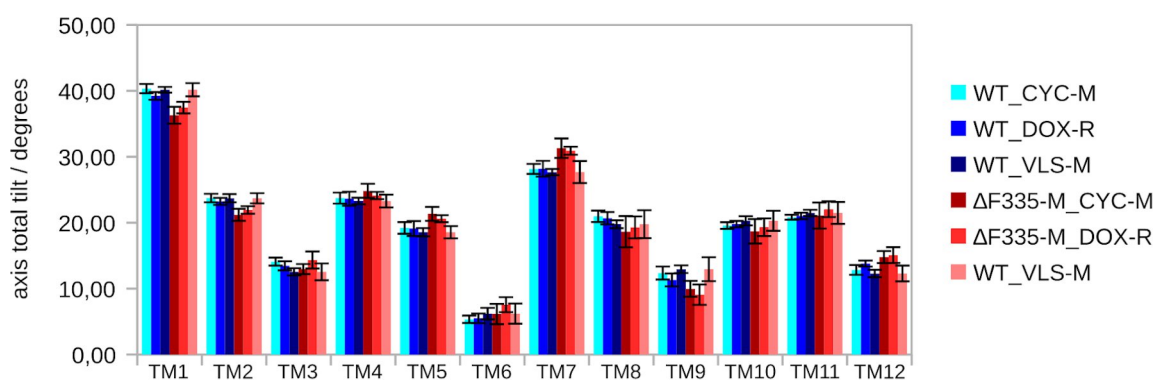**Figure S16:** Variation in the transmembrane axis length for the *holo* G830V variant (difference from the *holo* WT).

Table S9. Variation in the *z*-shift of the axis mid-points observed in the *holo* G830V variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative mean changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
G830V-R_ACT-R	–	–	–	–	+	ns	+	+	+	+	ns	ns
G830V-R_DOX-R	–	ns*	–	–	–	–	+	+	+	+	+	+

* ns, not significant
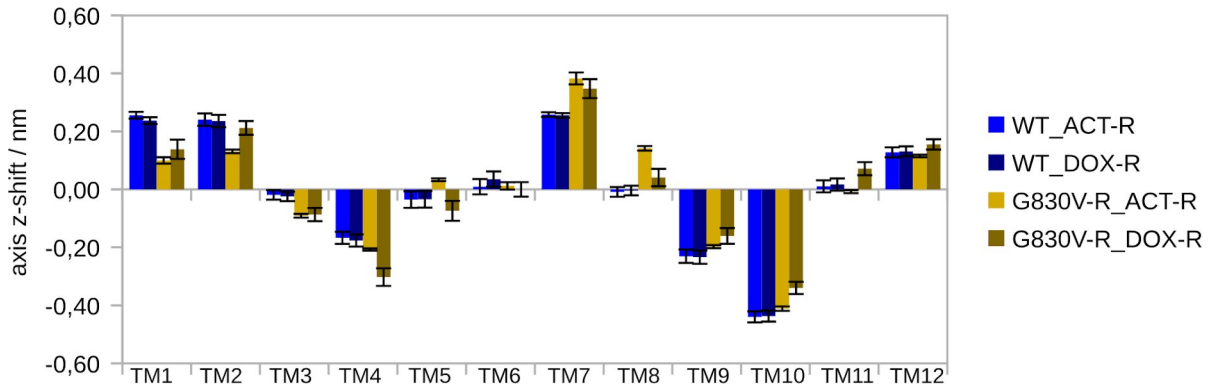
Figure S17: Comparison of individual transmembrane *z*-shift of the axis mid-points for the WT and G830V variant (*holo* systems).
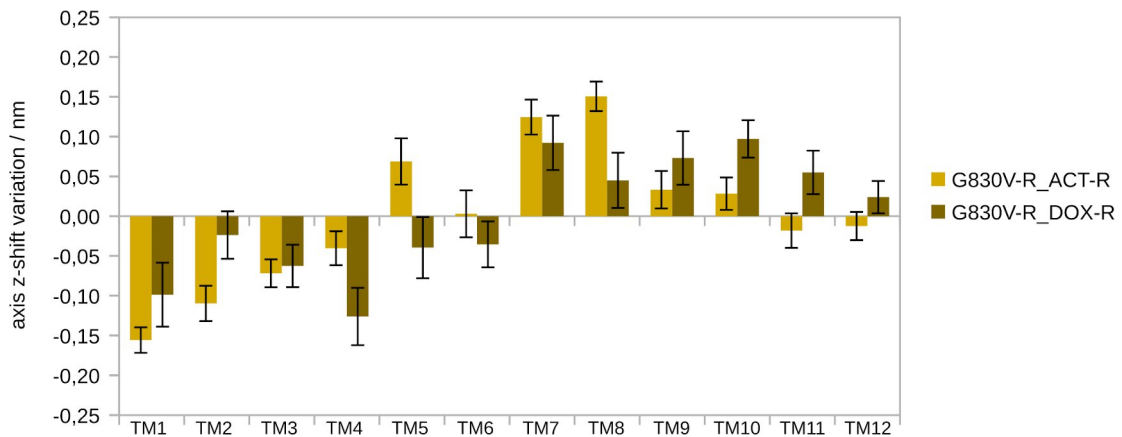
Figure S18: Variation in the transmembrane *z*-shift of the axis mid-points for the *holo* G830V variant (difference from the *holo* WT).

Table S10. Variation in the the axis total tilt observed in the *holo* G830V variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative mean changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
G830V-R_ACT-R	–	–	+	+	+	–	+	+	+	–	–	–
G830V-R_DOX-R	–	–	+	+	+	–	+	+	+	ns*	–	–

* ns, not significant
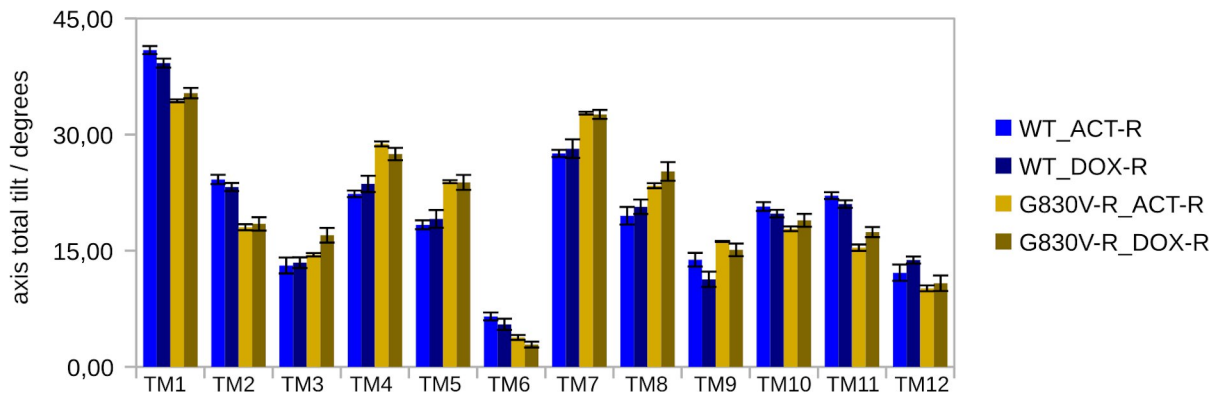
Figure S19: Comparison of individual transmembrane axis total tilt for the WT and *holo* G830V variant (*holo* systems).
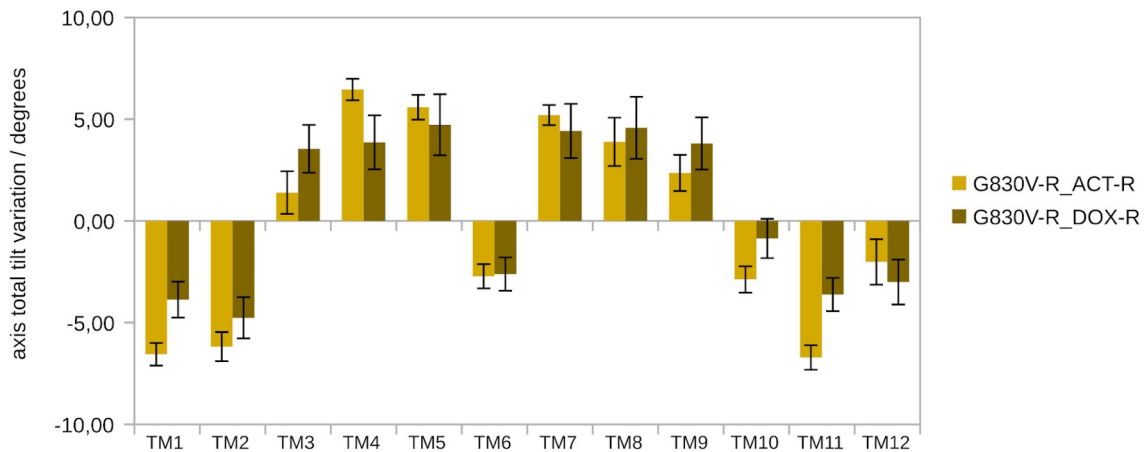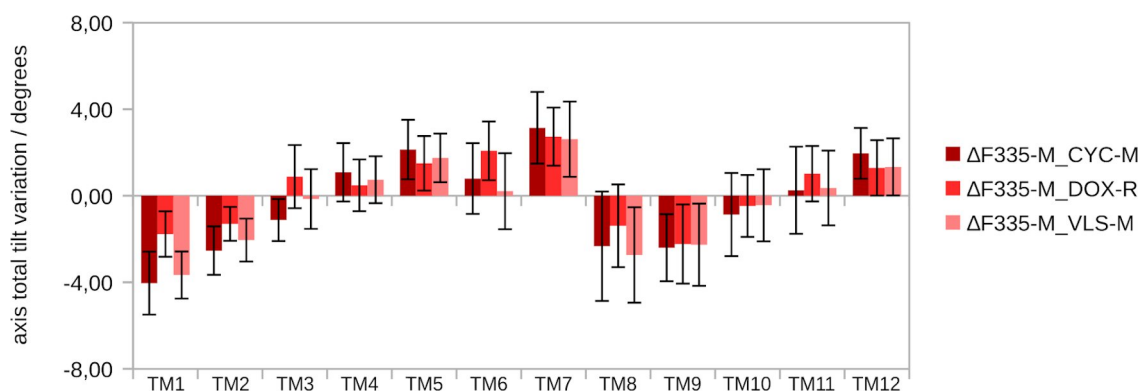

Figure S20: Variation in the transmembrane axis total tilt for the *holo* G830V variant (difference from the *holo* WT).

Table S11. Variation in the axis lateral tilt observed in the *holo* G830V variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative mean changes).

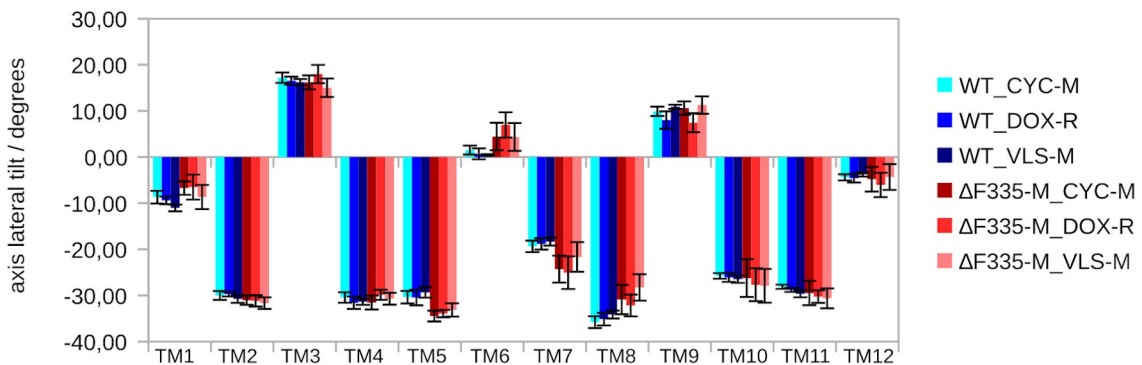|	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
G830V-R_ACT-R	+	+	–	–	–	–	–	–	+	+	+	+
G830V-R_DOX-R	+	+	+	–	–	ns*	–	–	+	–	+	+

* ns, not significant
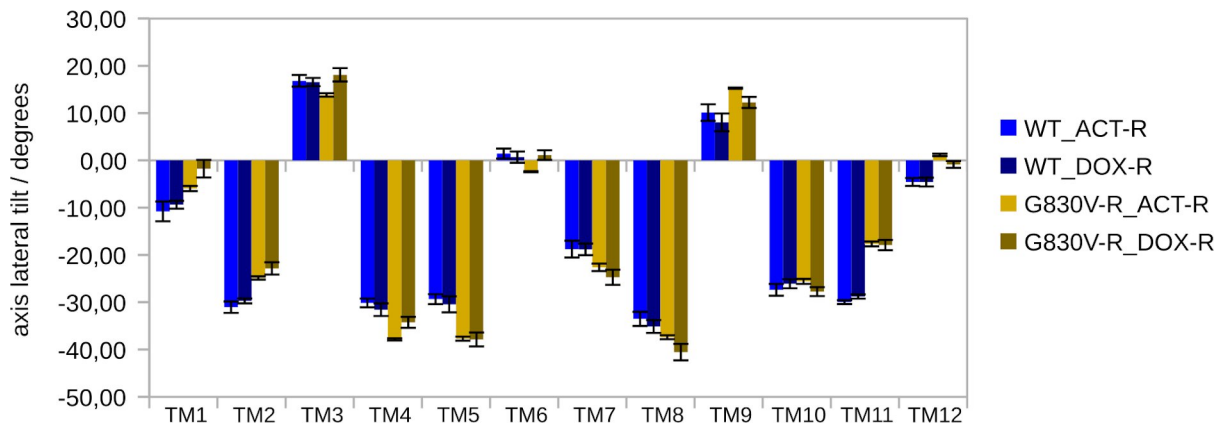
Figure S21: Comparison of individual transmembrane axis lateral tilt for the WT and G830V variant (*holo* systems).
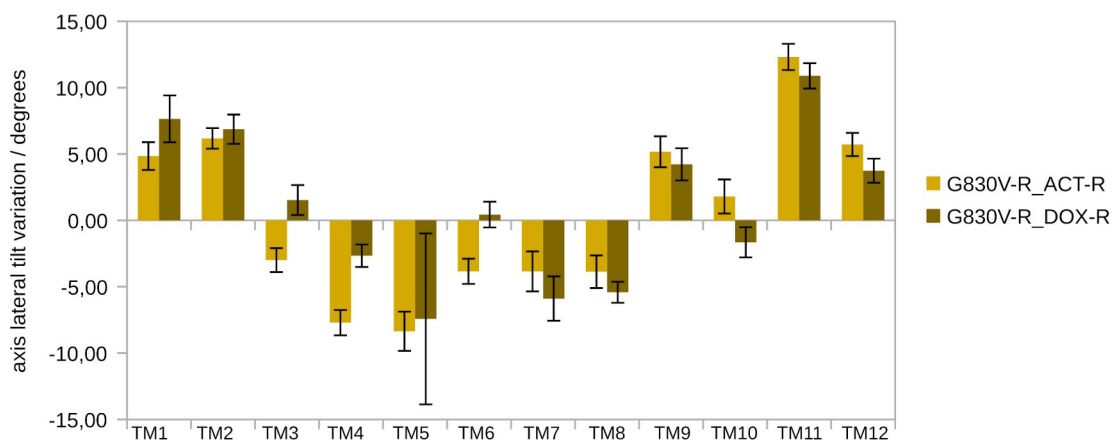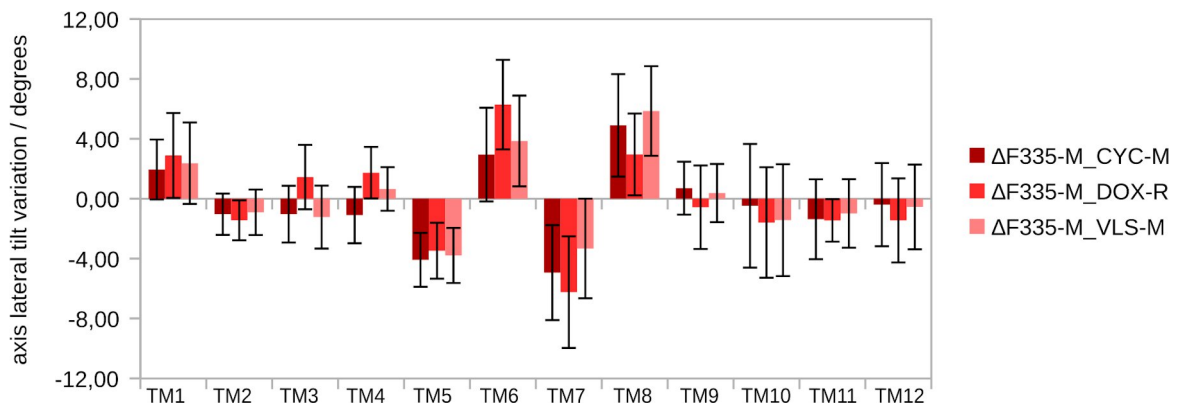

Figure S22: Variation in the transmembrane axis lateral tilt for the *holo* G830V variant (difference from the *holo* WT).

Table S12. Variation in the axis radial tilt observed in the *holo* G830V variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative mean changes).

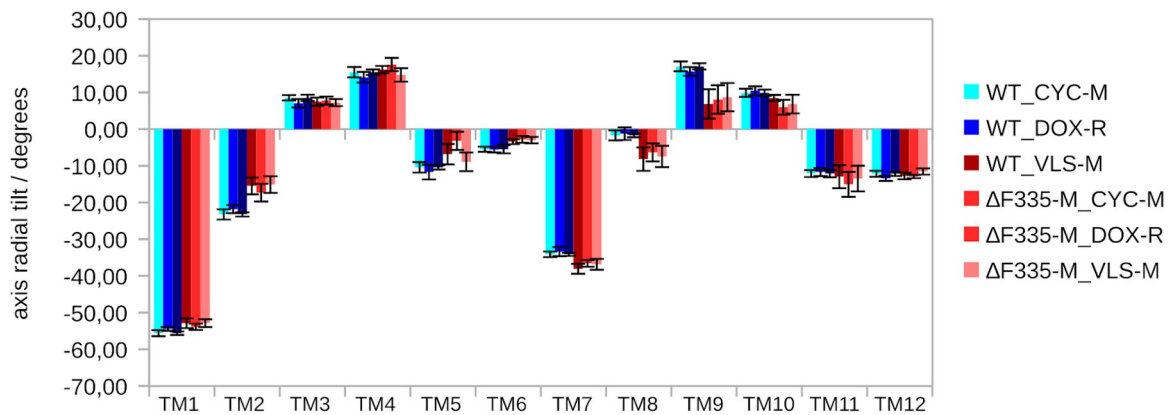|	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
G830V-R_ACT-R	+	+	+	+	ns*	+	–	–	+	ns	ns	+
G830V-R_DOX-R	+	+	+	+	+	+	–	–	+	–	–	+

* ns, not significant
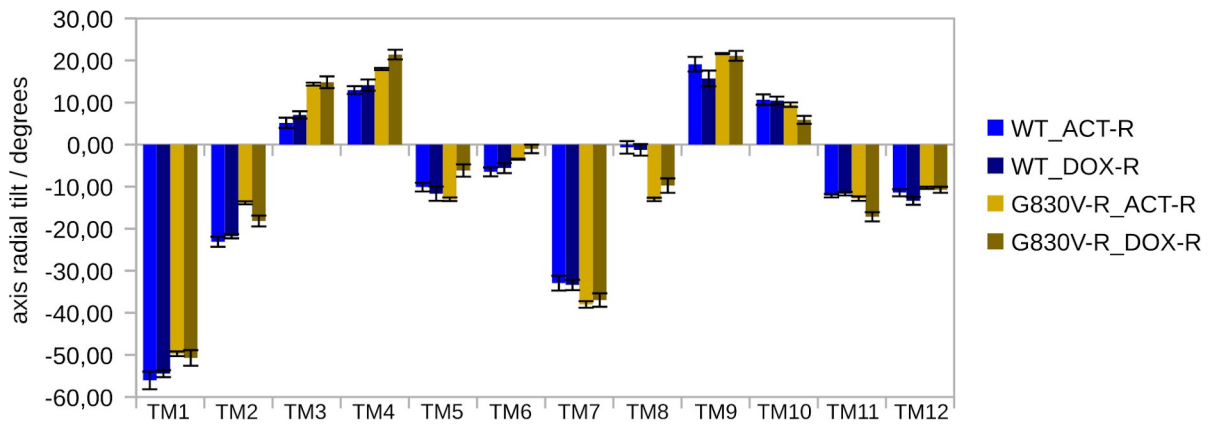
Figure S23: Comparison of individual transmembrane axis radial tilt for the WT and G830V varian (*holo* systems).
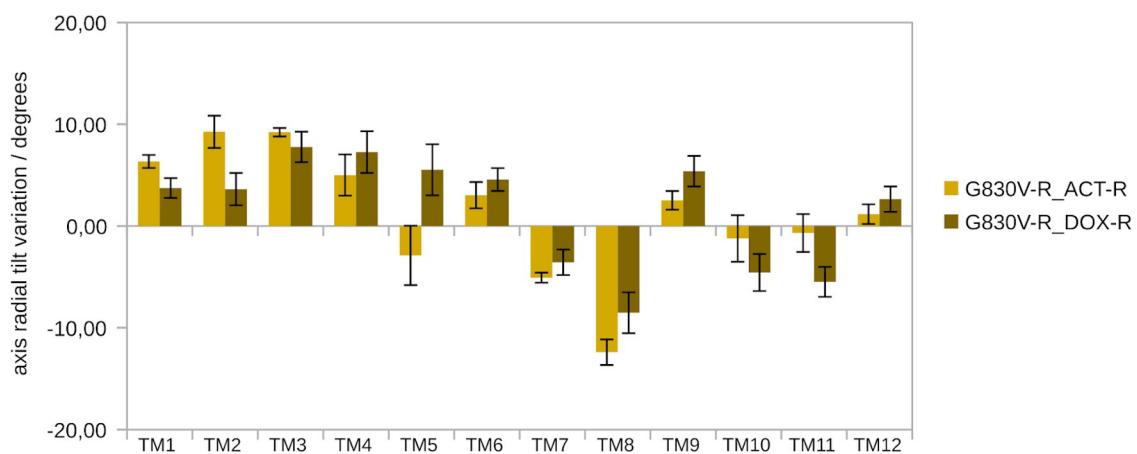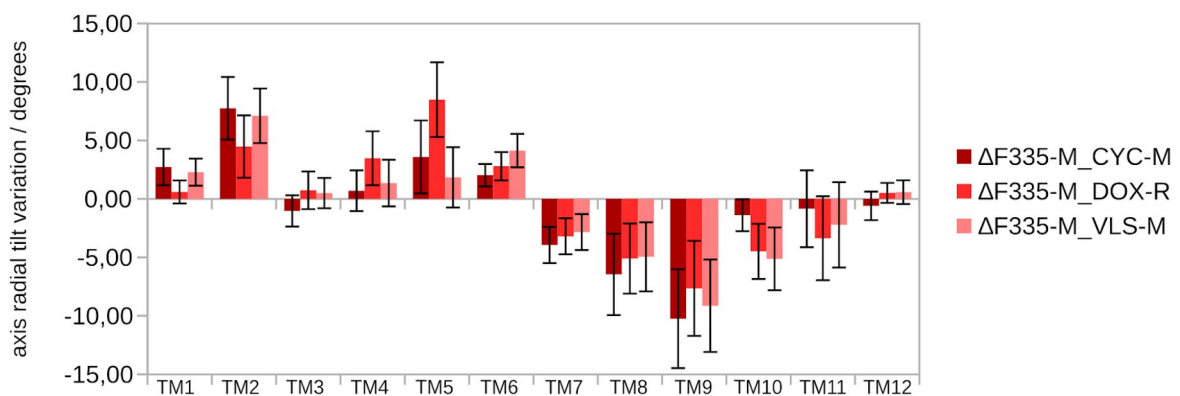

Figure S24: Variation in the transmembrane axis radial tilt for the *holo* G830V variant (difference from the *holo* WT).

Table S13. Variation in the axis distance observed in the *holo* F978A variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative mean changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
F978A-M_ACT-R	ns*	–	–	–	–	+	+	ns	–	+	+	+
F978A-M_CYC-M	+	–	ns	+	+	+	ns	–	–	+	+	+
F978A-M_VIN-H	ns	ns	–	–	ns	ns	+	ns	–	+	+	+

* ns, not significant
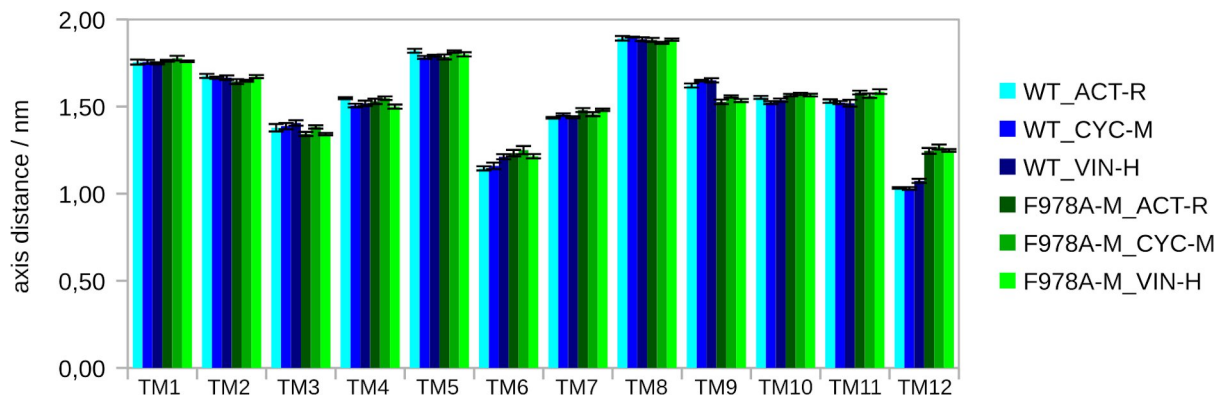
Figure S25: Comparison of individual transmembrane axis distance for the WT and F978A variant (*holo* systems).
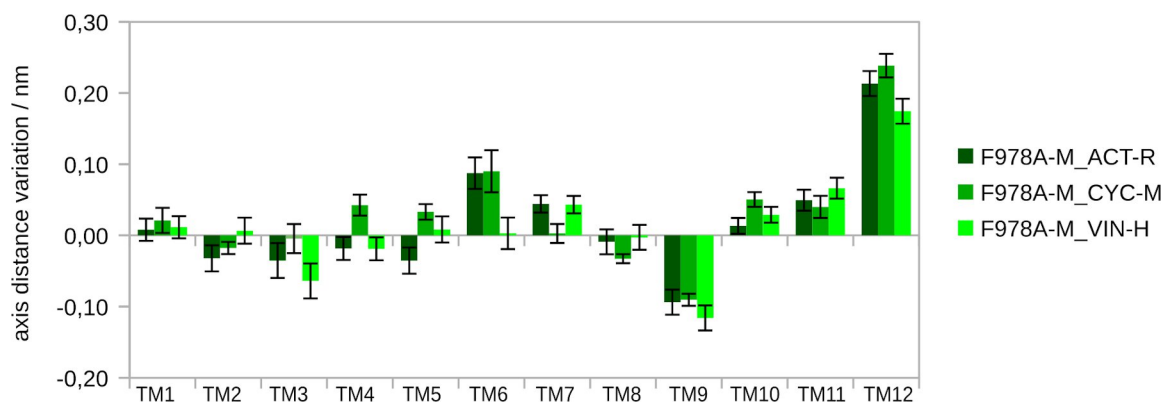
Figure S26: Variation in the transmembrane axis distance for the *holo* F978A variant (difference from the *holo* WT).

Table S14. Variation in the axis length observed in the *holo* F978A variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative mean changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
F978A-M_ACT-R	+	+	–	+	–	–	+	+	–	+	–	–
F978A-M_CYC-M	–	ns	–	+	ns	–	+	+	–	+	–	–
F978A-M_VIN-H	ns*	+	–	ns	–	–	+	ns	–	+	–	–

* ns, not significant
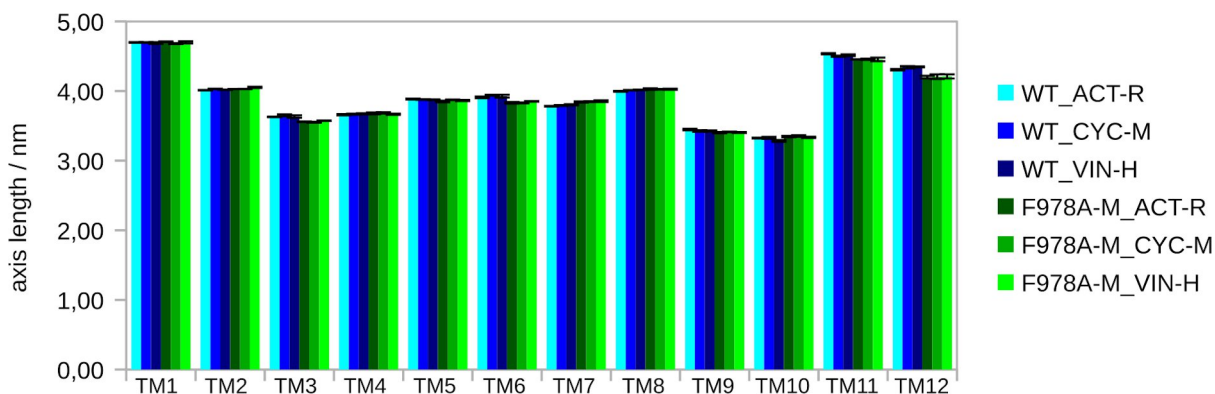
Figure S27: Comparison of individual transmembrane axis length for the WT and F978A variant (*holo* systems).
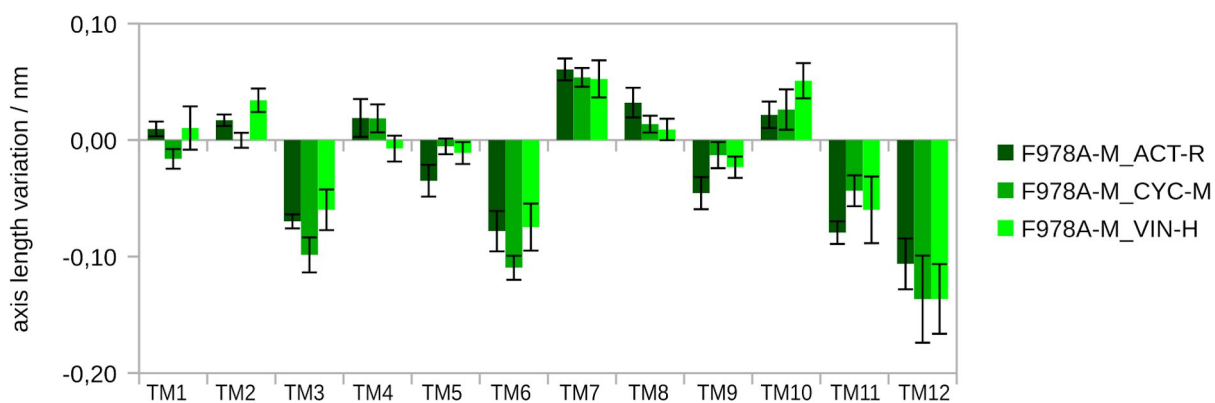
Figure S28: Variation in the transmembrane axis length for the *holo* F978A variant (difference from the *holo* WT).

Table S15. Variation in the *z*-shift of the axis mid-points observed in the *holo* F978A variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
F978A-M_ACT-R	–	–	–	–	–	–	+	ns	+	+	ns	+
F978A-M_CYC-M	–	ns	–	–	–	–	+	+	+	+	ns	+
F978A-M_VIN-H	ns*	ns	–	–	–	–	+	ns	ns	+	ns	+

* ns, not significant
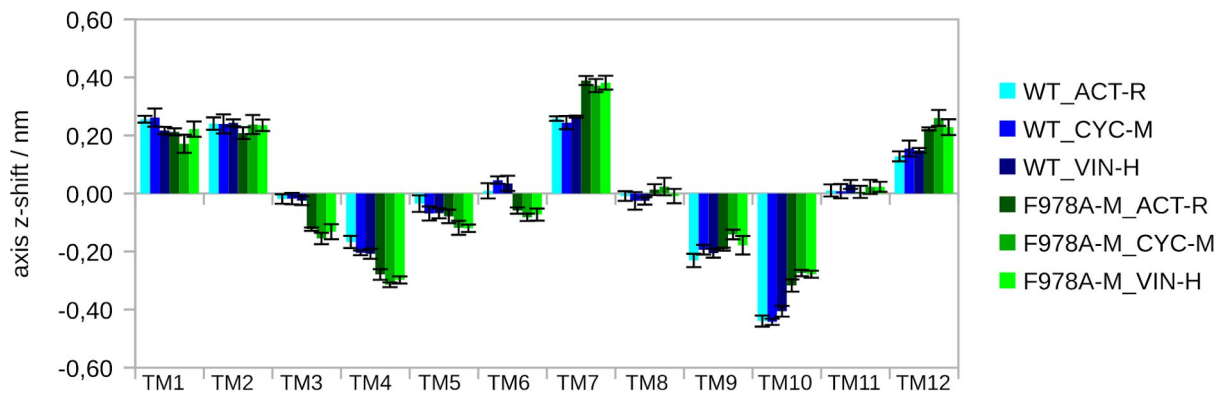
Figure S29: Comparison of individual transmembrane *z*-shift of the axis mid-points for the WT and F978A variant (*holo* systems).
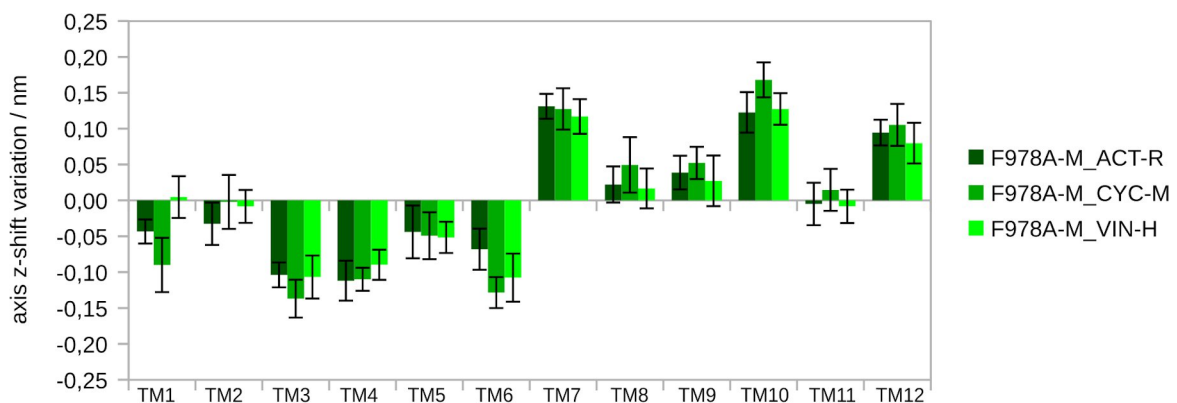
Figure S30: Variation in the transmembrane *z*-shift of the axis mid-points for the *holo* F978A variant (difference from the *holo* WT).

Table S16. Variation in the the axis total tilt observed in the *holo* F978A variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
F978A-M_ACT-R	ns*	–	+	+	+	ns	+	+	–	–	–	+
F978A-M_CYC-M	–	–	+	+	+	ns	+	+	–	–	–	+
F978A-M_VIN-H	+	–	+	+	+	ns	ns	+	–	ns	–	+

* ns, not significant
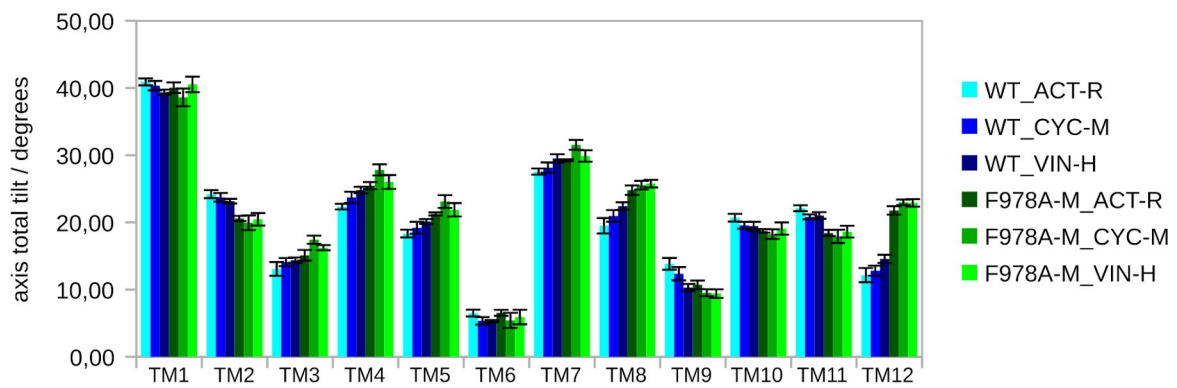
Figure S31: Comparison of individual transmembrane axis total tilt for the WT and F978A variant (*holo* systems).

Figure S32: Variation in the transmembrane axis total tilt for the *holo* F978A variant (difference from the *holo* WT).

Table S17. Variation in the taxis lateral tilt observed in the *holo* F978A variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative mean changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
F978A-M_ACT-R	+	+	ns*	–	–	+	–	–	ns	+	+	ns
F978A-M_CYC-M	+	+	ns	ns	–	+	–	–	ns	+	+	ns
F978A-M_VIN-H	+	+	+	ns	–	+	–	–	ns	ns	+	ns

* ns, not significant
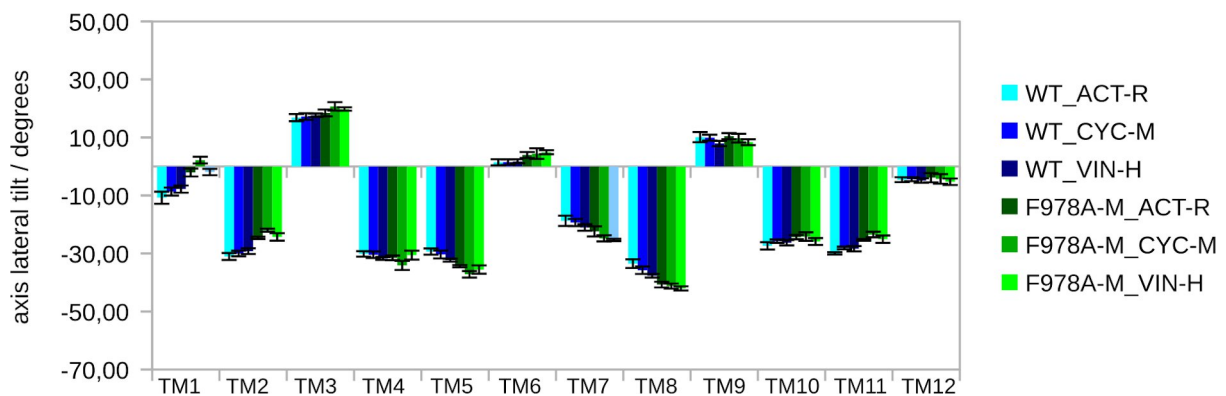
Figure S33: Comparison of individual transmembrane axis lateral tilt for the WT and F978A variant (*holo* systems).
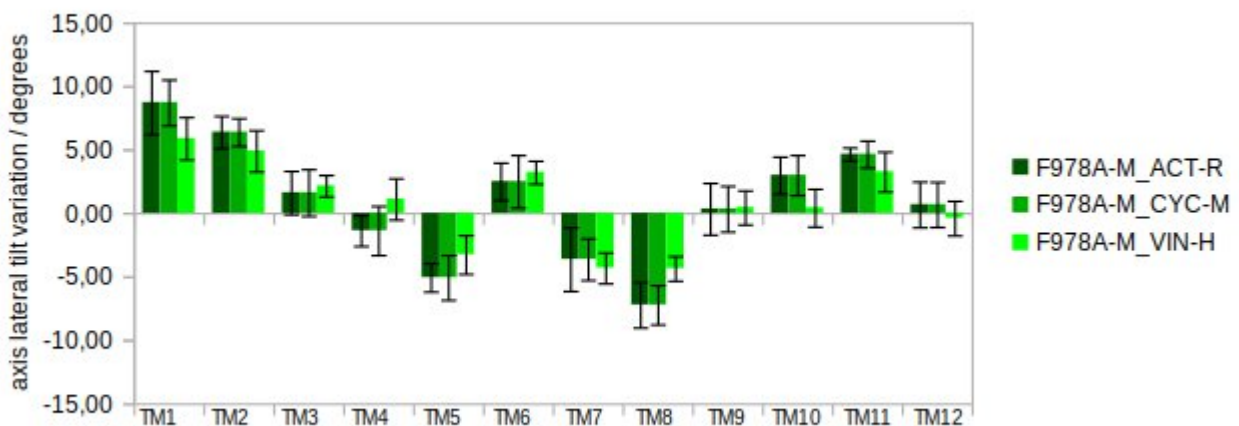
Figure S34: Variation in the transmembrane axis lateral tilt for the *holo* F978A variant (difference from the *holo* WT).

Table S18. Variation in the axis radial tilt observed in the *holo* F978A variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative mean changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
F978A-M_ACT-R	ns*	ns	+	+	+	ns	–	–	–	+	ns	–
F978A-M_CYC-M	ns	ns	+	+	+	+	–	–	–	+	–	–
F978A-M_VIN-H	ns	+	+	+	+	ns	ns	–	–	+	ns	–

* ns, not significant

Figure S35: Comparison of individual transmembrane axis radial tilt for the WT and F978A variant (*holo* systems).

Figure S36: Variation in the transmembrane axis radial tilt for the *holo* F978A variant (difference from the *holo* WT).

Table S19. Variation in the axis distance observed in the *holo* ΔF335 variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative mean changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
ΔF335-M_CYC-M	+	+	+	ns	+	+	+	ns	–	ns	+	–
ΔF335-M_DOX-R	+	+	ns*	–	ns	+	+	ns	–	+	+	–
ΔF335-M_VLS-M	+	+	–	–	+	ns	ns	–	–	+	+	–

* ns, not significant

Figure S37: Comparison of individual transmembrane axis distance for the WT and ΔF335 variant (*holo* systems).

Figure S38: Variation in the transmembrane axis distance for the *holo* ΔF335 variant (difference from the *holo* WT).

Table S20. Variation in the axis length observed in the *holo* ΔF335 variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative mean changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
ΔF335-M_CYC-M	ns*	–	+	+	–	+	ns	ns	–	+	–	–
ΔF335-M_DOX-R	–	ns	+	+	–	+	ns	ns	–	+	–	–
ΔF335-M_VLS-M	–	–	+	+	–	+	–	ns	ns	+	–	–

* ns, not significant

Figure S39: Comparison of individual transmembrane axis length for the WT and ΔF335 variant (*holo* systems).

Figure S40: Variation in the transmembrane axis length for the *holo* ΔF335 variant (difference from the *holo* WT).

Table S21. Variation in the *z*-shift of the axis mid-points observed in the *holo* ΔF335 variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative mean changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
ΔF335-M_CYC-M	–	–	ns	ns	+	–	+	ns	+	ns	ns	ns
ΔF335-M_DOX-R	ns*	–	+	–	ns	ns	+	–	+	ns	ns	ns
ΔF335-M_VLS-M	–	–	+	ns	ns	ns	+	ns	+	–	ns	ns

* ns, not significant

Figure S41: Comparison of individual transmembrane *z*-shift of the axis mid-points for the WT and ΔF335 variant (*holo* systems).
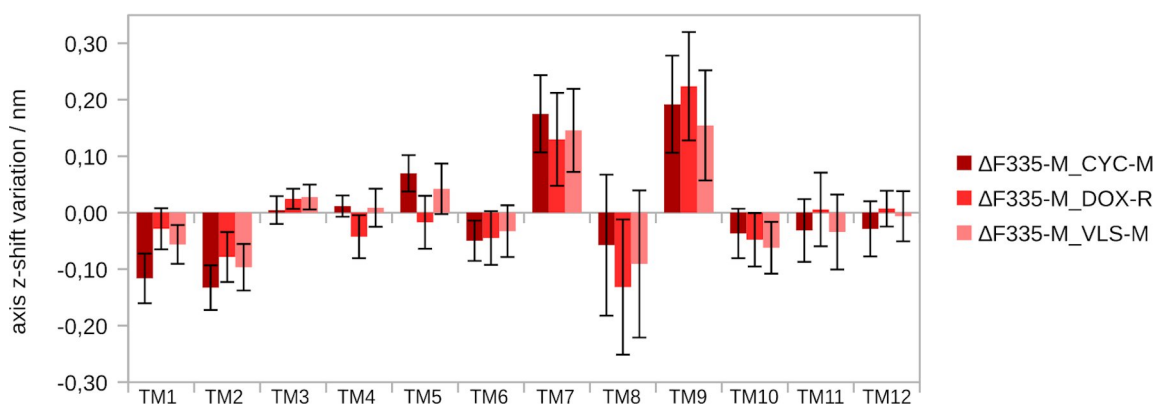
Figure S42: Variation in the transmembrane *z*-shift of the axis mid-points for the *holo* ΔF335 variant (difference from the *holo* WT).

Table S22. Variation in the the axis total tilt observed in the *holo* ΔF335 variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative mean changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
ΔF335-M_CYC-M	–	–	–	ns	+	ns	+	ns	–	ns	ns	+
ΔF335-M_DOX-R	–	–	ns*	ns	+	+	+	ns	–	ns	ns	ns
ΔF335-M_VLS-M	–	–	ns	ns	+	ns	+	–	–	ns	ns	ns

* ns, not significant

Figure S43: Comparison of individual transmembrane axis total tilt for the WT and ΔF335 variant (*holo* systems).

Figure S44: Variation in the transmembrane axis total tilt for the *holo* ΔF335 variant (difference from the *holo* WT).

Table S23. Variation in the taxis lateral tilt observed in the *holo* ΔF335 variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative mean changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
ΔF335-M_CYC-M	+	ns	ns	ns	–	+	–	+	ns	ns	ns	ns
ΔF335-M_DOX-R	+	ns	ns	+	–	+	–	+	ns	ns	ns	ns
ΔF335-M_VLS-M	ns*	ns	ns	ns	–	+	ns	+	ns	ns	ns	ns

* ns, not significant

Figure S45: Comparison of individual transmembrane axis lateral tilt for the WT and ΔF335 variant (*holo* systems).

Figure S46: Variation in the transmembrane axis lateral tilt for the *holo* ΔF335 variant (difference from the *holo* WT).

Table S24. Variation in the axis radial tilt observed in the *holo* ΔF335 variant transmembrane helices (compared to the *holo* WT; +, positive mean changes; –, negative mean changes).

	TM1	TM2	TM3	TM4	TM5	TM6	TM7	TM8	TM9	TM10	TM11	TM12
ΔF335-M_CYC-M	+	+	ns	ns	+	+	–	–	–	ns	ns	ns
ΔF335-M_DOX-R	ns*	+	ns	+	+	+	–	–	–	+	ns	ns
ΔF335-M_VLS-M	+	+	ns	ns	ns	+	–	–	–	+	ns	ns

* ns, not significant

Figure S47: Comparison of individual transmembrane axis radial tilt for the WT and ΔF335 variant (*holo* systems).

Figure S48: Variation in the transmembrane axis radial tilt for the WT and ΔF335 variant (difference from the *holo* WT).
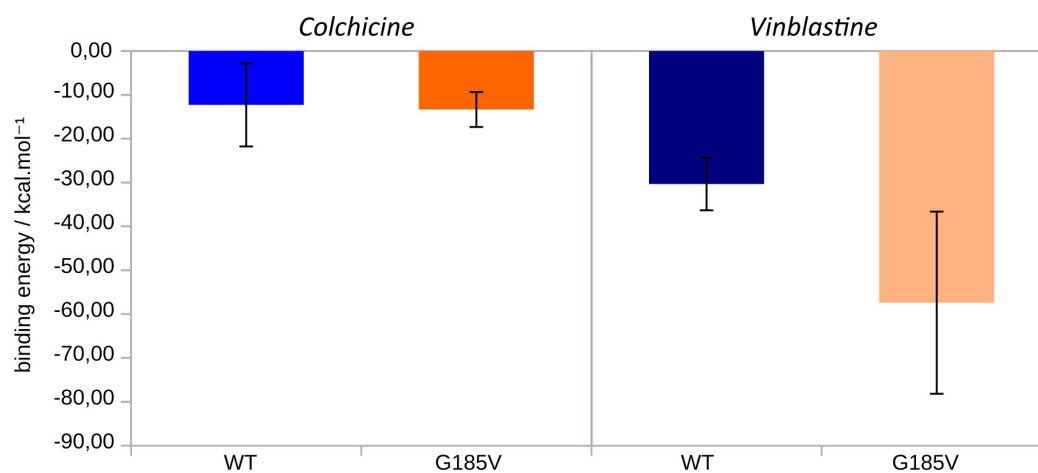
Figure S49: Relative Free-energies of binding (ΔG_{bind}) for colchicine and vinblastine in the WT and G185V variant.
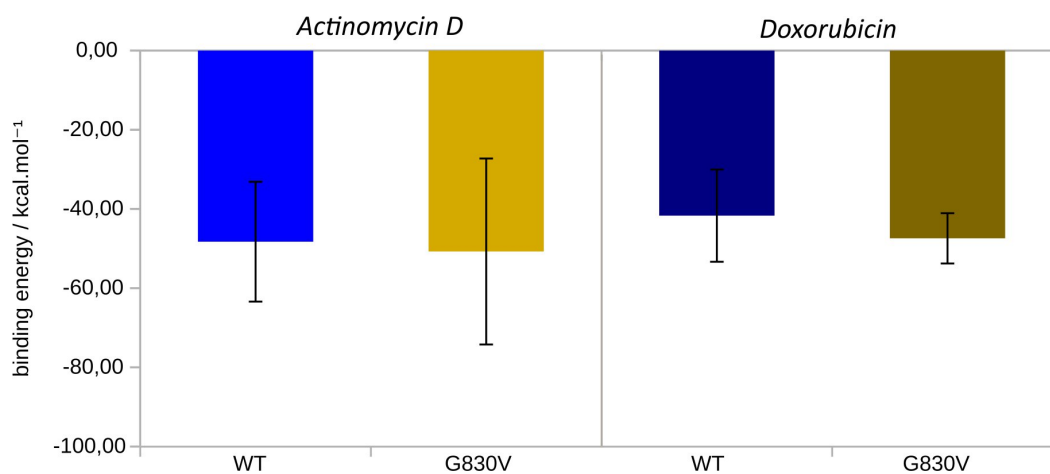
Figure S50: Relative Free energies of binding (ΔG_{bind}) for actinomycin D and doxorubicin in the WT and G830V variant.
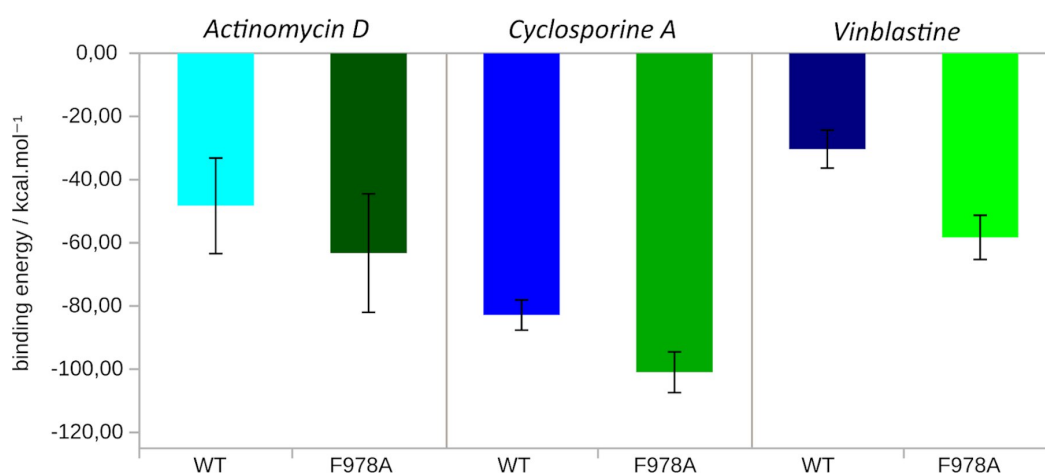
Figure S51: Relative Free energies of binding (ΔG_{bind}) for actinomycin D, cyclosporine A and vinblastine in the WT and F978A variant.
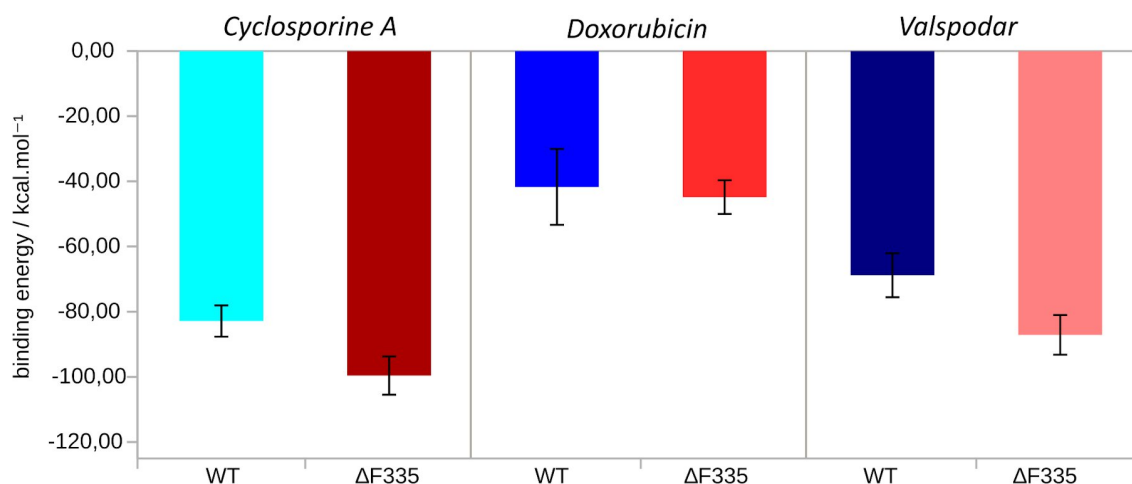
Figure S52: Relative Free energies of binding (ΔG_{bind}) for cyclosporine A, doxorubicin and valspodar in the WT and ΔF335 variant.

Table S25. Variation in the contact frequencies and hydrogen bond (HB) parameters for colchicine (G185V) and vinblastine (G185V and F978A), both located at the H-site.

Residues		WT_COL	G185V_COL		WT_VIN	G185V_VIN		F978A_VIN	
		Frequency	Frequency	Variation %	Frequency	Frequency	Variation %	Frequency	Variation %
189	LYS	0,106931	0,683168	538,89	0,471287	0,332673	-29,41	0,871287	84,87
192	MET					0,609901	NEW	0,263938	NEW
195	GLN					0,592079	NEW		
196	SER					0,655446	NEW		
344	SER					0,710891	NEW		
347	GLN							0,648342	NEW
832	ARG				0,522772		LOSS	0,131747	-74,80
938	PHE							0,824639	NEW
941	THR							0,726674	NEW
942	PHE	0,641584	0,269307	-58,02		0,847525	NEW	0,947510	NEW
945	THR	0,758416		LOSS					
946	GLN		0,825742	NEW		0,570297	NEW	0,730289	NEW

		WT_COL	G185V_COL	Variation %	WT_VIN	G185V_VIN	Variation %	F978A_VIN	Variation %
HB	life	1473,240	1063,813	-27,79	383,12	1353,856	253,38	1177,059	207,23
	<N>	1,706	2,729	59,93	1,32	0,604	-54,23	1,422	7,71
	ΔG	-22,713	-21,872	-3,70	-19,55	-21,846	11,72	-22,369	14,39

The variation percentage is estimated based on the mean frequencies of five replica systems for the *holo* P-gp systems. Green-shaded cells, no contact frequencies calculated for the specified residue.

Table S26. Variation in the contact frequencies and hydrogen bond (HB) parameters for actinomycin D (G830V and F978A) and doxorubicin (G830V and ΔF335), both located at the R-site.

Residues		WT_ACT	G830V_ACT		F978A_ACT		WT_DOX	G830V_DOX		ΔF335_DOX	
		Frequency	Frequency	Variation %	Frequency	Variation %	Frequency	Frequency	Variation %	Frequency	Variation %
232	TRP	0,758416	0,829703	9,40	0,897030	18,28	0,500990	0,823779	64,43	0,650401	29,82
291	LYS	0,477228		LOSS	0,304951	-36,10					
296	ASN	0,164357			0,912871	455,42	0,875247	0,600358	-31,41	1,000000	14,25
347	GLN	0,215842	0,576237	166,97							
770	PHE	0,306931	0,643564	109,68	0,514851	67,74					
773	GLN	0,209901	0,463366	120,75	0,619802	195,28	0,310891	0,905019	191,10	0,659088	112,00
777	PHE						0,750495	0,703243	-6,30	0,185971	-75,22
778	GLY						0,403960	0,561318	38,95		LOSS
831	SER	0,431683	0,960396	122,48	0,568317	31,65					
838	GLN	0,253465	0,895049	253,12	0,400000	57,81	0,546534	0,548412	0,34	0,715486	30,91

		WT_ACT	G830V_ACT	Variation %	F978A_ACT	Variation %	WT_DOX	G830V_DOX	Variation %	ΔF335_DOX	Variation %
HB	life	1062,751	525,676	-50,54	406,804	-61,72	397,677	561,385	41,17	1177,059	195,98
	<N>	2,638	2,443	-7,36	2,897	9,83	0,896	0,774	-13,64	1,422	58,68
	ΔG	-21,956	-19,878	-9,47	-19,694	-10,30	-19,649	-20,086	2,22	-22,369	13,85

The variation percentage is estimated based on the mean frequencies of five replica systems for the *holo* P-gp systems. Green-shaded cells, no contact frequencies calculated for the specified residue.

Table S27. Variation in the contact frequencies and hydrogen bond (HB) parameters for cyclosporine A (F978A and ΔF335) and valspodar (ΔF335), both located at the M-site.

Residues		WT_CYC Frequency	F978A_CYC Frequency	Variation %	ΔF335_CYC Frequency	Variation %	WT_VLS Frequency	ΔF335_VLS Frequency	Variation %
303	PHE	0,522772	0,615408	17,72	0,691089	32,20	0,841584	0,388119	-53,88
307	TYR	0,813861	0,765076	-5,99	0,514852	-36,74			
310	TYR	0,883168	0,662578	-24,98	0,368317	-58,30			
343	PHE	0,370297	0,678006	83,10	0,667327	80,21	0,419802	0,803960	91,51
347	GLN	0,819802	0,566839	-30,86	0,445544	-45,65	0,809901	0,536634	-33,74
721	ASN		0,604827	NEW					
725	GLN	0,164356	0,792936	382,45	0,663366	303,61	0,720792	0,283168	-60,71
759	PHE						0,605940		LOSS
766	SER		0,681897	NEW					
770	PHE	0,295050	0,460815	56,18					
946	GLN	0,174258			0,516832	196,59			
983	PHE						0,576238	0,368317	-36,08
990	GLN						0,097030	0,641584	561,22

		WT_CYC	F978A_CYC	Variation %	ΔF335_CYC	Variation %	WT_VLS	ΔF335_VLS	Variation %
HB	life	624,098	602,512	-3,46	644,905	3,33	887,101	770,965	-13,09
	<N>	2,879	2,822	-1,99	2,255	-21,69	1,500	1,774	18,29
	ΔG	-20,614	-20,587	-0,13	-20,811	0,95	-21,478	-21,510	0,15

The variation percentage is estimated based on the mean frequencies of five replica systems for the holo P-gp systems. Green-shaded cells, no contact frequencies calculated for the specified residue.

Table S28: Protein-ligand contact efficiency ratio for the WT and P-gp variants.

	WT			G185V_H		
Ligand	mean Total	Mean (>0,5)	Ratio	mean Total	Mean (>0,5)	Ratio
COL_H	8,80	3,80	0,43	9,80	4,20	0,43
VIN_H	13,20	4,00	0,30	8,00	5,60	0,70

	WT			G830V_R		
Ligand	mean Total	Mean (>0,5)	Ratio	mean Total	Mean (>0,5)	Ratio
ACT_R	19,60	4,00	0,20	18,20	6,20	0,34
DOX_R	19,80	5,80	0,29	20,60	6,80	0,33

	WT			F978A_M		
Ligand	mean Total	Mean (>0,5)	Ratio	mean Total	Mean (>0,5)	Ratio
ACT_R	19,60	4,00	0,20	19,80	5,20	0,26
CYC_M	27,60	3,80	0,14	26,60	7,60	0,29
VIN_H	13,20	4,00	0,30	16,60	7,00	0,42

	WT			ΔF335_M		
Ligand	mean Total	Mean (>0,5)	Ratio	mean Total	Mean (>0,5)	Ratio
CYC_M	27,60	3,80	0,14	26,60	6,60	0,25
DOX_R	19,80	5,80	0,29	19,20	5,20	0,27
VLS_M	22,60	5,40	0,24	25,20	6,40	0,25

The variation percentage is estimated based on the mean frequencies of five replica systems for the *holo* P-gp systems.
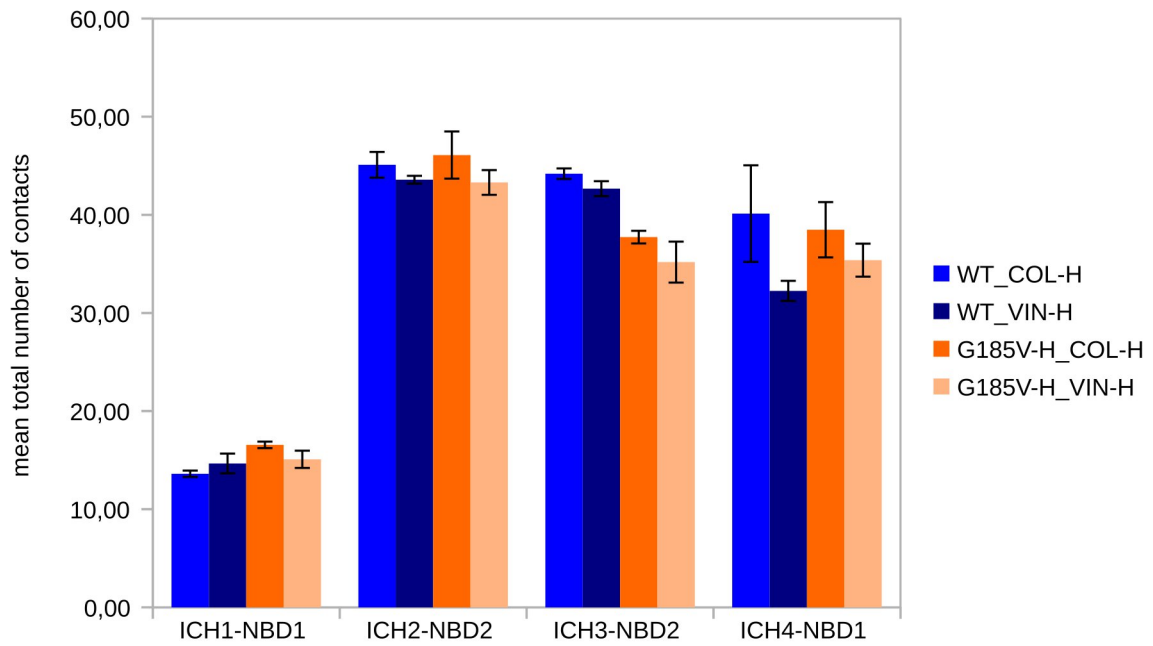
Figure S53: Comparison of the total number of contacts at each ICH-NBD interface between the WT and G185V variant (*holo* systems).
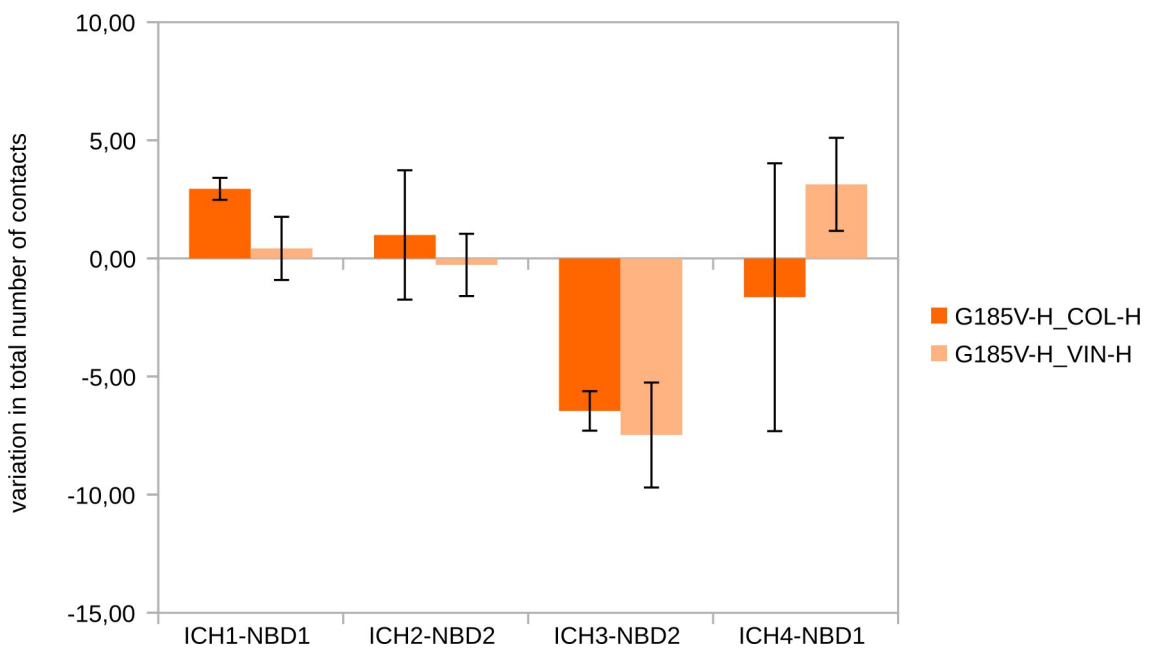
Figure S54: Variation in the total number of contacts at each ICH-NBD interface for the *holo* G185V variant (difference from the *holo* WT).
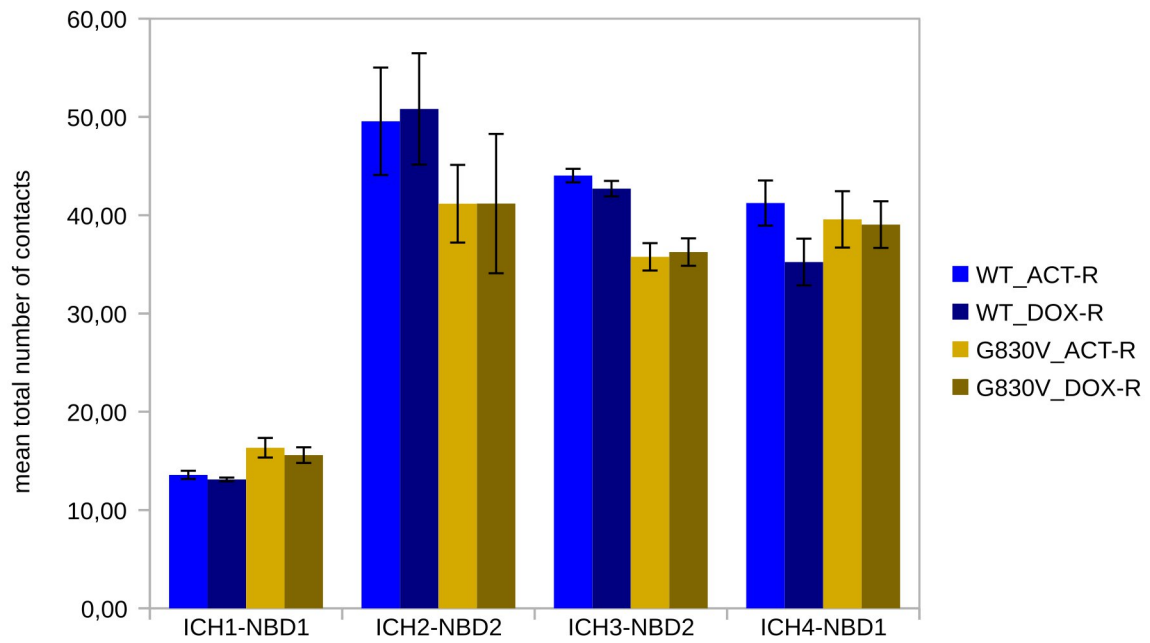
Figure S55: Comparison of the total number of contacts at each ICH-NBD interface between the WT and G830V variant (*holo* systems).
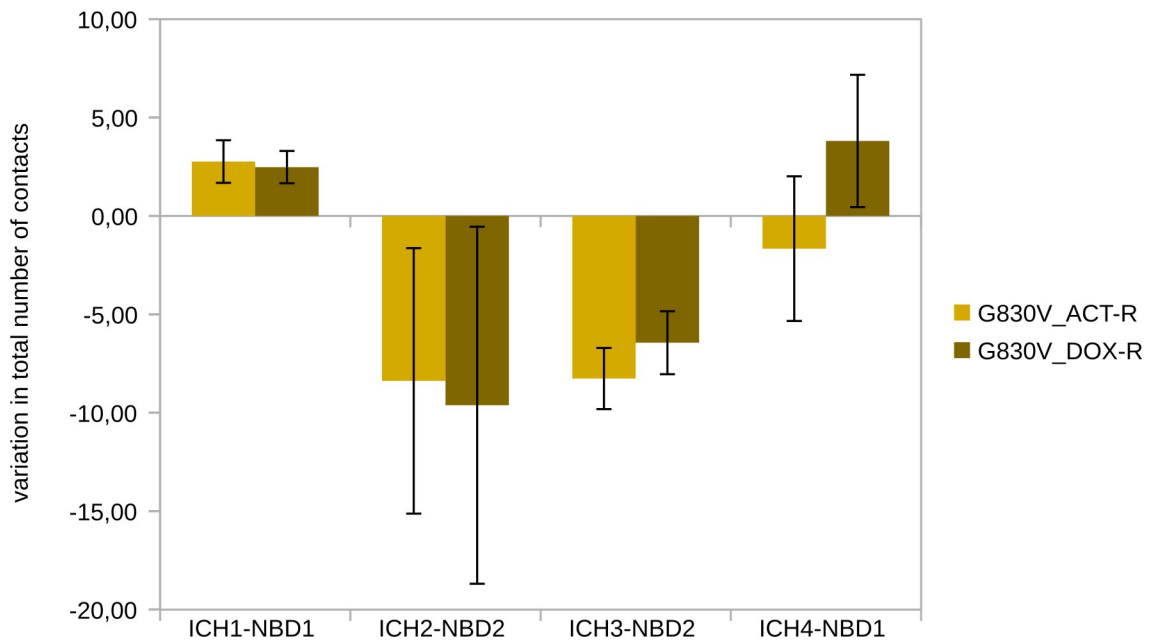
Figure S56: Variation in the total number of contacts at each ICH-NBD interface for the *holo* G830V variant (difference from the *holo* WT).
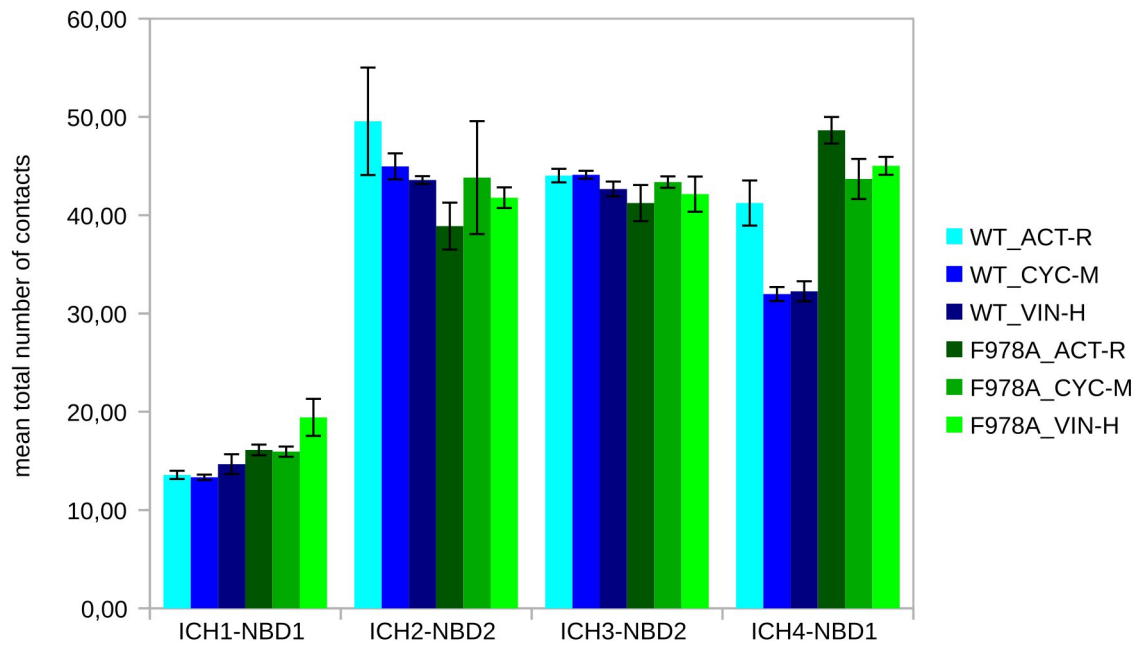
Figure S57: Comparison of the total number of contacts at each ICH-NBD interface between the WT and F978A variant (*holo* systems).
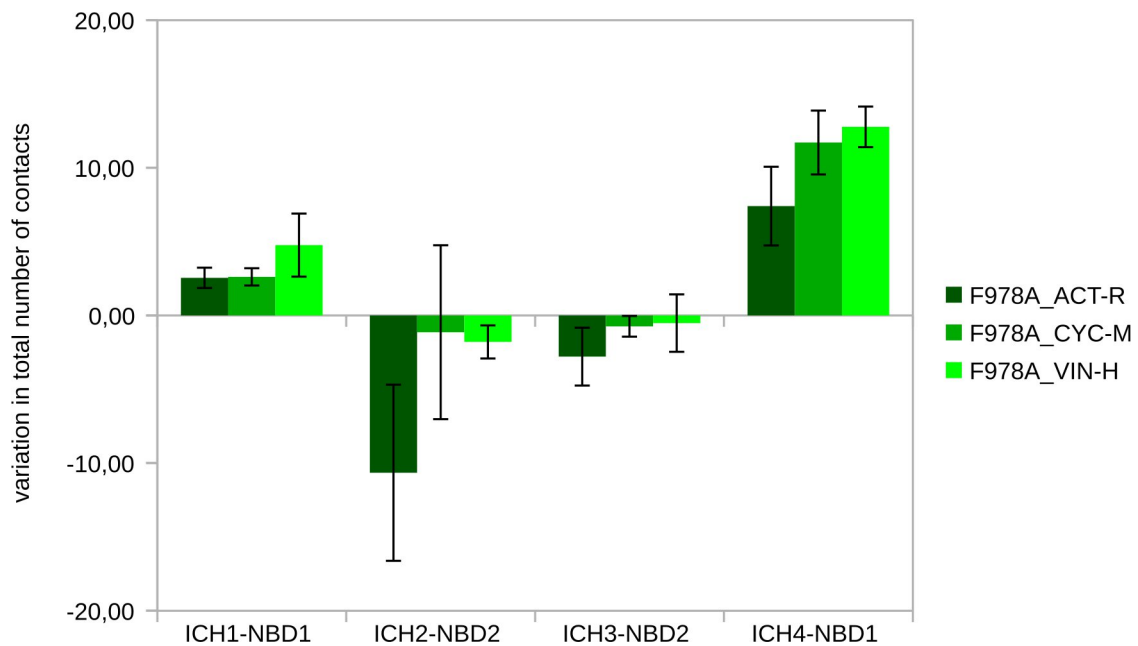
Figure S58: Variation in the total number of contacts at each ICH-NBD interface for the *holo* F978A variant (difference from the *holo* WT).
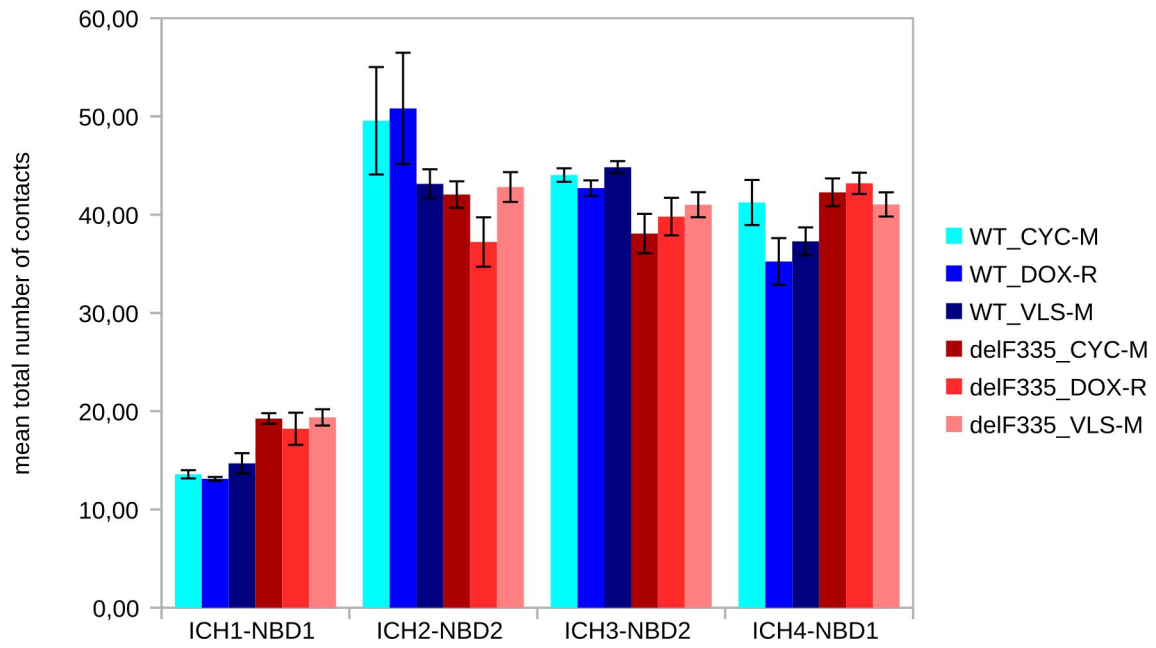
Figure S59: Comparison of the total number of contacts at each ICH-NBD interface between the WT and ΔF335 variant (*holo* systems).
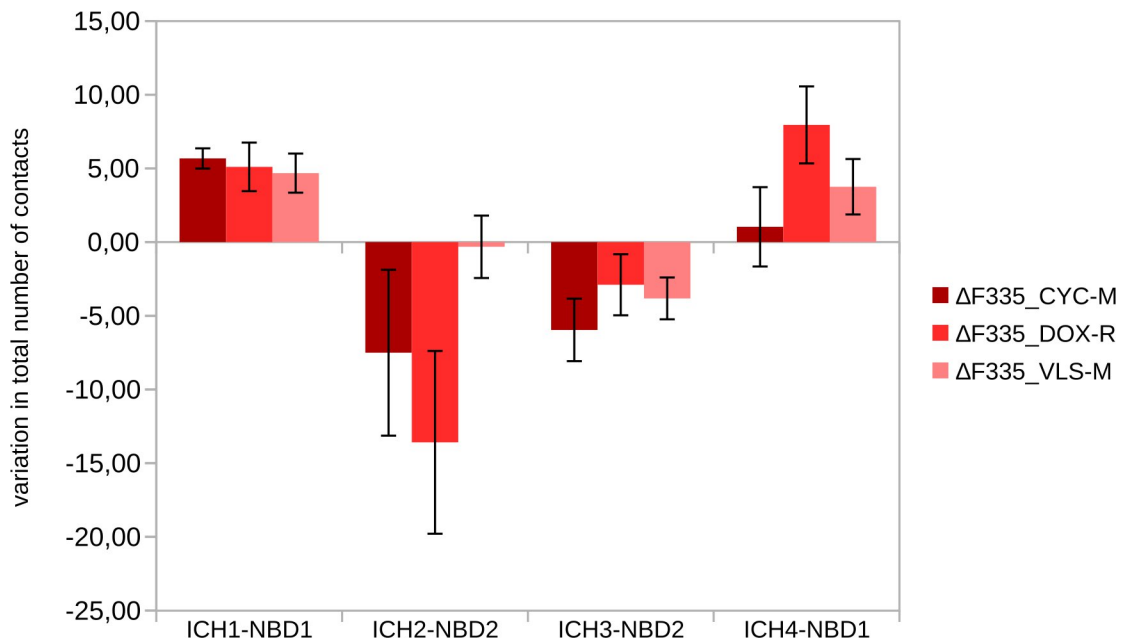
Figure S60: Variation in the total number of contacts at each ICH-NBD interface for the *holo* ΔF335 variant (difference from the *holo* WT).

Table S29. Variation in the contact frequencies and hydrogen bond (HB) parameters between the ICH-NBD residues for the WT and G185V variants (*holo* systems).

TMD-NBD interface				WT_COL	G185V_COL		WT_VIN	G185V_VIN	
ICH1		NBD1		Frequency	Frequency	Variation %	Frequency	Frequency	Variation %
160	ILE	443	LEU	0,188119	0,748515	297,89	0,356436	0,502970	41,11
164	ASP	404	ARG	0,946535	0,982178	3,77	0,970297	0,859406	-11,43
164	ASP	405	LYS	0,629505	0,984158	56,34	0,960396	0,786139	-18,14

Hydrogen Bonds			WT_COL	G185V_COL	Variation %	WT_VIN	G185V_VIN	Variation %
ICH1-NBD1		life	570,612	709,366	24,32	426,906	369,544	-13,44
		<N>	1,911	2,442	27,77	2,188	1,976	-9,69
		DG	-20,376	-20,730	1,74	-19,734	-19,247	-2,47

TMD-NBD interface				WT_COL	G185V_COL		WT_VIN	G185V_VIN	
ICH2		NBD2		Frequency	Frequency	Variation %	Frequency	Frequency	Variation %
262	ARG	1086	PHE	0,425743	0,576238	35,35	0,502970	0,568317	12,99
263	THR	1117	SER	0,635643	0,603168	-5,11	0,649505	0,611881	-5,79
263	THR	1118	GLN	0,035644	0,423762	1088,89	0,031683	0,512871	1518,75
263	THR	1200	ASP	0,198020	0,578218	192,00	0,031683	0,673267	2025,00
265	ILE	1110	ARG	0,758416	0,657426	-13,32	0,639604	0,623762	-2,48
266	ALA	1086	PHE	0,596039	0,483169	-18,94	0,596039	0,522772	-12,29
267	PHE	1110	ARG	0,504911	0,332674	-34,11	0,471287	0,526733	11,76
267	PHE	1115	ILE	0,936634	0,908911	-2,96	0,897030	0,920792	2,65

Hydrogen Bonds			WT_COL	G185V_COL	Variation %	WT_VIN	G185V_VIN	Variation %
ICH2-NBD2		life	452,519	393,494	-13,04	335,881	940,684	180,06
		<N>	1,396	1,861	33,30	1,095	1,600	46,13
		DG	-19,839	-19,057	-3,94	-18,981	-20,133	6,07

TMD-NBD interface				WT_COL	G185V_COL		WT_VIN	G185V_VIN	
ICH3		NBD2		Frequency	Frequency	Variation %	Frequency	Frequency	Variation %
800	ASP	1086	PHE	0,736633	0,826733	12,23	0,699010	0,782178	11,90
800	ASP	1087	TYR	0,994059	0,995050	0,10	0,990099	0,995050	0,50
801	VAL	1086	PHE	0,902970	0,883663	-2,14	0,914851	0,873762	-4,49
801	VAL	1087	TYR	0,469307	0,512376	9,18	0,485149	0,502475	3,57
802	SER	1087	TYR	0,687128	0,646039	-5,98	0,693069	0,695544	0,36
805	ASP	1044	TYR	0,912871	0,908416	-0,49	0,922772	0,888614	-3,70
805	ASP	1046	THR	1,000000	0,878713	-12,13	1,000000	0,626238	-37,38

Hydrogen Bonds			WT_COL	G185V_COL	Variation %	WT_VIN	G185V_VIN	Variation %
ICH3-NBD2		life	981,591	690,718	-29,63	702,618	339,698	-51,65
		<N>	2,814	1,790	-36,37	2,871	1,956	0,96
		DG	-20,899	-20,815	-0,40	-20,761	-19,098	-20,10

TMD-NBD interface				WT_COL	G185V_COL		WT_VIN	G185V_VIN	
ICH4		NBD1		Frequency	Frequency	Variation %	Frequency	Frequency	Variation %
905	ARG	441	GLN	1,000000	1,000000	0,00	1,000000	1,000000	0,00
908	VAL	467	ARG	0,736633	0,697029	-5,38	0,609901	0,631683	3,57
909	SER	441	GLN	0,998020	0,839604	-15,87	1,000000	0,887129	-11,29
911	THR	467	ARG	0,724752	0,000000	Lost	0,459406	0,057426	-87,50
912	GLN	464	ARG	0,623762	0,318812	-48,89	0,435644	0,366337	-15,91
912	GLN	467	ARG	0,150495	0,572277	280,26	0,180198	0,394059	118,68

Hydrogen Bonds			WT_COL	G185V_COL	Variation %	WT_VIN	G185V_VIN	Variation %
ICH4-NBD1		life	475,834	365,229	-23,24	454,402	284,478	-37,40
		<N>	3,063	2,673	-12,73	2,556	1,991	-22,14
		DG	-19,742	-18,657	-5,50	-19,648	-18,789	-4,37

The variation percentage is estimated based on the mean frequencies of five replica systems for the WT and G185V variant (*holo* systems).

Table S30. Variation in the contact frequencies and hydrogen bond (HB) parameters between the ICH-NBD residues for the WT and G830V variants (*holo* systems).

TMD-NBD interface				WT_ACT	G830V_ACT		WT_DOX	G830V_DOX	
ICH1		NBD1		Frequency	Frequency	Variation %	Frequency	Frequency	Variation %
164	ASP	404	ARG	0,984139	0,986149	0,20	0,986119	0,992060	0,60
164	ASP	405	LYS	0,991074	0,841535	-15,09	0,940535	0,828688	-11,89

Hydrogen Bonds			WT_ACT	G830V_ACT	Variation %	WT_DOX	G830V_DOX	Variation %
ICH1-NBD1		life	779,200	677,000	-13,12	724,200	756,000	4,39
		<N>	2,035	2,065	1,50	2,020	2,009	-0,55
		DG	-21,072	-20,895	-0,84	-21,022	-21,339	1,50

TMD-NBD interface				WT_ACT	G830V_ACT		WT_DOX	G830V_DOX	
ICH2		NBD2		Frequency	Frequency	Variation %	Frequency	Frequency	Variation %
262	ARG	1077	SER	0,463366	0,269307	-41,88	0,000000	0,000000	0,00
262	ARG	1081	GLN	0,475248	0,401980	-15,42	0,449505	0,334500	-25,58
262	ARG	1086	PHE	0,457426	0,520792	13,85	0,481188	0,448378	-6,82
263	THR	1117	SER	0,540594	0,316832	-41,39	0,744554	0,374684	-49,68
265	ILE	1110	ARG	0,695049	0,740594	6,55	0,716831	0,860754	20,08
266	ALA	1086	PHE	0,623762	0,459386	-26,35	0,532673	0,462209	-13,23
266	ALA	1110	ARG	0,421783	0,548515	30,05	0,459406	0,415108	-9,64
267	PHE	1110	ARG	0,461386	0,429703	-6,87	0,477228	0,327416	-31,39
267	PHE	1115	ILE	0,916832	0,827723	-9,72	0,934653	0,884585	-5,36

Hydrogen Bonds			WT_ACT	G830V_ACT	Variation %	WT_DOX	G830V_DOX	Variation %
ICH2-NBD2		life	301,800	1782,600	490,66	501,600	496,800	-0,96
		<N>	1,142	1,104	-3,38	1,309	0,980	-25,19
		DG	-18,851	-22,547	19,61	-20,198	-19,927	-1,34

TMD-NBD interface				WT_ACT	G830V_ACT		WT_DOX	G830V_DOX	
ICH3		NBD2		Frequency	Frequency	Variation %	Frequency	Frequency	Variation %
800	ASP	1086	PHE	0,754455	0,780198	3,41	0,718812	0,794452	10,52
800	ASP	1087	TYR	1,000000	0,998020	-0,20	0,968317	0,994059	2,66
801	VAL	1086	PHE	0,928713	0,853465	-8,10	0,891089	0,885029	-0,68
801	VAL	1087	TYR	0,526733	0,091089	-82,71	0,510891	0,032127	-93,71
802	SER	1087	TYR	0,736633	0,299010	-59,41	0,655445	0,279891	-57,30
805	ASP	1044	TYR	0,904950	0,469307	-48,14	0,930693	0,504916	-45,75
805	ASP	1046	THR	0,998020	0,998020	0,00	0,998020	0,803960	-19,44

Hydrogen Bonds			WT_ACT	G830V_ACT	Variation %	WT_DOX	G830V_DOX	Variation %
ICH3-NBD2		life	858,200	773,800	-9,83	565,400	623,800	10,33
		<N>	2,849	2,979	4,56	2,686	2,997	11,59
		DG	-21,282	-21,316	0,16	-20,742	-20,789	0,22

TMD-NBD interface				WT_ACT	G830V_ACT		WT_DOX	G830V_DOX	
ICH4		NBD1		Frequency	Frequency	Variation %	Frequency	Frequency	Variation %
905	ARG	438	GLN	0,451485	0,146535	-67,54	0,000000	0,000000	0,00
905	ARG	441	GLN	1,000000	0,998020	-0,20	0,998020	0,996040	-0,20
908	VAL	467	ARG	0,566337	0,720811	27,28	0,560396	0,708928	26,50
909	SER	441	GLN	0,992079	0,912871	-7,98	0,980198	0,868163	-11,43
910	LEU	467	ARG	0,000000	0,936634	NEW	0,000000	0,828883	NEW
911	THR	467	ARG	0,617830	0,376238	-39,10	0,000000	0,000000	0,00
912	GLN	464	ARG	0,643564	0,467327	-27,38	0,000000	0,000000	0,00

Hydrogen Bonds			WT_ACT	G830V_ACT	Variation %	WT_DOX	G830V_DOX	Variation %
ICH4-NBD1		life	621,200	441,600	-28,91	365,200	705,400	93,15
		<N>	2,828	2,495	-11,78	2,544	2,368	-6,90
		DG	-20,458	-19,820	-3,12	-19,324	-20,750	7,38

The variation percentage is estimated based on the mean frequencies of five replica systems for the WT and G830V variant (*holo* systems).

Table S31. Variation in the contact frequencies and hydrogen bond (HB) parameters between the ICH-NBD residues for the WT and F978A variants (*holo* systems).

TMD-NBD interface				WT_ACT	F978A_ACT		WT_CYC	F978A_CYC		WT_VIN	F978A_VIN	
ICH1		NBD1		Frequency	Frequency	Variation %	Frequency	Frequency	Variation %	Frequency	Frequency	Variation %
160	ILE	443	LEU	0,000000	0,000000	0,00	0,140594	0,485149	245,07	0,356436	0,661386	85,56
164	ASP	404	ARG	0,984139	0,936614	-4,83	0,974257	0,964356	-1,02	0,970297	0,980198	1,02
164	ASP	405	LYS	0,991074	0,776238	-21,68	0,990099	0,871287	-12,00	0,960396	0,778218	-18,97

Hydrogen Bonds			WT_ACT	F978A_ACT	Variation %	WT_CYC	F978A_CYC	Variation %	WT_VIN	F978A_VIN	Variation %
ICH1-NBD1		life	779,200	442,800	-43,17	484,600	520,400	7,39	426,9058	392,800	-7,99
		<N>	2,035	2,038	0,15	1,929	2,131	10,46	2,1882	2,442	11,59
		DG	-21,072	-19,931	-5,41	-20,150	-20,238	0,44	-19,7344	-19,572	-0,82

TMD-NBD interface				WT_ACT	F978A_ACT		WT_CYC	F978A_CYC		WT_VIN	F978A_VIN	
ICH2		NBD2		Frequency	Frequency	Variation %	Frequency	Frequency	Variation %	Frequency	Frequency	Variation %
262	ARG	1044	TYR	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,5148514	0,403961	-21,54
262	ARG	1077	SER	0,463366	0,435643	-5,98	0,447525	0,560396	25,22	0,3841586	0,463367	20,62
262	ARG	1081	GLN	0,475248	0,495050	4,17	0,407921	0,594099	45,64	0,3960398	0,550495	39,00
262	ARG	1086	PHE	0,457426	0,465347	1,73	0,514851	0,439604	-14,62	0,5029704	0,550495	9,45
263	THR	1117	SER	0,540594	0,562376	4,03	0,740594	0,633663	-14,44	0,6495048	0,526733	-18,90
265	ILE	1110	ARG	0,695049	0,396040	-43,02	0,738614	0,467327	-36,73	0,6396038	0,370297	-42,11
266	ALA	1086	PHE	0,623762	0,487129	-21,90	0,600000	0,510891	-14,85	0,5960394	0,459406	-22,92
266	ALA	1110	ARG	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,4534656	0,437624	-3,49
267	PHE	1110	ARG	0,461386	0,336634	-27,04	0,560396	0,457426	-18,37	0,4712874	0,310891	-34,03
267	PHE	1115	ILE	0,9168316	0,8534652	-6,91	0,944574	0,851485	-9,86	0,8970296	0,853465	-4,86

Hydrogen Bonds			WT_ACT	F978A_ACT	Variation %	WT_CYC	F978A_CYC	Variation %	WT_VIN	F978A_VIN	Variation %
ICH2-NBD2		life	301,800	325,000	7,69	497,600	398,800	-19,86	335,881	317,8	-5,38
		<N>	1,142	1,053	-7,79	1,367	1,774	29,83	1,0948	1,285	17,37
		DG	-18,851	-19,120	1,43	-20,078	-19,578	-2,49	-18,9814	-18,648	-1,76

TMD-NBD interface				WT_ACT	F978A_ACT		WT_CYC	F978A_CYC		WT_VIN	F978A_VIN	
ICH3		NBD2		Frequency	Frequency	Variation %	Frequency	Frequency	Variation %	Frequency	Frequency	Variation %
800	ASP	1086	PHE	0,754455	0,720792	-4,46	0,734653	0,726732	-1,08	0,699010	0,768317	9,92
800	ASP	1087	TYR	1,000000	0,996040	-0,40	0,998020	1,000000	0,20	0,990099	0,984158	-0,60
801	VAL	1086	PHE	0,928713	0,893069	-3,84	0,899010	0,893069	-0,66	0,914851	0,869247	-4,98
801	VAL	1087	TYR	0,526733	0,433663	-17,67	0,489109	0,508911	4,05	0,485149	0,516831	6,53
802	SER	1087	TYR	0,736633	0,516831	-29,84	0,685148	0,643564	-6,07	0,693069	0,568317	-18,00
805	ASP	1044	TYR	0,904950	0,851485	-5,91	0,936634	0,904950	-3,38	0,922772	0,869307	-5,79
805	ASP	1046	THR	0,998020	0,998020	0,00	0,998020	0,998020	0,00	1,000000	1,000000	0,00

Hydrogen Bonds			WT_ACT	F978A_ACT	Variation %	WT_CYC	F978A_CYC	Variation %	WT_VIN	F978A_VIN	Variation %
ICH3-NBD2		life	858,200	471,800	-45,02	822,600	388,800	-52,74	702,618	547,400	-22,09
		<N>	2,849	2,709	-4,93	2,841	2,568	-9,61	2,871	2,509	-12,61
		DG	-21,282	-19,803	-6,95	-21,464	-19,488	-9,21	-20,761	-19,792	-4,67

TMD-NBD interface				WT_ACT	F978A_ACT		WT_CYC	F978A_CYC		WT_VIN	F978A_VIN	
ICH4		NBD1		Frequency	Frequency	Variation %	Frequency	Frequency	Variation %	Frequency	Frequency	Variation %
905	ARG	401	TYR	0,174258	0,732655	320,44	0,217822	0,617822	183,64	0,220222	0,691089	213,81
905	ARG	438	GLN	0,451485	0,855426	89,47	0,215842	0,772277	257,80	0,245545	0,920792	275,00
905	ARG	441	GLN	1,000000	1,000000	0,00	0,964356	1,000000	3,70	1,000000	0,998020	-0,20
908	VAL	467	ARG	0,566337	0,752475	32,87	0,568317	0,726732	27,87	0,609901	0,679208	11,36
909	SER	441	GLN	0,992079	1,000000	0,80	0,998020	0,996040	-0,20	1,000000	0,994059	-0,59
911	THR	467	ARG	0,617830	0,904950	46,47	0,352475	0,976238	176,97	0,459406	0,847525	84,48
912	GLN	464	ARG	0,643564	0,4752474	-26,15	0,409901	0,400000	-2,42	0,435644	0,398040	-8,63

Hydrogen Bonds			WT_ACT	F978A_ACT	Variation %	WT_CYC	F978A_CYC	Variation %	WT_VIN	F978A_VIN	Variation %
ICH4-NBD1		life	621,200	446,400	-28,14	476,000	347,800	-26,93	454,402	324,200	-28,65
		<N>	2,828	3,743	32,33	2,457	3,434	39,73	2,556	3,531	38,12
		DG	-20,458	-19,897	-2,74	-19,526	-19,268	-1,32	-19,648	-18,998	-3,31

The variation percentage is estimated based on the mean frequencies of five replica systems for the WT and F978A variant (*holo* systems).

Table S32. Variation in the contact frequencies and hydrogen bond (HB) parameters between the ICH-NBD residues for the WT and ΔF335 variant (*holo* systems).

TMD-NBD interface				WT_CYC	ΔF335_CYC		WT_VLS	ΔF335_VLS		WT_DOX	ΔF335_DOX	
ICH1		NBD1		Frequency	Frequency	Variation %	Frequency	Frequency	Variation %	Frequency	Frequency	Variation %
160	ILE	443	LEU	0,187183	0,613861	227,95	0,396040	0,568317	43,50	0,183518	0,513282	179,69
164	ASP	404	ARG	0,984139	0,984158	0,00	0,910891	0,976238	7,17	0,986119	0,980198	-0,60
164	ASP	405	LYS	0,991074	0,000000	LOST	0,825743	0,000000	LOST	0,940535	0,000000	LOST

Hydrogen Bonds		WT_CYC	ΔF335_CYC	Variation %	WT_VLS	ΔF335_VLS	Variation %	WT_DOX	ΔF335_DOX	Variation %
ICH1-NBD1	life	779,200	217,800	-72,05	427,400	227,200	-46,84	724,200	440,200	-39,22
	<N>	2,035	2,121	4,22	2,109	2,067	-1,97	2,020	1,957	-3,09
	DG	-21,072	-18,168	-13,78	-19,843	-18,241	-8,07	-21,022	-18,994	-9,65

TMD-NBD interface				WT_CYC	ΔF335_CYC		WT_VLS	ΔF335_VLS		WT_DOX	ΔF335_DOX	
ICH2		NBD2		Frequency	Frequency	Variation %	Frequency	Frequency	Variation %	Frequency	Frequency	Variation %
262	ARG	1077	SER	0,447525	0,536633	19,91	0,461386	0,536633	16,31	0,000000	0,000000	0,00
262	ARG	1081	GLN	0,407921	0,558416	36,89	0,415842	0,570297	37,14	0,449505	0,335606	-25,34
262	ARG	1086	PHE	0,514851	0,362376	-29,62	0,000000	0,000000	0,00	0,481188	0,308911	-35,80
263	THR	1117	SER	0,740594	0,544554	-26,47	0,544548	0,631677	16,00	0,744554	0,547729	-26,44
265	ILE	1110	ARG	0,738614	0,762376	3,22	0,758416	0,831683	9,66	0,716831	0,552326	-22,95
266	ALA	1086	PHE	0,600000	0,516826	-13,86	0,580198	0,512872	-11,60	0,532673	0,456622	-14,28
266	ALA	1110	ARG	0,000000	0,000000	0,00	0,000000	0,000000	0,00	0,459406	0,382176	-16,81
267	PHE	1110	ARG	0,560396	0,613861	9,54	0,520792	0,431683	-17,11	0,477228	0,489893	2,65
267	PHE	1115	ILE	0,944574	0,887129	-6,08	0,912871	0,899010	-1,52	0,934653	0,763404	-18,32

Hydrogen Bonds		WT_CYC	ΔF335_CYC	Variation %	WT_VLS	ΔF335_VLS	Variation %	WT_DOX	ΔF335_DOX	Variation %
ICH2-NBD2	life	497,600	349,400	-29,78	303,600	376,200	23,91	501,600	461,200	-8,05
	<N>	1,367	1,307	-4,36	1,151	1,451	26,14	1,309	1,217	-7,09
	DG	-20,078	-19,129	-4,72	-18,746	-19,387	3,42	-20,198	-19,201	-4,93

TMD-NBD interface				WT_CYC	ΔF335_CYC		WT_VLS	ΔF335_VLS		WT_DOX	ΔF335_DOX	
ICH3		NBD2		Frequency	Frequency	Variation %	Frequency	Frequency	Variation %	Frequency	Frequency	Variation %
800	ASP	1086	PHE	0,734653	0,760396	3,50	0,734653	0,720792	-1,89	0,718812	0,838371	16,63
800	ASP	1087	TYR	0,998020	0,988119	-0,99	1,000000	0,998020	-0,20	0,968317	0,988419	2,08
801	VAL	1086	PHE	0,899010	0,914911	1,77	0,914851	0,932673	1,95	0,891089	0,913151	2,48
801	VAL	1087	TYR	0,489109	0,277228	-43,32	0,526731	0,453465	-13,91	0,510891	0,290996	-43,04
802	SER	1087	TYR	0,685148	0,475248	-30,64	0,708911	0,575577	-18,81	0,655445	0,467719	-28,64
805	ASP	1044	TYR	0,936634	0,623762	-33,40	0,922772	0,744554	-19,31	0,930693	0,534672	-42,55
805	ASP	1046	THR	0,998020	0,998020	0,00	0,996040	0,998020	0,20	0,998020	0,994059	-0,40

Hydrogen Bonds		WT_CYC	ΔF335_CYC	Variation %	WT_VLS	ΔF335_VLS	Variation %	WT_DOX	ΔF335_DOX	Variation %
ICH3-NBD2	life	822,600	288,000	-64,99	445,000	290,000	-34,83	565,400	360,200	-36,29
	<N>	2,841	2,658	-6,47	2,836	2,511	-11,45	2,686	2,590	-3,58
	DG	-21,464	-18,857	-12,15	-19,752	-18,792	-4,86	-20,742	-19,240	-7,24

TMD-NBD interface				WT_CYC	ΔF335_CYC		WT_VLS	ΔF335_VLS		WT_DOX	ΔF335_DOX	
ICH4		NBD1		Frequency	Frequency	Variation %	Frequency	Frequency	Variation %	Frequency	Frequency	Variation %
905	ARG	401	TYR	0,217822	0,669307	207,27	0,215842	0,641584	197,25	0,255446	0,740949	190,06
905	ARG	438	GLN	0,215842	0,770297	256,88	0,304951	0,661380	116,88	0,271287	0,848459	212,75
905	ARG	441	GLN	0,964356	1,000000	3,70	1,000000	0,998020	-0,20	0,998020	0,992266	-0,58
908	VAL	467	ARG	0,568317	0,786138	38,33	0,532674	0,790099	48,33	0,560396	0,784506	39,99
909	SER	441	GLN	0,998020	1,000000	0,20	0,986139	0,996040	1,00	0,980198	0,996133	1,63
911	THR	467	ARG	0,352475	0,974257	176,40	0,398020	0,978218	145,77	0,233664	0,974351	316,99
912	GLN	464	ARG	0,000000	0,000000	0,00	0,576238	0,259406	-54,98	0,000000	0,000000	0,00

Hydrogen Bonds		WT_CYC	ΔF335_CYC	Variation %	WT_VLS	ΔF335_VLS	Variation %	WT_DOX	ΔF335_DOX	Variation %
ICH4-NBD1	life	476,000	225,000	-52,73	448,200	319,000	-28,83	365,200	229,800	-37,08
	<N>	2,457	3,254	32,40	2,980	3,176	6,59	2,544	3,491	37,23
	DG	-19,526	-18,173	-6,93	-19,721	-18,856	-4,39	-19,324	-18,230	-5,66

The variation percentage is estimated based on the mean frequencies of five replica systems for the WT and ΔF335 variant (*holo* systems).

ANNEX S4

"Probing the allosteric modulation of P-glycoprotein: A medicinal chemistry approach towards the identification of non-competitive P-gp inhibitors."

TABLE OF CONTENTS

MATERIAL AND METHODS

1. Computational Fragment-Based Drug discovery.

1.1 Fragments selection and parametrization. Fragments were obtained from predicted ICH–NBD binders, drawn in MarvinSketch v19.18 and exported to MOE software for further protonation and minimization using the MMFF94x force-field (adjusting hydrogen and lone pairs by default). Following, each fragment was parametrized in the PRODRG online server and manually curated, according to the GROMOS96 54a7 force field, adjusting partial charges to the AM1-BCC scheme calculated with Antechamber v1.27.

1.2 Construction of the molecular dynamics (MD) systems. Based on the previously published human P-gp homology model, only the cytoplasmic portion of the N- and C-terminal (concerning NBD1 and NBD2, respectively) were used. The preparation of the initial structures was performed using VMD and MOE software packages. Each NBD was inserted in a simulation box with dimensions xyz of $10 \times 10 \times 9$ nm³. Following, all systems were solvated with an adequate number of SPC water molecules. For each NBD, five MD systems were built, each one comprising six fragment types and five copies of each (for a total of 30 fragments), by randomly inserted them in the surrounding water environment using *gmx insert-molecules* module available in GROMACS software. Any overlapping waters were automatically removed, and any charge excess were neutralized by adding an adequate number of counter-ions (sodium or chlorine).

1.3. Simulated-annealing molecular dynamics simulations (saMD). After an energy minimization step, a short 10 ps *NVT* ensemble run followed by a 5 ns *NpT* ensemble MD simulations were performed for equilibrating temperature (303 K) and pressure (1 bar), respectively, while keeping all protein atoms spatially restrained. Following, a 50 ns MD run was performed for each system, using a simulated annealing protocol applied to the fragments, solvent and ions to decrease the probability of non-specific binding to each of the NBDs. Herein, two heating cycles (303 K to 323 K) were performed, increasing the temperature during 1 ns ($t = 0$ ns and $t = 25$ ns) and decreasing over a period of 3 ns ($t = 2$ ns and $t = 27$ ns). Due to the simulated annealing procedure, the protein's alpha carbons were kept restrained throughout the whole MD simulation to prevent unfoldings. For each system, five replicates were performed (25 MD systems per NBD, in a total of 1.25 μs of simulation time per NBD).

1.4 Analysis and identification of allosteric drug-binding site(s). The last 20 ns of each replicate were concatenated using *gmx trajcat* tool available in GROMACS to obtain a single trajectory file of

100 ns per NBD. An isosurface cut-off of 0.1 was used to fine-tune the generated occupancy maps. Visual inspection of the fragments' occupancy maps was performed in VMD, and only those in close vicinity of each TMD-NBD interface were considered for the identification of possible binding sites next to the ICH domains. Finally, the properties of each putative aDBS such as lining residues, pocket volume, residues distribution and mean polarities were evaluated using the EPOS[BP] software (default parameters), and compared with the modulator site (M-site).

2. Molecular Docking studies. Molecular docking was performed with the AutoDock VINA v1.1.2 in both isolated NBDs. Considering the location of the selected occupancy maps, a docking box was defined to include all regions between the ICHs and the whole ATP-binding site, with grid dimensions *xyz* of $18.75 \times 18.75 \times 18.75$ Å^3 (NBD1) or $18.75 \times 22.50 \times 18.75$ Å^3 (NBD2), and centered at the *xyz* dimensions of $40.27 \times 53.20 \times 55.31$ Å^3 (NBD1) or $63.40 \times 47.61 \times 55.70$ Å^3 (NBD2), respectively. The database consisted in small in-house libraries of flavonoids ($n = 28$). Additionally, due to experimental evidences that BUM and NPA bind in a region located at the ICH-NBD interfaces, these molecules were also included in the docking studies as references. Ten docking poses were generated for each ligand. Molecules were manipulated as previously stated, but saved in the PDB format. For further usage in the AutoDock VINA docking software, both receptor (P-gp) and ligands were converted to the PDBQT format with AutoDockTools v1.5.6rc. Visual inspection of the docking poses was made in MOE, and the protein-ligand interactions were calculated using BINANA v2.1 and LigPlot v4.5.3 software packages.

3. Free-energy calculations and analysis of compound 23. After a short energy minimization run to minimize clashes between the ligand and the protein, a 100 ns MD run was performed using the top-ranked docking pose of compound **23** at the NBD2, using the final configuration of the NBD2 after the MD simulations. From the initial MD simulation, four replicates of 50 ns MD run each were performed by selecting distinct snapshots along the simulation (based on the RMSD evolution during the first MD run), namely at 40, 70 and 90 ns MD simulation run. Only the last 30 ns of each MD simulation were considered as production run and used for analysis. The visual inspection of the binding mode(s) was performed in MOE, and relative free-energies of binding (ΔG_{MD}) were calculated using the *g_mmpbsa* tool.

4. MD simulation parameters. All MD simulations were done with GROMACS v2016.x package. All *NVT* equilibration runs were performed at 303 K using the Velocity-rescale (V-rescale) thermostat. The Nosé-Hoover thermostat and the Parrinello-Rahman barostat for temperature (303 K) and

pressure (1 bar), respectively, were applied in all *NpT* runs. Pressure equilibration was achieved through an isotropic pressure coupling, with the systems' compressibility set to 4.5×10^{-5} bar^{-1}. All bond lengths were constrained using the LINCS or SETTLE (for water molecules) algorithms and PBC conditions were applied to all dimensions. The Particle Mesh Ewald (PME) with cubic interpolation was employed, with a cut-off radius of 12 Å for both electrostatic and van der Waals interactions and an FFT grid spacing of 0.16 for long range electrostatics. Group-based and Verlet cut-off schemes were applied for the calculation of non-bonded interactions on CPU or GPU, respectively.

5. Experimental studies: ATPase assays

5.1 Compounds

A small in-house library of thioxanthone (*n* = 8) and flavanone (*n* = 28) derivatives were used to evaluate the effect of such compounds in P-gp ATPase activity using PgpGlo^{TM} P-gp ATPase assays. The compounds were tested at the initial concentration of 200 µM (thioxanthones) and 100 µM (flavonoids). Additionally, the compounds BUM (1 mM) and spiropedroxodiol (50 µM) were also included in the ATPase assays as positive controls for allosteric and competitive P-gp inhibition mechanisms, respectively. All compounds were dissolved in DMSO.

5.2 ATPase assays — general concepts

The ATPase activity of the P-gp ATP-dependent drug efflux pump was measured by using the PgpGlo^{TM} Assay Systems kit according to the manufactures' recommendation. Briefly, this assay relies on the ATP dependence of the light-generating reaction of firefly luciferase using recombinant human Pgp in a cell membrane fraction. The recombinant human P-gp in cell membrane fraction (25 µg/well) was incubated with 5mM of MgATP for a period of 40 minutes at 37 °C. The reaction is stopped by adding the luciferase-based ATP Detection Reagent. At this stage, luciferase reaction is initiated, and the luminescence signal is produced. Luminescence changes in direct proportion to the unmetabolized ATP concentration (ATP not consumed by P-gp). Reactions where no ATP is consumed generate the brightest signals (e.g. sodium orthovanadate; Na_3VO_4), while reactions where ATP has been consumed have relatively lower luminescent signals, reflecting lower unmetabolized ATP concentrations (e.g. verapamil).

5.3 ATPase assays layout and analysis of the results

The assay was performed in a 96-well plate layout, which comprised four different sample types: i)

non-treated (NT), ii) treated with Na$_3$VO$_4$, iii) treated with verapamil (VER), and iv) treated with test compound (TC) in the presence of the substrate verapamil (TC+VER). All the samples were tested in triplicate according to the technical bulletin.

A first screening was performed to evaluate the capacity of compounds to inhibit the P-gp drug-stimulated ATPase activity in the presence of 500 µM of verapamil, a P-gp substrate that stimulates P-gp ATPase activity. In the following step, the compounds that showed the highest inhibition of the VER-stimulated ATPase activity were selected for calculating the respective maximal inhibitory concentration (IC$_{50}$) values by decreasing the test compounds' concentration using a two-fold dilution, while maintaining verapamil concentration unchanged (500 µM), in agreement with the experimental protocol provided by the manufacturer.

The luminescence signal was read in average relative light units (RLU), and compared with the Na$_3$VO$_4$ – a strong P-gp inhibitor – treated samples. The difference between the average luminescent signals from Na$_3$VO$_4$-treated samples (RLU$_{NA3VO4}$) and NT samples (RLU$_{NT}$) corresponds to the basal Pgp ATPase activity, and calculated as follows:

$$RLU_{Na_3VO_4} - RLU_{NT} = \Delta RLU_{basal}$$

The difference between the average luminescent signals from Na$_3$VO$_4$-treated samples (RLU$_{NA3VO4}$) and VER-treated samples (RLU$_{VER}$) reflects the VER-stimulated ATPase, and can be determined as follows:

$$RLU_{Na_3VO_4} - RLU_{VER} = \Delta RLU_{VER}$$

Finally, the difference between the average luminescent signals from Na$_3$VO$_4$-treated samples (RLU$_{Na3VO4}$) and TC+VER-treated samples (RLU$_{TC}$) reflects the ability of TC to inhibit VER-stimulated ATPase activity, and can be calculated by the same equation:

$$RLU_{Na_3VO_4} - RLU_{TC+VER} = \Delta RLU_{TC+VER}$$

Table S1: Fragments obtained from each molecule.
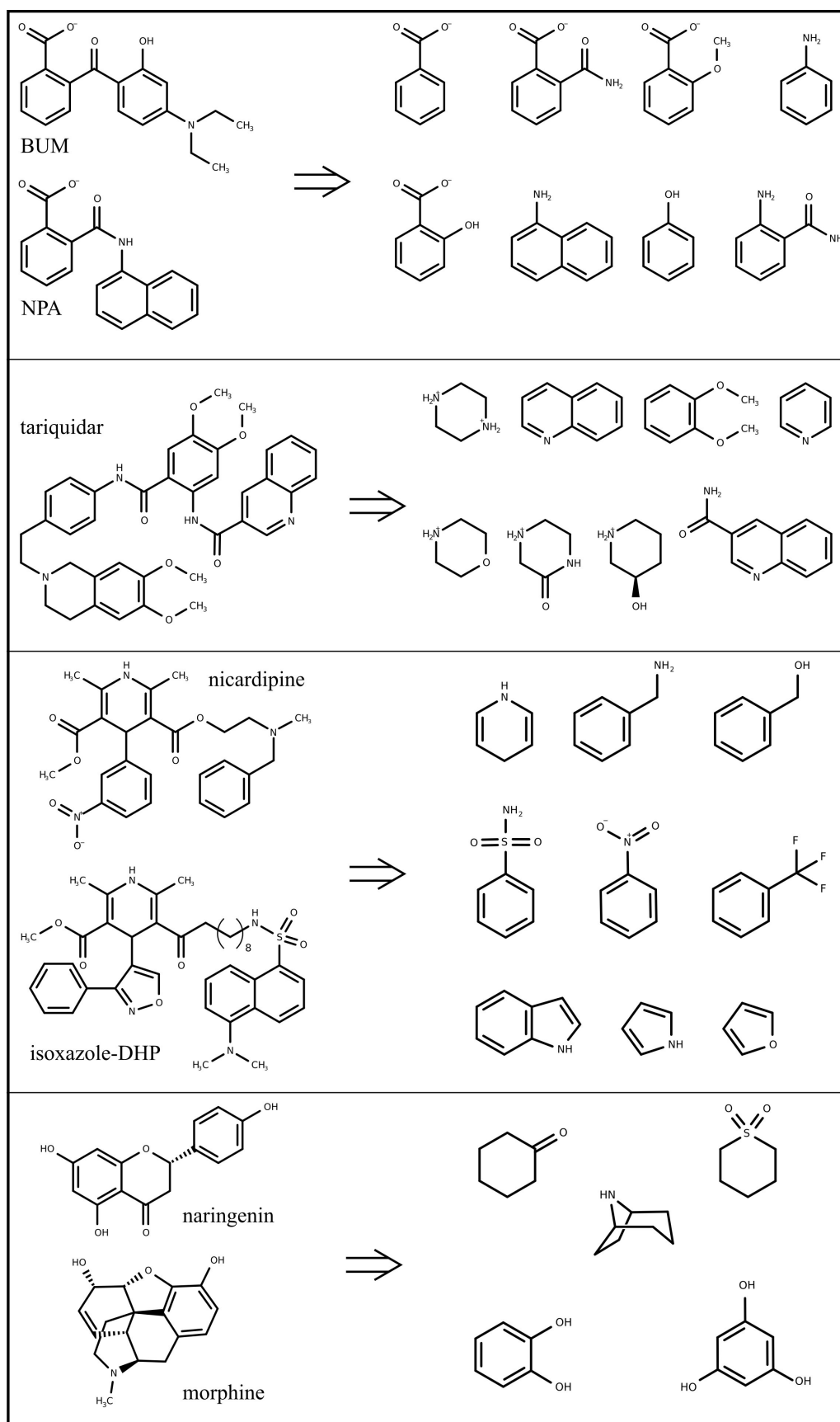
Table S2: Fragments included in each MD system.

System	Fragments (IUPAC name)	Mol.Formula	SMILES
1	Naphthalen-1-amine	$C_{12}H_{15}N$	NC1=CC=CC2=CC=CC=C12
	Aniline	C_6H_7N	NC1=CC=CC=C1
	Pyridine	C_5H_5N	C1=CC=NC=C1
	Benzoate	$C_7H_5O_2^-$	[O-]C(=O)C1=CC=CC=C1
	Phenol	C_6H_6O	OC1=CC=CC=C1
	1,4-dihydropyridine	C_5H_7N	C1C=CNC=C1
2	Benzene-1,2-diol	$C_6H_6O_2$	OC1=C(O)C=CC=C1
	Benzene-1,3,5-triol	$C_6H_6O_3$	OC1=CC(O)=CC(O)=C1
	1H-indole	C_8H_7N	N1C=CC2=C1C=CC=C2
	Nitrobenzene	$C_6H_5NO_2$	[O-][N+](=O)C1=CC=CC=C1
	1H-pyrrole	C_4H_5N	N1C=CC=C1
	Quinoline-3-carboxamide	$C_{10}H_8N_2O$	NC(=O)C1=CC2=CC=CC=C2N=C1
3	2-carbamoylbenzoate	$C_8H_6NO_3^-$	NC(=O)C1=C(C=CC=C1)C([O-])=O
	1,2-dimethoxybenzene	$C_8H_{10}O_2$	COC1=C(OC)C=CC=C1
	Furan	C_4H_4O	O1C=CC=C1
	Phenylmethanamine	C_7H_9N	NCC1=CC=CC=C1
	Quinoline	C_9H_7N	C1=CC2=CC=CN=C2C=C1
	Trifluormethylbenzene	$C_7H_5F_3$	FC(F)(F)C1=CC=CC=C1
4	2-aminobenzamide	$C_7H_8N_2O$	NC(=O)C1=C(N)C=CC=C1
	2-carboxyphenolate	$C_7H_5O_3^-$	OC1=C(C=CC=C1)C([O-])=O
	Cyclohexanone	$C_6H_{10}O$	O=C1CCCCC1
	2-methoxybenzoate	$C_8H_7O_3^-$	COC1=C(C=CC=C1)C([O-])=O
	Phenylmethanol	C_7H_8O	OCC1=CC=CC=C1
	Piperazine-1,4-diium	$C_4H_{12}N_2^{2+}$	C1C[NH2+]CC[NH2+]1
5	Morpholin-4-ium	$C_4H_{10}NO^+$	C1COCC[NH2+]1
	Benzenesulfonamide	$C_6H_7NO_2S$	NS(=O)(=O)C1=CC=CC=C1
	piperidin-1-ium-3-ol	$C_5H_{12}NO^+$	OC1CCC[NH2+]C1
	3-oxopiperazine-1,4-diium	$C_4H_9N_2O^+$	O=C1C[NH2+]CCN1
	8-azoniabicyclo[3.2.1]octane	$C_7H_{14}N^+$	C1CC2CCCC1[NH2+]2
	Thiane 1,1-dioxide	$C_5H_{10}O_2S$	O=S1(=O)CCCCC1

Table S3: Experimentally evaluated compounds.
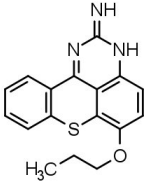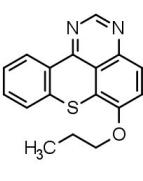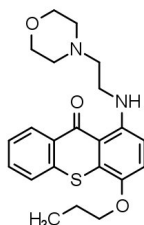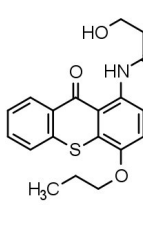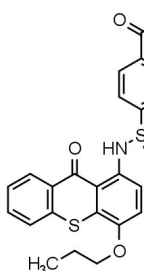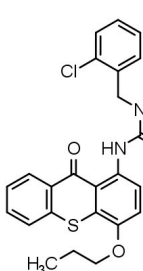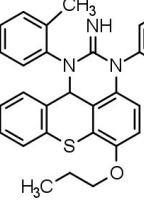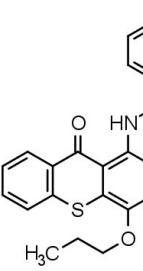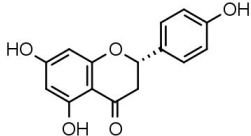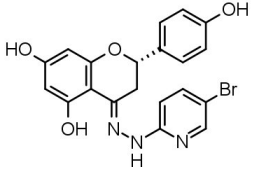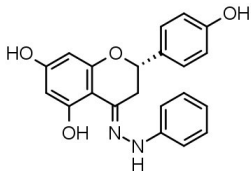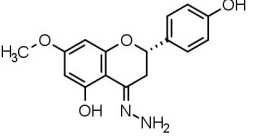
TX1		TX2	
TX3		TX4	
TX5		TX6	
TX7		TX8	
Naringenin		2	
3		4	
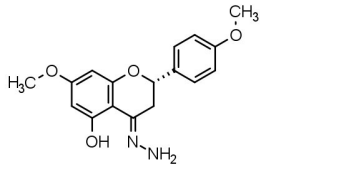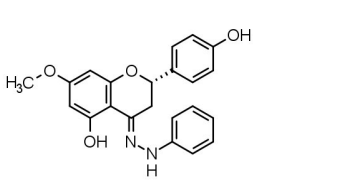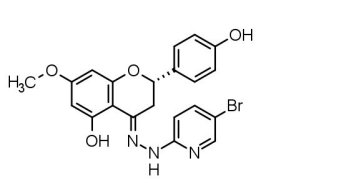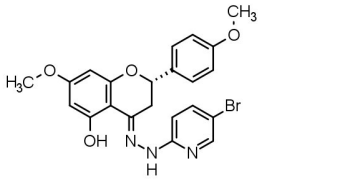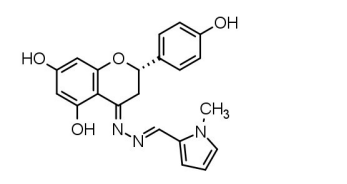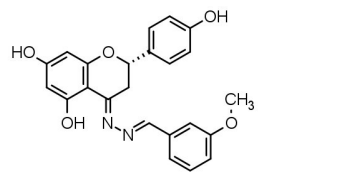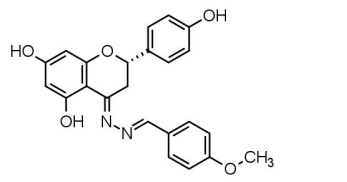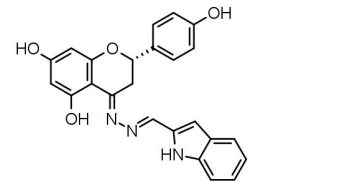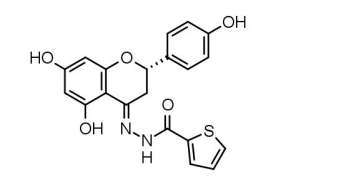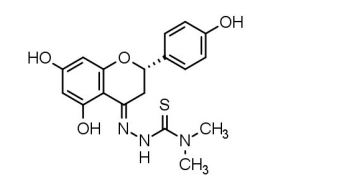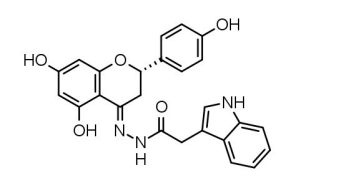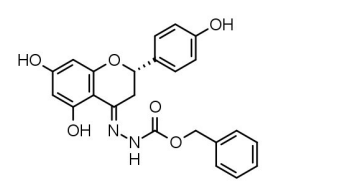
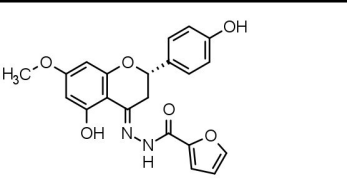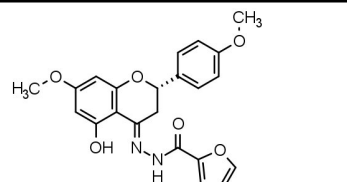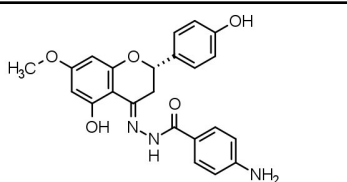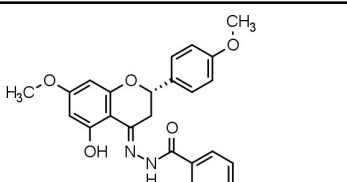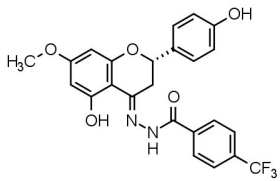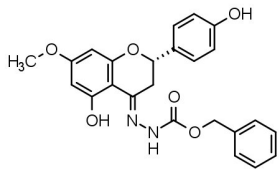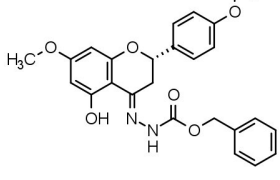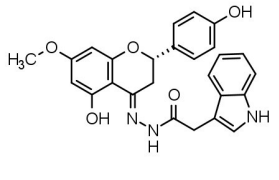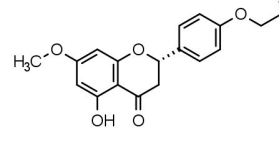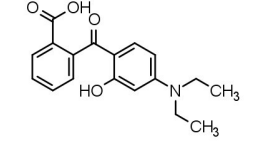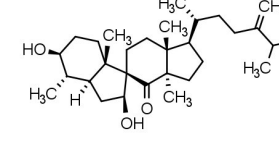| 5		6	
7		8	
9		10	
11		12	
13		14	
15		16	
17		18	
29		20	

Table S3 (continued)
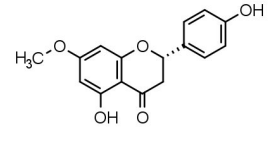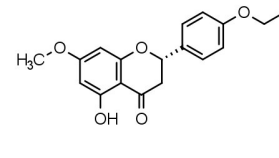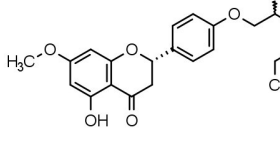
21		22	
23		24	
25		26	
27		28	
BUM		spiropedroxodiol	

Table S4. N-terminal residue interactions with top-ranked docked poses of all tested molecules and total number of contacts.

Molecule Name	binding energy (kcal/mol)	Hydrogen Bonds		Hydrophobic Contacts		Pi-pi	T-stacking	Cation-pi	Salt Bridges	Contacts	
		ICH	NBD	ICH	NBD					ICH	NBD
REFS BUM	-5.8	--	S434	E902, N903	V437, D555, I585	--	--	--	--	15	12
NPA	-7.1	--	--	E902, N903	V437, S474, D555, I585	--	--	--	--	6	17
Naringenin (1)	-6.6	I160	--	I160, F163, V168, E902, F904, R905	--	--	--	R905	--	14	10
7-OMe-Naringenin (25)	-6.5	--	--	E902, N903, T906	V472, V437, S474, D555, I585	--	--	--	--	9	11
Thioxantones TX1	-7.1	D164	--	I160, F163, D164, V168, I901, E902, F904, R905	--	--	--	R905	--	10	16
TX2	-6.8	--	--	I160, F163, D164, V168, I901, E902, F904, R905	--	--	--	R905	--	10	16
TX3	-6.4	R905	S434, E476	--	G430, S434, V437, V472, S474, E476, D555, I585	--	--	--	--	5	22
TX4	-6.3	--	E476	--	V437, S474, E476, D555	--	--	--	--	4	9
TX5	-7.1	R905	--	E902, N903, R905, T906	V437, S474, E476, D555, I585	--	--	--	--	15	17
TX6	-7.8	R905	--	F163, I160, E902, F904, R905	S474, E476	--	--	R905	--	14	11
TX7	-6.7	--	--	E902, N903	V437, S474, E476, D555	--	--	R905	--	8	10
TX8	-6.8	--	--	E902, N903, R905, T906	S434, V437, S474, I585	--	--	--	--	17	10
Hydrazones 2	-7.8	--	--	N903	S434, V437, V472, S474, D555, I585	--	--	--	--	7	17
3	-7.8	--	S434, E476, D555	R905	S434, V437, V472, S474, D555, I585	--	--	R905	--	8	17
4	-6.7	--	--	N903	V437, V472, S474, D555, I585	--	--	--	--	7	11
5	-6.4	--	--	E902, N903	V437, V472, S474, L553, L554, D555	--	--	--	--	9	19
6	-7.6	R905	S434	N903, R905	S434, V437, V472, S474, D555, I585	--	--	R905	--	8	14
7	-7.5	--	S434, D555	R905	V437, V472, S474, E476, D555, I585	--	--	--	--	7	16
8	-7.3	--	S434	E902, N903, T906	V437, V472, S474, L553, L554, D555, I585	--	--	--	--	9	17
Azines 9	-7.8	--	E476	R905, T906	Y401, G430, C431, S434, V437, V472, S474, D555, I585	--	--	--	--	8	26
10	-7.4	N903	S434	N903	G430, S434, V437, S474, D555, I585	--	--	--	--	5	26
11	-7.5	--	--	E902, N903, R905	R404, V437, S474, E476	--	--	R905	--	14	18
12	-7.1	--	--	E902	S434, V437, V472, S474, D555, I585	--	--	--	--	6	19
Carbohydrazides 13	-7.1	R905	E476	N903, R905	V437, V472, S474, Q475, D555, I585	--	--	--	--	12	15
14	-7.1	N903	--	R905	C431, S434, V437, V472, S474, D555, I585	--	--	--	--	7	20
15	-7.6	--	--	R905	S434, S474, D555, I585	--	--	--	--	9	14
16	-7.6	R905	--	E902	G430, S434, V437, V472, S474, D555, I585	--	--	--	--	5	19
17	-7.4	--	S434	N903, T906	V437, V472, S474, D555, I585	--	--	--	--	9	14
18	-7.3	--	S434	E902, N903, T906	V437, V472, S474, L553, L554, D555, I585	--	--	--	--	11	20
19	-7.5	R905	V473, E476	--	G430, S434, V437, V472, S474, L554, D555, I585	--	--	--	--	7	22
20	-7.7	--	--	E902, N903	V437, V472, S474, E476, D555, I585	--	--	--	--	11	21
21	-7.4	N903	Y401	R905	Y401, S403, S434, V437, D555, I585	--	--	--	--	10	14
22	-8.1	--	--	E902, N903	S434, V437, E476, D555, I585	--	--	R905	--	9	17
23	-8.1	--	S434	E902, N903	S434, V437, D555, I585	--	--	R905	--	10	20
24	-7.8	R905	S434, E476	R905, T906	Q475, S434, V437, S474, V437, I585	--	--	--	--	10	17
4'-O derivatives 26	-6.3	--	--	E902, T906	S434, V437, S474, E476, I585	--	--	--	--	12	15
27	-6.3	--	--	E902, N903, R905, T906	S434, V437, S474, D555, I585	--	--	--	--	14	23
28	-6.4	--	--	R905, T906	V437, S474, E476, D555, I585	--	--	--	--	9	16

Table S5. NBD2 residue interactions with top-ranked docked poses of all tested molecules and total number of contacts.

Molecule Name		binding energy (kcal/mol)	Hydrogen Bonds		Hydrophobic Contacts		Pi-pi	T-stacking	Cation-pi	Salt Bridges	Contacts	
			ICH	NBD	ICH	NBD					ICH	NBD
REFS	BUM	-6.4	T263	--	A260, R262	E1119, I1121	--	--	R262	R262	20	5
	NPA	-7.2	A259, R262	--	A259, A260, R262	S1117, Q1118	--	--	--	R262	19	7
	Naringenin (**1**)	-6.5	T263	Q1118	R262, A260, I261, T811	--	--	--	R262	--	26	3
	7-OMe-Naringenin (**25**)	-6.2	T263	--	E256, A269	Q1118, E1119	--	--	--	--	17	6
Thioxantones	TX1	-6.2	T263	--	E256	Q1118, D1200	--	--	--	--	7	7
	TX2	-6.2	--	E1119	A260, R262	S1117, Q1118, E1119, A1205	--	--	--	--	9	8
	TX3	-6.2	R262	--	E256, A260, R262, V264	S1117, Q1118, P1120, I1121	--	--	--	--	22	12
	TX4	-6.2	--	I1121	E256, A259, A260, V264	S1117, Q1118, P1120	--	--	--	--	13	14
	TX5	-7.1	--	--	E256, A259, A260, R262, T263, V264	Q1118, E1119	--	--	--	E1119	12	7
	TX6	-6.5	T263	--	A260, I261, R262, T811, D805	--	--	--	R262	--	23	5
	TX7	-6.4	--	--	E256, A259, A260, I261, R262	Q1118, E1119	--	--	--	--	14	5
	TX8	-7.1	--	--	A260, I261, R262	Q1118, E1119, I1121, A1205, L1206	--	--	R262	D1200	15	9
Hydrazones	2	-6.8	--	--	I261, R262, D805	Y1044, T1046, C1074	--	--	R262	--	20	9
	3	-7.2	--	--	A259, A260, I261, R262, T263, T811	--	--	--	R262	--	30	3
	4	-6.5	T263	--	E255, E256, A259, A260	Q1118	--	--	--	--	14	12
	5	-6.3	T263	--	E255, E256, A259, A260	Q1118	--	--	--	--	14	9
	6	-7.0	--	--	A259, A260, I261, R262, T263, D805, T811	--	--	--	R262	--	33	1
	7	-6.7	L258, T811	--	I261, R262, A260, A259	--	--	--	R262	--	22	1
	8	-6.8	L258, T811	--	A259, A260, R262, T263, T811	T1046, D1200	--	--	R262	--	20	5
Azines	9	-6.9	T811	--	T811, D805, I261, A259, R262, T263	--	--	--	R262	--	29	2
	10	-7.0	--	--	A259, A260, I261, R262, T263, D805, T811	Q1118, D1200	--	--	--	--	32	5
	11	-6.9	T263	--	A260, R262, D805, T811	Y1044	--	--	R262	--	25	4
	12	-7.5	R262, T263	Q1118	A260, I261, R262, T811, T810	--	--	--	--	--	27	3
Carbohydrazides	13	-7.1	R262	--	A259, I261, R262, T263, D805, T811	--	--	--	R262	--	40	3
	14	-7.0	--	I1121	E256, A259, A260	S1117, Q1118, E1119, I1121	--	--	--	--	8	13
	15	-8.5	A259, T263, D805	--	A259, A260, I260, D805, T811	--	--	--	--	--	25	2
	16	-8.3	T263, D805	--	A260, I261, R262, D805, T811	G1073	--	--	R262	--	27	3
	17	-7.4	T263	E1119, D1200	A259, A260, I261, R262	Q1118	--	--	R262	--	21	11
	18	-7.1	T263	S1117, Q1118, D1200	A259, A260, E256, T263, V264	S1117, Q1118, P1120, I1121	--	--	--	--	15	19
	19	-7.7	R262, T263	E1119, D1200	A259, A260, R262	Q1118	--	--	R262	--	23	13
	20	-7.6	R262, T263	E1119, D1200	A259, A260, R262	Q1118, A1205	--	--	R262	--	22	14
	21	-7,0	--	E1119	E256, A259, A260, R262	Q1118, E1119	--	--	R262	--	15	7
	22	-7.4	F804	--	E255, A259, A260, D805, T811	--	--	--	--	--	19	3
	23	-7.5	--	--	A260, I261, T811	Q1118	--	--	--	--	17	6
	24	-8.4	T263	E1119, D1200	E256, A260, I261, R262	Q1118, E1119, I1121	--	--	R262	--	22	18
4'-O derivatives	26	-6.8	T263, F804	Q1118, D1200	A260, I261, R262, F804, D805, T810, T811	--	--	--	R262	--	40	4
	27	-6.6	T263	Q1118	A259, A260, I261, R262, F804, D805, T810, T811	--	--	--	R262	--	41	6
	28	-7.4	R262, T263, T811	--	A259, A260, I261, R262, T811	T1046, Q1118	--	--	--	--	41	9

Table S6: ΔRLU values for first-screening compounds.

Thioxanthone derivatives				
Compound	**ΔRLU**		**Compound**	**ΔRLU**
BASAL	3000		**TX4**	19000
VERAPAMIL	23667		**TX5**	19333
TX1	22667		**TX6**	21000
TX2	17667		**TX7**	19000
TX3	32333		**TX8**	16500

Flavanone derivatives				
Compound	**ΔRLU**		**Compound**	**ΔRLU**
BASAL	3000		**15**	12333
VERAPAMIL	15667		**16**	7000
1	12000		**17**	6333
2	13333		**18**	10000
3	8667		**19**	10000
4	10667		**20**	20000
5	7333		**21**	n.d. *
6	12333		**22**	8667
7	17000		**23**	6333
8	12667		**24**	6333
9	13000		**25**	9667
10	12667		**26**	9667
11	11667		**27**	3667
12	8667		**28**	1667
13	8000		**BUM**	7300
14	15000		**spiropedroxodiol**	11967

* not determined due to solubility issues.

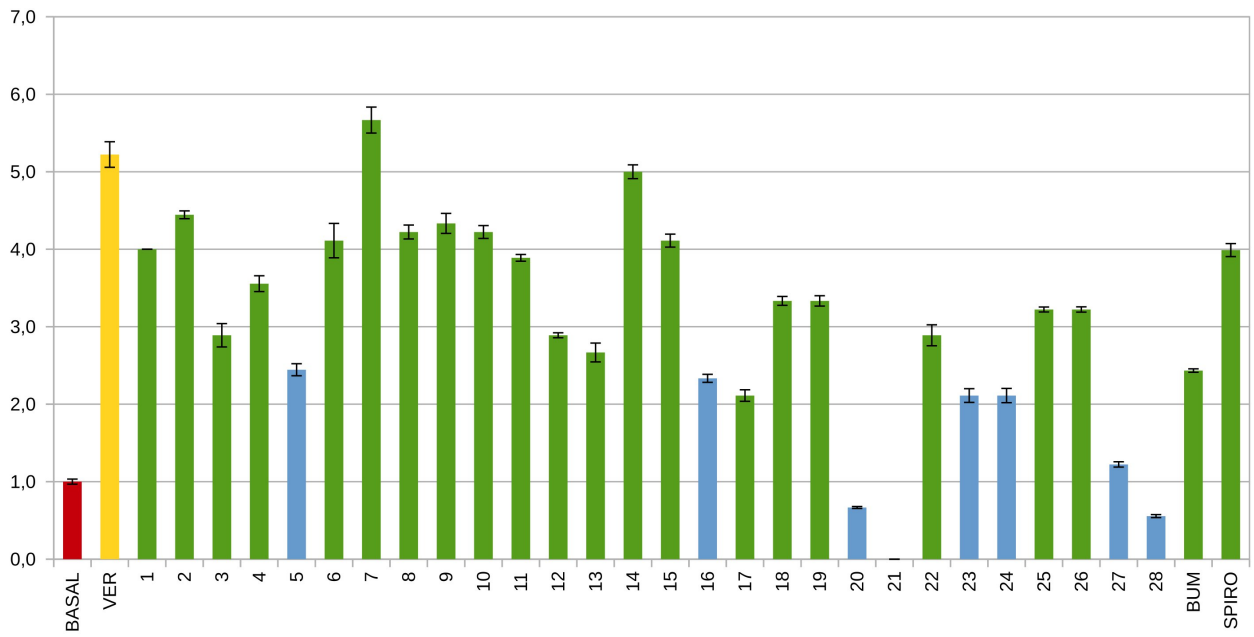**Figure S1:** Normalized RLU values for first-screening of compounds (all compounds, flavonoids).

Figure S2: Normalized RLU values for first-screening of compounds (all compounds, thioxanthones).