Structural and functional study of efflux pumps involved in drug resistance

Lorena Marcela Martinez Jaramillo

Université Claude Bernard Lyon 1

UNIVERSITÉ DE LYON

N° d'ordre : 17 - 2014 Année 2014

THÈSE DE L'UNIVERSITÉ DE LYON

Délivrée par

L'UNIVERSITÉ CLAUDE BERNARD LYON 1

ÉCOLE DOCTORALE INTERDISCIPLINAIRE SCIENCES ET SANTE

DIPLÔME DE DOCTORAT
en
BIOCHIMIE

(arrêté du 7 août 2006)

soutenue publiquement le 14 février 2014

par

Lorena Marcela MARTINEZ JARAMILLO

Structural and functional study of efflux pumps involved in drug resistance

Directeur de thèse
Dr Pierre FALSON

JURY	M. le Docteur Manuel GARRIGOS	Rapporteur
	M. le Docteur Jean-Michel JAULT	Rapporteur
	M. le Professeur Marc LE-BORGNE	Examinateur
	M. le Docteur François ANDRÉ	Invité
	M. le Professeur Geoffrey CHANG	Invité
	M. le Docteur Pierre FALSON	Directeur de thèse

UNIVERSITÉ CLAUDE BERNARD - LYON 1

Président de l'Université	**M. François-Noël GILLY**
Vice-président du Conseil d'Administration	M. le Professeur Hamda BEN HADID
Vice-président du Conseil des Études et de la Vie Universitaire	M. le Professeur Philippe LALLE
Vice-président du Conseil Scientifique	M. le Professeur Germain GILLET
Directeur Général des Services	M. Alain HELLEU

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard	Directeur : M. le Professeur J. ETIENNE
Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux	Directeur : Mme la Professeure C. BURILLON
Faculté d'Odontologie	Directeur : M. le Professeur D. BOURGEOIS
Institut des Sciences Pharmaceutiques et Biologiques	Directeur : Mme la Professeure C. VINCIGUERRA
Institut des Sciences et Techniques de la Réadaptation	Directeur : M. le Professeur Y. MATILLON
Département de formation et Centre de Recherche en Biologie Humaine	Directeur : M. le Professeur P. FARGE

COMPOSANTES ET DÉPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies	Directeur : M. le Professeur F. DE MARCHI
Département Biologie	Directeur : M. le Professeur F. FLEURY
Département Chimie Biochimie	Directeur : Mme le Professeur H. PARROT
Département GEP	Directeur : M. N. SIAUVE
Département Informatique	Directeur : M. le Professeur S. AKKOUCHE
Département Mathématiques	Directeur : M. le Professeur A. GOLDMAN
Département Mécanique	Directeur : M. le Professeur H. BEN HADID
Département Physique	Directeur : Mme S. FLECK
Département Sciences de la Terre	Directeur : Mme la Professeure I. DANIEL
UFR Sciences et Techniques des Activités Physiques et Sportives	Directeur : M. C. COLLIGNON
Observatoire des Sciences de l'Univers de Lyon	Directeur : M. B. GUIDERDONI
Polytech Lyon	Directeur : M. P. FOURNIER
Ecole Supérieure de Chimie Physique Electronique	Directeur : M. G. PIGNAULT
Institut Universitaire de Technologie de Lyon 1	Directeur : M. C. VITON
Institut Universitaire de Formation des Maîtres	Directeur : M. A. MOUGNIOTTE
Institut de Science Financière et d'Assurances	Administrateur provisoire : M. N. LEBOISNE

"Science is the simple word we use to describe a method of organizing our curiosity"

Tim Minchin

REMERCIEMENTS

Ces dernières lignes présagent la fin proche de cette aventure. Laquelle fut difficile par moments et agréable par d'autres, mais surtout très riche personnellement. Elle fut aussi l'occasion de belles rencontres... Bien que ce manuscrit soit présenté comme un travail personnel, je dois mes réussites à de nombreuses collaborations. Ainsi, c'est avec un très grand plaisir que je tiens à remercier toutes les personnes qui m'ont soutenu et qui ont contribué, de près ou de loin, à mener à bien cette thèse. Même si toutes ces personnes ne sont pas citées ici, je ne les ai pas oublié et je les remercie du fond du cœur.

Ma reconnaissance s'adresse en premier lieu à mon directeur de thèse, le Dr. Pierre FALSON, pour avoir pris le risque d'encadrer en thèse un ingénieur chimiste de Colombie, qui n'avait qu'une connaissance très vague en biochimie. Je le remercie de m'avoir fait bénéficier de son expérience dans le domaine des transporteurs ABC mais surtout pour ses conseils, son enthousiasme contagieux, sa motivation débordante et sa persévérance. Je n'oublierai jamais notre complicité scientifique, laquelle nous a permis de travailler de façon complémentaire et efficace pendant ces années.

Je remercie également à le Dr Attilio DI PIETRO, co-directeur de l'équipe « mécanisme et modulation de la résistance aux médicaments », de m'avoir accueillie au sein de son équipe ainsi que les conseils qu'il m'a apportés lors des nombreuses conversations à table.

Je tiens ensuite à remercier très chaleureusement les rapporteurs de cette thèse, les Docteurs Jean-Michel JAULT et Manuel GARRIGOS pour l'intérêt qu'ils ont porté à mon travail et le temps précieux qu'ils y ont consacré. Je suis également très reconnaissante envers le Professeur Marc LE-BORGNE, le Dr. François ANDRE et le Professeur Geoffrey CHANG de m'avoir fait l'honneur d'examiner mon travail. Un grand merci aux membres du comité suivi de thèse, Dr. Jean-Michel JAULT et Dr. Loris G. BAGGETTO qui ont suivi mon travail de très près et dont les conseils m'ont beaucoup servi.

Je veux exprimer ma reconnaissance à la Région Rhône-Alpes et la Ligue Nationale Contre le Cancer pour leurs soutiens financiers qui ont rendu possible ce travail de thèse.

I would like to thank Pr. Geoffrey CHANG from the University of California, San Diego (CA, USA) to give me the opportunity to work with him, and all staff member of its team for welcoming me in the lab. Thank you for sharing with me your knowledge, your skills but also your "American-Indian-Chinese" culture. A special thanks to: Koustav, per be my translator in several times and his lovely wife, Anindita, per her kindness. Rupak, per his helpful discussions and nice concerts, and his charming wife, kinnarys, for giving me her great pizza recipe. Cristina, per the nice margaritas drinks. Carey, my tea partner, for not letting me die in the cold-room. Beverly, for the best shopping experience ever. Finally, to my buddy, Mark, for his friendship, for the rides, and of course for taking care of Mamut.

I spent a wonderful time with you, thanks guys!

Je voudrais également remercier mes collaborateurs d'avoir contribué au bon déroulement de ce travail : À Guillaume Gros et Dr. J. Hasserodt (ENS-Lyon) pour leur implication dans le VIH projet, à Vincent Pavot et Dr. Bernard Verrier (IBCP) pour m'avoir confié leurs super nanoparticules, à Emilie Henin pour son aide précieuse avec toutes ses équations d'enzymo et avec le programme R.

Un grand MERCI à l'ensemble des membres de l'IBCP et des plateaux techniques pour l'aide qu'ils m'ont apporté, en particulier à Sébastien Dussergey et Thibault Andrieu du cytométrie en flux qui m'ont toujours accueilli avec une grande gentillesse. À Isabelle Grosjean du plateau de production et d'analyse de protéines pour m'avoir appris la culture cellulaire, mais surtout pour l'aide qu'elle m'a apportée avec mes soucis de contamination. À tout le personnel administratif et aux gens que j'ai côtoyés pendant ces années… À mes lulus : Alex et Burcu qui sont devenues mes amis.

Je tiens aussi à mentionner le plaisir que j'ai eu à travailler au sein de mon équipe, et j'en remercie tous ses membres, anciens et nouveaux. Hélène pour sa bonne humeur, pour son écoute et les discussions abordées, à Vincent pour son aide et conseils sur Pymol. À Elodie, Sandrine et Rima pour m'avoir encadré durant toute la période de mon stage. À Ophélie, Sarah et Laura, pour le soutient et enthousiasme portés au projet des QZ. À Doriane, pour avoir partagé ses connaissances avec moi. À nos brésiliens, Glaucio, Evelyn, Gustavo et Nathalia pour m'avoir apporté un petit bout de chez moi, à Arnaud pour me faire rire dans les moments difficiles, à Laurianne pour sa gentillesse, à Anne pour les moments de rigolade au labo. À mes stagiaires, Marine et Raphaël, avec qui j'ai partagé et échangé de nombreuses idées. J'espère que je vous ai bien encadré !!!

A Lucia, mi compañera de postre, por todos los momentos que compartimos y por los que vendrán. À Agnès pour son humour et pour avoir fait des congrès ABC une expérience inoubliable. À Mylène pour sa disponibilité, ses conseils, ses corrections et relectures attentives. À Charlotte, ma Chachu, qui a vécu cette thèse à mes côtés et qui m'a toujours prêté une oreille attentive, accordé un sourire et avec qui j'ai partagé des larmes et beaucoup de discussions. Je te considère comme ma sœur !

Merci pour tous les souvenirs qui vont me rester, et les soirées que j'ai passé à vos côtés.…

J'adresse des remerciements particuliers à ceux qui ont eu le courage de relire des «tranches» de cette thèse: Pierre, Odile, Mylène, Burku, Rupak, Chachu. Vous m'avez gentiment torturé !!!

Je remercie tous mes amis (Farah, Brice, Washa, Jean, Diana, Edison, Victor, Lili, Gladis, Alex, José) qui ont toujours été là pour moi.

También me gustaría agradecer a toda mi familia, en especial a mi papá, mamá y hermanas, los cuales siempre han creído en mí y me han apoyado en mis proyectos. A mi segunda familia, a Myriam, Alonso y Salomón, por hacerme hecho sentir en casa desde el principio.

Para terminar, esta aventura no hubiera sido posible sin mi alma gemela, David. Gracias por hacer de mi alguien mejor y darme tanta felicidad.

FOREWORD

Multidrug resistance (MDR) ATP-binding-cassette (ABC) transporters confer a drug resistance phenotype cells against infectious and anticancer agents by extruding drugs out of cells. They act as molecular « pumps » that actively translocate drugs through the plasma membrane by using the energy gained from ATP hydrolysis. The most prevalent of these MDR pumps is the P-glycoprotein (P-gp), followed by multidrug resistance-associated Protein (MRP1 or ABCC1) and the breast cancer resistance protein (BCRP or ABCG2). A functional understanding of these transporters, and strategies to block them is a challenge in the fight against drug resistance to chemotherapeutic treatments for many years.

As part of my thesis, I studied especially P-gp and tried to contribute to the understanding of its function along three axes, which have been the subject of scientific publications and presented at the end of each chapter:

- The first axe attempted to find a molecule capable of inhibiting both HIV-1 protease and P-gp without being transported by P-gp (chapter I). In this way, this molecule will not be evacuated by P-gp and would be accumulated in normally inaccessible areas due to the presence of these pumps (e.g. brain). Additionally, this strategy would prevent the administration of several molecules (e.g. HIV inhibitor and P-gp inhibitor).

- The second axe was the structural study of P-gp by generating new x-ray crystallography of P-gp in complex with several substrates and inhibitors. As well as trapping P-gp in other conformation (e.g., outward-facing conformation) by using molecules that prevent ATP hydrolysis (non-hydrolysable ATP-analogues) or phosphate release (sodium orthovanadate) (chapter II). The assembly of this information would provide a mechanism at the molecular level of drug transport, essential to develop new treatments.

- The last axe was the molecular characterization of two drug-substrate binding sites of P-gp, exemplified with Hoechst 33342 (H-site) and daunorubicin (R-site) (chapter III). The goal was to determine if the inhibitors QZ59 enantiomers (crystallized with P-gp in 2009) have a distinct or common location with the H- and R- drug- binding sites by doing enzymology and docking experiments. This information is essential for designing a new generation of modulators. In addition, data will provide new information about the enzymatic mechanism of P-gp-mediated drug efflux.

But first of all, I will recall some features of multi-drug resistance (MDR), ABC transporters and the three ABC transporters involved in the MDR phenotype: P-gp, ABCG2 and MRP1 (bibliographic review). Then, I will describe the experimental methods used in this study (materials and methods).

This thesis ends with a general discussion summarizing the main results of each chapter and showing their contribution to the understanding of the functional mechanisms of P-glycoprotein

(general conclusion and future directions). In addition, all references cited in the different parts are listed at the end of memory (references).

CONTENTS

ABBREVIATIONS AND UNITS USED

°C	Degree celsius
Å	Angstrom (1 Å = 10^{-10} m)
A. fulgidus	*Archaeoglobus fulgidus*
A. acidocaldarius	*Alicyclobacillus acidocaldarius*
ABC	ATP-binding cassette
ADP	Adenosine-5'-diphosphate
AIC	Akaike's information criterion
AMPPNP	Adenosine-5’(βγ-imido)triphosphate
ATP	Adenosine-5'-triphosphate
BCRP	Breast cancer resistance protein
C. elegans	*Caenorhabditis elegans*
CGM	Cell growth medium
CMC	Critical micelle concentration
CNS	Central nervous system
CSF	Cerebrospinal fluid
DNA	Deoxyribonucleic acid
E. coli	*Escherichia coli*
FDA	Food and drug administration
h	Hill coefficient
H. influenza	H*aemophilus influenza*
H. sapiens	*Homo sapiens*
H-site	Hoechst 33342 transport site
K_A	Activation constant
K_I	Inhibition constant
K_m	Michaelis-Menten constant
K_{SI}	Substrate inhibition constant
L. lactis	*Lactococcus lactis*
M. acetivorans	*Methanosarcina acetivorans*
M. jannaschii	*Methanocaldococcus jannaschii*
M. musculus	*Mus musculus*
MDR	Multidrug resistance
MRP1	Multidrug resistance-associated protein 1
NBD	Nucleotide binding domain
P. furiosus	*Pyrococcus furiosus*
P. horikoshii	*Pyrococcus horikoshii*
P. pastoris	*Pichia pastoris*
PDB	Protein data bank
P-gp	P-glycoprotein

RNA	Ribonucleic acid
R-site	Rhodamine 123 transport site
S. aureus	*Staphylococcus aureus*
S. solfataricus	*Sulfolobus solfataricus*
S. tiphymurium	*Salmonella tiphymurium*
SBP	Substrate binding protein
T. aquaticus	*Thermus aquaticus*
T. litoralis	*Thermococcus litoralis*
T. maritima	*Thermotoga maritima*
TMD	Transmembrane domain
V. cholerae	*Vibrio cholerae*
V_m	*Maximal drug efflux rate*
WT	Wild type
Y. pestis	*Yersinia pestis*

FIGURES

TABLES

BIBLIOGRAPHIC REVIEW

MULTIPLE DRUG RESISTANCE (MDR) PHENOTYPE

"Drug resistance follows the drug like a faithful shadow"
Paul Ehrlich

The ability of cancer cells, bacteria or virus to survive to cytotoxic drugs involves many mechanisms. Some are intrinsic, *i.e.* not responding to drugs from the beginning, others are acquired, due to a resistance coming from drug pressure. Ineffectiveness of chemotherapies are also related to host factors and specific genetic (mutational) or epigenetic (not mutational) target modifications. Usually, after a long-term drug use, resistance appears not only to that drug but also to a series of structurally-unrelated drugs. This phenomenon is known as multi-drug resistance (MDR) [1]. This pervasive and insidious problem becomes increasingly significant in the treatment of various diseases such as cancer as well as fungal, viral and bacterial infections. The following sections will deal with those resistance patterns, particularly the resistance of cancer cells to chemotherapies, the resistance of bacteria to antibiotics and resistance of human immunodeficiency virus (HIV) to anti-HIV-1 chemotherapy.

Cancer cells resistance

Cell division is a prerequisite for reproduction, growth, repair tissue, etc. In prokaryotic organisms together with some lower eukaryotes, the cell cycle occurs *via* a process called binary fission. In eukaryotic organisms the process can occur through mitosis and meiosis depending on the purpose, division or sexual reproduction. This process is carefully controlled by internal factors such as kinetochore attachment, and external factors e.g. growth factors, hormones[a]. However, each division may induce mutations which, in a context of multistep genetic events, can weaken the cell cycle controls and lead to a mass of cells known as tumor. Tumors may be benign, *i.e*. not spreading from their site of origin such as warts. They unfortunately may also be malignant, *i.e.* unregulated and spreading. There are over 100 different types of cancer, each is referred by the type of cell that is initially affected (e.g., Hepatoma, Hepato = liver)[b]. Deaths due to cancer in 2008 were estimated to 7.6 millions; they are projected to 13.1 millions in 2030 [2]. Such increase is due to the high levels of carcinogens (agent that causes cancer) in our life (food, alcohol, tobacco, sedentary life, pollutants in environment, exposure to sunlight, radiation, viruses, etc.); it is also due to the longer time life of some populations.

Local cancers are commonly removed by surgery and radiotherapy, but those for which cells are spreading the treatment is chemotherapy, using targeted drugs, hormones and antibodies. Chemotherapy uses cytotoxic drugs to kill the cancer cells or interfere with one or several checkpoints involved in the cell division process (damage DNA, inhibition of the synthesis of new DNA strands, mitosis arrest, etc.). However such drugs also target healthy cells, causing side effects

[a] According to the website: https://www.boundless.com/biology/cancer-and-disease/the-cell-cycle-is-tightly-regulated/factors-instructing-cells-to-divide-or-stop-dividing/
[b] According to the website: http://www.medicalnewstoday.com/info/cancer-oncology/

such as feeling sick, hair lost, infection, etc. These drugs can be divided in two groups in respect of their impact, broad or specific, on the cell cycle (Figure 1). These groups can be subdivided in smaller groups, based on factors such as the chemical structure, the cytotoxic mechanism and the relationship with another drug. Some of them act in several ways and thus belong to more than one group (Appendix Table 1).

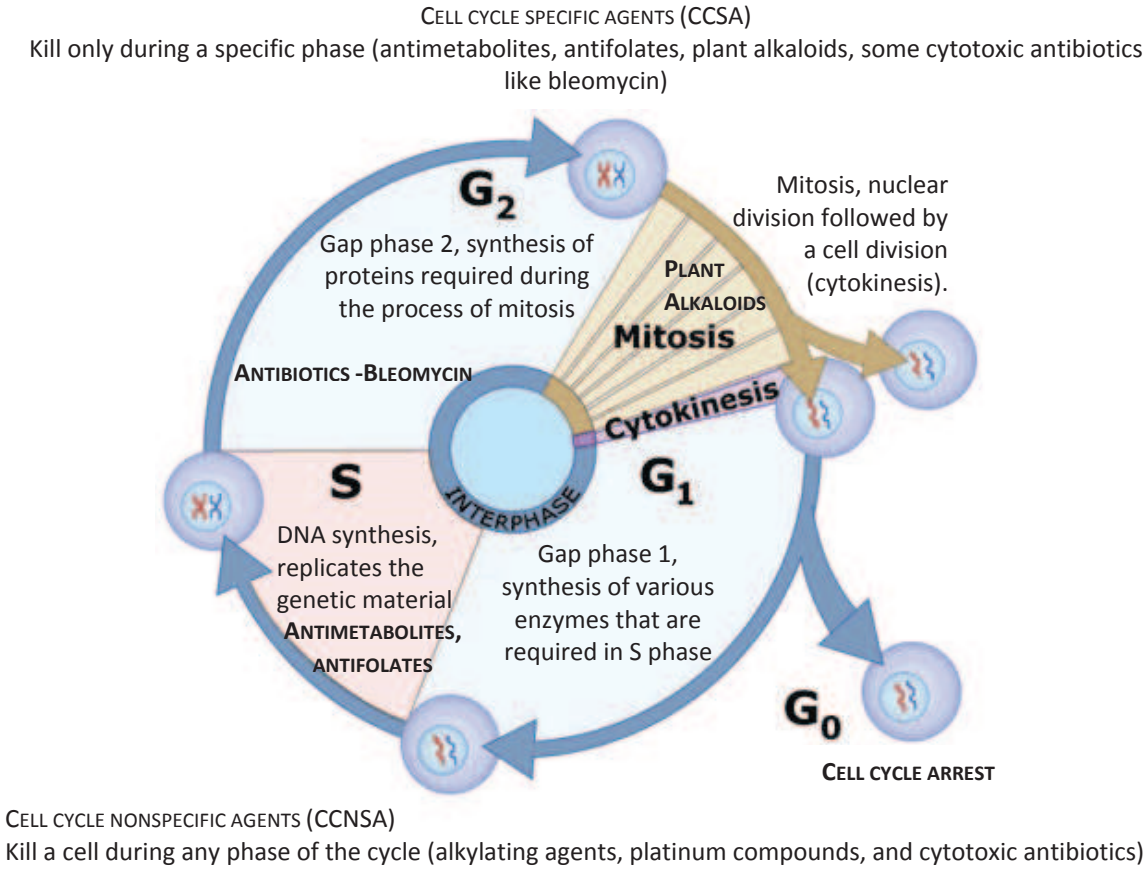


Figure 1. Representation of the cell division cycle with the cell cycle specific and nonspecific agents groups.
Adapted from [2, 3].

And human cancer cells continue to replicate despite the presence of chemotherapy agents

Cancer cells resist to chemotherapies by cellular mechanisms such as escape to drug-induced apoptosis but also non-cellular mechanisms, e.g. physiological environment [4-7]. They are summarized below and in Figure 2.

1. Drug entry: a drug (D) penetrates into cells by passive diffusion, endocytosis or importers. This can be reduced by either decreasing uptake (e.g., OATP: organic anion-transporting polypeptide; OCT: organic cation transporter) or by alterations in lipid metabolism or by increasing drug efflux pumps (such as the ABC transporters family). Among the ABC transporters' family the P-glycoprotein (P-gp, ABCB1) is the most involved in this MDR

phenotype due to its detoxification function. Others have also been described, such as the breast cancer resistance protein (BCRP, ABCG2) and the multidrug resistance-associated proteins (MRP1, 2, 4) although each of them play a specific physiological role. This family and their most relevant members will be introduced in detail later.

2. Drug-metabolizing enzymes (DMEs): these enzymes are involve in the drugs elimination process. Such metabolization involves three phases: modification, conjugation, and excretion. In the phase I, a variety of enzymes introduce reactive and polar groups into the drug. This is mediated mainly by cytochrome P450 (CYPs) and flavin monooxygenase (FMO) enzymes. In the phase II these metabolites are conjugated with charged species such as glutathione S-transferases (GSTs), UDP-glucuronosyltransferases (UGTs), sulfotransferases (SULTs) and N-acetyltransferases (NATs). In many resistant tumor cell lines, the GST system is overexpressed for protecting the cell from chemotherapeutic agents. Finally, in Phase III multidrug resistance-associated proteins (MRPs) expel the conjugated metabolites out of the cell.

3. Drug sequestration: some types of drugs can be trapped in subcellular organelles (e.g., lysosomes and endosomes) and then excluded from the cell by non-ABC efflux transporters (e.g., ATP7A/BRLIP76).

4. Mechanisms activated after nuclear entry: usually prodrugs escape the above mechanisms but once they have reached the nucleus they can be handled by Major Vault Proteins and then either sequestered in vesicles or translocated by ABC pumps. If the drug stays into the nucleus, mechanisms are triggered to counteract its effect.

5. Escape to drug-induced apoptosis: cancer cells have many different ways to change the apoptotic signaling pathways, either by reducing the level of proteins involved in apoptosis (e.g., MAPK: mitogen activated protein kinase; P53, etc.), or by increasing anti-apoptotic proteins (Bcl-2/Bcl-x_L). Increased expression levels of bcl-2 have been associated with cellular resistance towards a number of cytotoxic agents including doxorubicin, taxol and mitoxantrone.

6. Micro-environment: solid tumors generally have inadequate vascularization, leading to a lack of nutrition and hypoxia. This creates an environment of unfavorable physicochemical conditions for many chemotherapeutic drugs. For example, the effectiveness for some prodrugs that requires oxidation to become cytotoxic is reduce in this condition. Further, cancer cell under this condition activate the HIF-1 transcription factor which up-regulates the expression of numerous MDR-linked genes (ABC transporters, Bcl2 family genes, glutathione, metallothionein).

7. Signal transduction pathways: cancer cells have signaling cascades in which integrin receptors, growth factor receptors, frizzled receptors and smoothened-patched receptors have a positive or negative feedback, triggering the expression of proteins involved in specific cellular functions.

Although these mechanisms can occur separately, they are more often linked and synergistic [8].

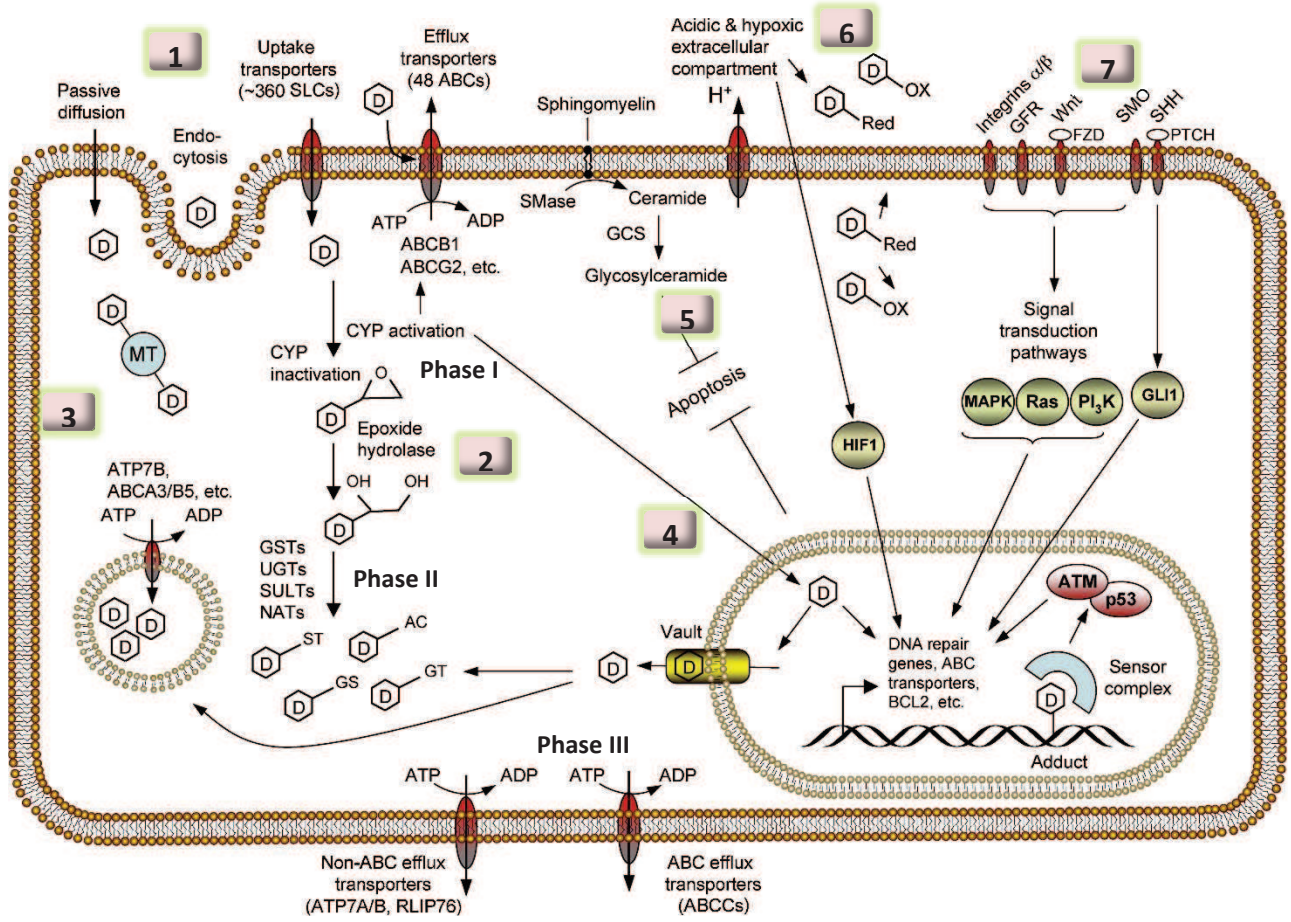


Figure 2. Drug resistance mechanisms in cancer.
SLCs: solute carriers; ABCs: ATP-binding cassette transporters; SMase: sphingomyelinase; GFR: growth factor receptor; Wnt: wingless; FZD: frizzled; Smo: smoothened; SHH: sonic hedgehog; PTCH: patched; MT: Metallothionein; GSTs: glutathione- S-transferases; UGTs: UDP-glucuronosyltransferases; SULTs: sulfo- transferases; NATs: arylamine N-acetyltransferases; GCS: glucosyl-ceramide synthase; ABCCs: ABC transporters subfamily C.
Adapted from to [7].

Bacterial resistance

Since their discovery in 1940, antibiotics are widely used in medicine and agriculture. Unfortunately, such a broad use leads to an increasing resistance of microorganisms. At the beginning this was not alarming because pharmaceutical companies produced new antibiotics, however the rate of discovery became lower that the ability of bacteria to develop resistance. One main consequence of this situation is the emergence of nosocomial infections.

Antibiotics can be generated by microorganisms (oregano oil), or produced from semi (ampicillin) or full (quinolones) synthesis pathways. They have the capacity to kill microorganisms without disturbing eukaryotic cells (host). Antibiotics can have a broad or narrow spectra of action;

they can be bactericidal or bacteriostatic. Based on their type of pharmacological effect they are divided into five groups, as illustrated in Figure 3:

1. Cell wall synthesis inhibitors: cell wall is crucial for the life and survival of bacterial species. It is made of peptidoglycans and lipids. Several classes of antibiotics target enzymes involved in its biosynthesis, such as transpeptidases, and transglycosylases.

2. Cytoplasmic membrane synthesis inhibitors: plasma membrane is the physiological barrier tightly segregating and regulating the flow of substances. As its damage leads to death, the use of such antibiotics (e.g., polymyxins) is restricted to dermal application because their broad toxicity for the mammalian host.

3. Inhibition of protein biosynthesis: several types of antibacterial agents alter the protein biosynthesis in the ribosome.

4. Inhibition of nucleic acids biosynthesis: antibiotics which interfere with their synthesis compromise bacterial multiplication and survival. Some antibiotics such as the metronidazole target topoisomerases, other like quinolones block the DNA gyrase while the rifampicin blocks the RNA polymerase.

5. Antimetabolite activities: these antibiotics mimics a natural substrate, inhibiting the enzyme target once bound. For example, sulfonamides bind to dihydropteroate synthase (DHP) and trimethoprim inhibits dihydrofolate reductase (DHF), interrupting the synthesis of folic acid (vitamin synthesized by bacteria, but not humans).

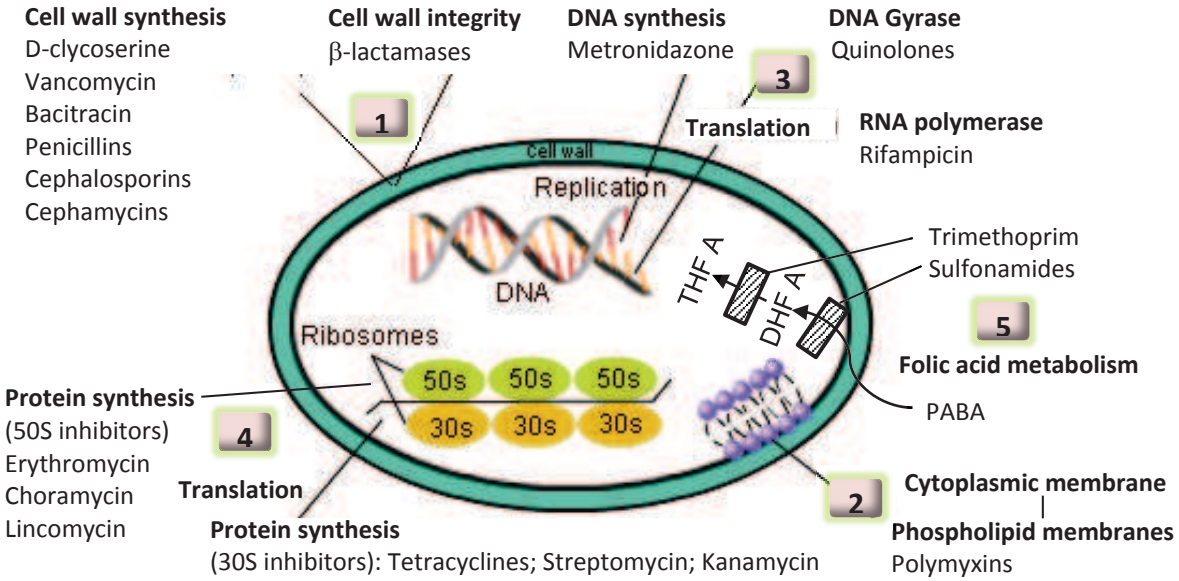


Figure 3. Mechanisms of action of antibiotics.
PABA: p-aminobenzoic acid; DHF A: dihydrofolate acid; THF A: tetrahydrofolic acid.
Adapted from the website: http://chemistry.tutorvista.com/biochemistry/antibiotics.html

And bacteria evade the activity of antibiotics

In his 1945 Nobel Prize lecture, Alexander Fleming reminded the audience to the dangers of an inappropriate use of penicillin: "*… I would like to sound one note of warning. Penicillin is to all intents and purposes non-poisonous so there is no need to worry about giving an overdose and poisoning the patient. There may be a danger, though, in underdosage. It is not difficult to make microbes resistant to penicillin in the laboratory by exposing them to concentrations not sufficient to kill them, and the same thing has occasionally happened in the body* [c]"

Factors that drive the antibiotic resistance can either take place prior to the drug treatment (natural resistance) or may develop following the exposure to a drug (acquired resistance). While natural resistance is stable, affects all cells of all strains, is transmitted to the offspring and depends on its biochemical constitution, the acquired resistance is less stable, involves only a few strains of a given species and depends not only of its biochemical constitution but also of mutations in DNA.

The various mechanisms of resistance illustrated in Figure 4 can be due to:

1. Reducing the permeability of the bacterial cell wall: the bacteria such as *Pseudomonas aeruginosa* prevent the access of antibiotics to intracellular targets by mutations affecting porins (proteins involved in antibiotic uptake) channels in the cell membrane.

2. Drug inactivation: antibiotics can be inactivated by enzymatic cleavage. For example, penicillin and cephalosporins can be inactivated by β-lactamases which destroy the β-lactam ring (three carbons and one nitrogen) through hydrolysis.

3. Target mutations: one of the most common mechanisms of resistance is the mutation of the target enzyme which prevents the binding of the antibiotic. For example, the ribosome of staphylococci can become insensitive to erythromycin following specific enzymatic modifications of rRNA.

4. Failure to metabolize drugs: some drugs need to be converted by the host into an active form. This is the case for *Bacteroides fragilis* who does not metabolise the nitroimidazole derivative metronidazole to the active metabolite and are therefore resistant to this drug.

5. Biosynthesis or increasing the level of transporters proteins: when compounds cannot be obtained by diffusion (e.g., sugars, vitamins, metal ions, etc.), bacteria depend on ABC importers which transport them into the cell. However, bacteria also have active transport systems (efflux pumps as AcrAB pump in *S. typhimurium*) for removing antibiotics (such as tetracyclines, macrolides, and quinolones) out the cell.

[c] Available from the website: http://www.nobelprize.org/nobel_prizes/medicine/laureates/1945/fleming-lecture.pdf.

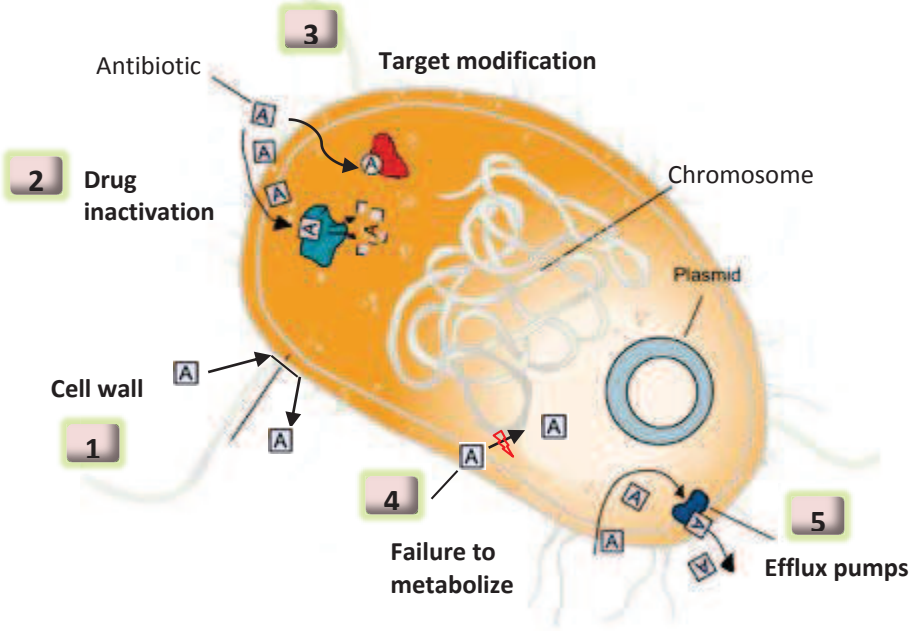


Figure 4. Common mechanisms of antibiotics resistance.
Adapted from the website:
http://www.wiley.com/college/pratt/0471393878/student/activities/bacterial_drug_resistance/

Human immunodeficiency virus (HIV) resistance

Acquired immunodeficiency syndrome (AIDS) is a pandemic having killed about 25 million people between 1981, the year of the first cases were reported, and January 2006[d]. The human immunodeficiency virus (HIV) responsible for this disease was isolate in 1983 at the Pasteur Institute [9] and few months later in U.S. [10]. Luc Montagnier and Françoise Barré-Sinoussi were awarded for their discovery with the Nobel Price in Physiology or Medicine in 2008, twenty-five years later after their discovery. Gallo and co-workers are recognized to be first to have demonstrated that HIV causes AIDS.

Description of HIV

HIV is a lentivirus belonging to the family of retroviruses (implying a slow progression of the disease). It uses the cellular machinery of the immune cells (lymphocytes T4, macrophages, monocytes ...) to replicate, which leads to the destruction of CD4+ cells. Below to 200 CD4+ cells/mm^3 in the blood (for a non-infected person this number is normally between 500 and 1200), humans are attacked by various opportunistic infections which lead to death [11].

Up to now, the HIV-1 and HIV-2 species have been identified. They invaded the human population by multiple cross-species transmissions from Simian Immunodeficiency Virus (SIV) infected non-human primates. HIV-1 is believed to have evolved from SIVcpz, a SIV-variant present in chimpanzees while HIV-2 originates from a variant (SIVsmm) present in sooty mangabeys. Both HIV-1 and HIV-2 are very similar in terms of transmission routes and disease symptoms. However,

[d] Global Report on the AIDS epidemic 2006 (2006 Report on the global AIDS epidemic).

HIV-1 is more contagious and most predominant in the world, while the HIV-2 takes a longer time to develop AIDS and is concentrated in West Africa [9, 12, 13].

Genes and structure of HIV-1

The HIV-1 particle is spherical with a diameter of around 100-120 nm. It contains two copies of single stranded RNA as well as the viral enzymes protease, reverse transcriptase and integrase. These proteins are tightly bound to nucleocapsid protein p7 and the core is protected by the viral p24 (CA) protein. The viral cone itself is surrounded by a spherical matrix comprised of p17 (matrix, MA) proteins which are enclosed by the viral envelope, a phospholipid bilayer. A schematic illustration of the viral structure is shown in Figure 5.

The single stranded RNA encodes for 15 different proteins, some by overlapping reading frames. They are synthesized from nine genes: three genes that encode for polyprotein precursors (Gag, Pol, Env) and six accessory genes with regulatory and auxiliary functions: Vif, Vpr, Vpu, Tat, Rev, and Nef. The genomic structure of HIV-1 is depicted in appendix Table 2.

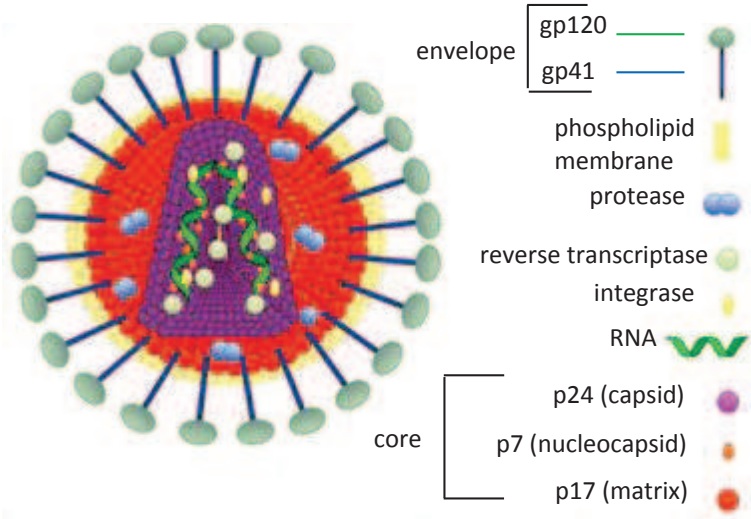


Figure 5. Schematic representation of an HIV-1 particle.

HIV-1 replication cycle

The HIV life-cycle goes through multiple steps to reproduce itself and create many more virus particles. This process can be divided into the following steps [14] illustrated in Figure 7:

1. Entry: this begins with the binding between the glycoprotein gp120 of the viral envelope and the CD4 receptor on the host cell. However, the gp120-CD4 complex is not enough and required to interact with a co-receptor of the cell surface, such as chemokine receptors CCR5, R5-tropic strains, or CXCR4, X4-tropic strains. The fusion in the membrane occurs through the "hairpin" domains of the glycoprotein gp41 who pulls the virus and host cell membranes together (Figure 6).

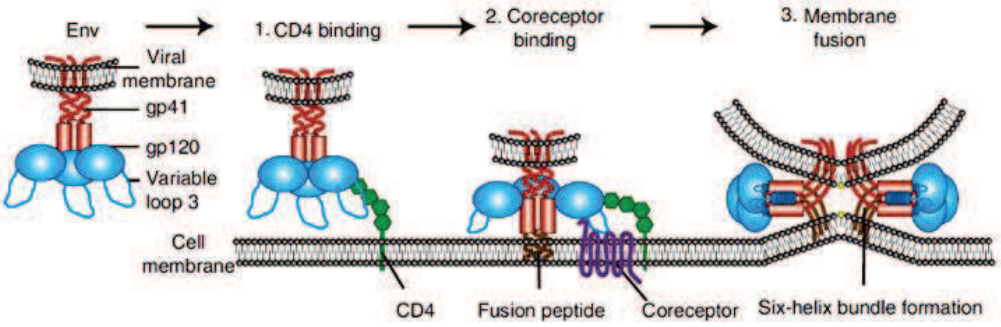


Figure 6. Scheme of HIV-1 entry.
According to the website: http://www.sinobiological.com/Basic-HIV-Human-Immunodeficiency-Virus-Virology-a-6386.html

2. Uncoating: after fusion of the viral envelope with the plasma membrane of the target cell, the capsid core is disrupted releasing the RNA and enzymes in the cytoplasm of the target cell.

3. Reverse transcription: in the cytoplasm of the cell, the viral RNA is reversely transcribed into a DNA-RNA hybrid by the viral reverse transcriptase (RT). Degradation of RNA by RNase H allows the release of single-stranded DNA is then synthesized double-stranded DNA. The reverse transcriptase is constantly detached and re-attached to the viral DNA, which is a source of errors, creating numbers genomic and phenotypic variants of the virus.

4. Integration: the viral DNA moves into the nucleus where it is integrated into the host chromosomal DNA by the integrase enzyme. This integration is an irreversible process (marking an important point of HIV infection) and occurs randomly in the cellular chromosome. At this stage, the virus is known as a provirus and may remain latent for hours to years before becoming active through transcription (copying of DNA into RNA).

5. Transcription and translation: when the host cell becomes activated, the transcription of the proviral DNA into RNA begin, controlled by a number of proteins including Tat and cellular DNA transcription factors. The unspliced, partially spliced, and fully spliced versions of viral RNA may be transported out of the nucleus and translated.

6. Assembly, budding and proteolytic maturation: the new produced proteins together with viral RNA are assembled at the plasma membrane (known as an immature virion) and the new viral particles bud off the host cell. Before the virus becomes infectious, the viral protease enzyme cleaves the Gag and Gag-Pol polyprotein into functional structural and enzymatic components (PR: protease; RT: reverse transcriptase; and IN: integrase).

Anti-HIV therapy

In the absence of vaccine (none of those which have been tested were successful, e.g., AIDSVAX vaccine candidate), antiretroviral therapy is the only way to fight against HIV. If it does not allow the complete eradication of this virus in the body, it still reduces mortality and morbidity through prevention and restoration of immune efficiency. This drugs fighting the virus at different stages of the replication cycle of HIV to prevent infection of new cells and the production of virions. Thus, they can be classified into several groups according to the steps in the viral replication cycle with which they aim to interfere: there are protease inhibitors, reverse transcriptase inhibitors, integrase inhibitors, and entry inhibitors (Figure 7 and appendix Table 3).

Up to now there are 25 antiretroviral drugs have been approved for treating HIV by the US Food and Drug Administration (FDA) and the European Medicines Agency (EMEA). As HIV mutates quickly for resisting to drugs, drugs are absorbed in combination with 2 or more agents belonging to one or more different classes of antiretroviral. This therapy is called highly active anti-retroviral therapy (HAART). Today, the life time for HIV positive patient receiving this therapy exceeds many years compared to one year in 1987 with only zidovudine (AZT) [15].

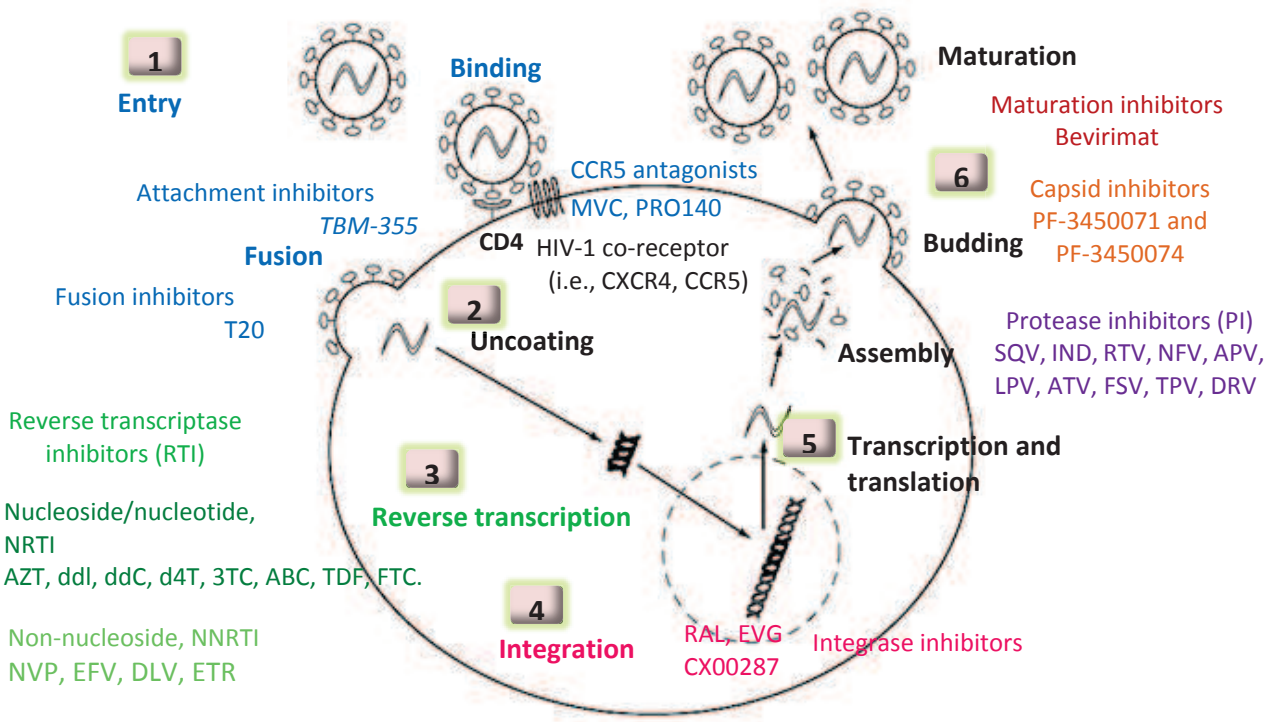


Figure 7. HIV life-cycle and mechanisms of action of antiretroviral drugs.

T20: enfuvirtide; TMB-355: ibalizumab; MVC: maraviroc; PRO140
AZT: zidovudine; ddI: didanoside; ddC: zalcitabine; d4T: stavudine; 3TC: lamivudine; ABC: abacavir; TDF: tenofovir; FTC: emcitrabine
NVP: nevirapine; EFV: efavirenz; DLV: delavirdine; ETR: etravirine
RAL: raltegravir; EVG: elvitegravir; CX00287
SQV: saquinavir; RTV: ritonavir; IND: indinavir; NFV: nelfinavir; APV: amprenavir; LPV: lopinavir; ATV: atazanavir; FSV: fosamprenavir; TPV: tipranavir; DRV: darunavir

Adapted from [16, 17]

Although success of HAART, human immunodeficiency virus still persist

In the absence of a treatment able to fully eradicate the virus, a HIV positive people has to be treated forever. This represents several risks such as poor absorption, toxic effects, drug-drug interactions and drug resistance, leading to treatment failure and disease progression. In general, the first failure is the virology (viral load doesn't drop), followed by the immunologic (CD4 cell doesn't rise or it drops) and then clinical progression (the patient have symptoms of HIV disease despite taking the treatment). There are three manly factors associated to a treatment failure [18]:

1. HIV related factors: drug resistance usually originates from genomic variation due to misincorporated nucleotides during reverse transcription, making new strains of HIV and anti-retroviral agents ineffective by altering the ability of the drugs to bind the target. The appendix Table 4, shows some of these mutations in the active site or outside of HIV that are targeted by antiretroviral drugs (e.g., reverse transcriptase, integrase, and protease). These mutations were identified by *in vitro* experiments (e.g., site-directed mutagenesis), tests of clinical isolates (e.g., susceptibility and sequencing of viruses from patients in whom the drug is failing) and correlation studies between genotype at baseline and virology response in patients exposed to the drug [19].

   A main obstacle to the complete eradication of HIV is the presence of viral reservoirs that can be considered at the cellular and tissue levels. At the cellular level the virus has the capacity to infect cells latently; in the memory cells as CD4 cell, in long half-life cell as dendritic cells, etc. At the tissue level, the virus can also penetrate through anatomical sites isolated from the rest of the body, e.g., in the brain, microglia cell may be infected with HIV. These reservoirs lead the possibility of new resistant variants.

2. Host related factors: patients with risk factors such as obesity, abnormal liver or renal functions are more likely to develop adverse drug reactions (ADRs). Some pre and co-existing diseases may also increase drugs interactions and toxicity.

   Patients that do not follow the treatment regimen as prescribed in terms of dose, time and with or without food, can experience an increase in viral load and drug resistance. In order for the regime to be effective the patient need be up to 95 % adherent [20].

   Other patients like children's and pregnant women challenges the antiretroviral therapy because they require dose modification and the number of approval of appropriate pediatric drug is limited.

3. Anti-retroviral (ART) related factors: ART may have,

   - Adverse effects and toxicity: occurs when drugs affect the body in ways other than those intended. This may cause poor adherence in the patient. For example, Ritonavir cause abdominal pain; Abacavir may cause headache, nausea, vomiting, diarrhea, etc. In

general, patients receiving ART have a high risk of cardiovascular and cerebrovascular diseases.

- Food restrictions: necessary for some drugs to be absorbed properly. For example, Lopinavir/Ritonavir should be taken with a meal or light snack [21].
- Complex pharmacokinetics properties: influencing drug disposition at several levels as absorption, distribution, metabolization and elimination of the drug. Two family of proteins are involve in the disposition of antiretroviral drugs; the ABC efflux transporters and solute carrier (SLC) uptake transporters [16, 17].
- Drug interactions: interactions between drugs can alter the effectiveness of antiretroviral therapy. For example, Garlic capsules (natural treatment for heart health and circulation) stop Saquinavir from working properly. The combination of Saquinavir with ritonavir may cause an abnormal heart rhythm. This interactions are linked to changes in the activity of cytochrome P450 (CYP) enzymes. However, ABC and SCL transporters can contribute [16, 17].

A common mechanism: drug efflux pumps

"All living cells contain genes that code for efflux pumps…"
[22]

As described above, all mechanisms of resistance highlight the implication of efflux pumps to decrease the effective concentration of drugs in the intended target, leading to multidrug-resistance (MDR). Several membrane transporters are implicated in MDR phenomenon. These membrane proteins play an important role in all organisms, working either as importers, bringing nutrients and other molecules in the cell (e.g., amino acids, sugars, and essential metals), or as exporters, actively transporting a wide range of molecules across the lipid membrane, such as biologicals substances (e.g., lipids, peptides and hormones) exogenous toxic substances (heavy metals, contaminants ...) and also a wide range of therapeutic drugs [23].

The efflux pumps-mediated resistance in bacteria

According to the number of proteins components that the pump has, single or multiple (usually Gram-negative bacteria efflux pumps have several components), and the type of energy which drives them, the efflux pumps-mediated resistance in bacteria may be classified in five families as follow [24, 25]:

The major facilitator superfamily (MFS), the resistance nodulation division (RND) family and the small multidrug resistance (SMR) family are driven by a proton gradient. The MSF consists of membrane transport proteins which are found from bacteria to higher eukaryotes and are involved in the symport, antiport, or uniport of various substrates, such as sugars, Krebs cycle intermediates,

phosphate esters, oligosaccharides, and antibiotics. Some examples are the proteins EmrD of *E. coli*; NorA and QacA of *S. aureus*. The SMR proteins are the smallest secondary drug transporters, which usually exists as a part of a tripartite complex. Some of the well characterized examples of this family include Mmr of *Mycobacterium tuberculosis*, QacE of *Klebsiella aerogenes* and EmrE of *E. coli*. The RND proteins interact with a membrane fusion protein and an outer membrane protein to allow drug transport across both the inner and outer membranes of Gram-negative bacteria. AcrB of *E. coli*, MexB of *P. aeruginosa* and MtrD of *Neisseria gonorrhoeae* are members of this transporter family.

The multidrug and toxic-compound extrusion (MATE) family is driven by a proton gradient and is similar to the MFS family with members as NorM of *Vibrio parahaemolyticus* and YdhE of *E. coli*.

The ATP-binding cassette (ABC) superfamily is driven by ATP hydrolysis and has similarities with the human ABC transporters, for example, LmrA from *Lactococcus lactis* is homolog of the human multidrug transporter MDR1 [26].

The following figure and table show some efflux pumps from each family.

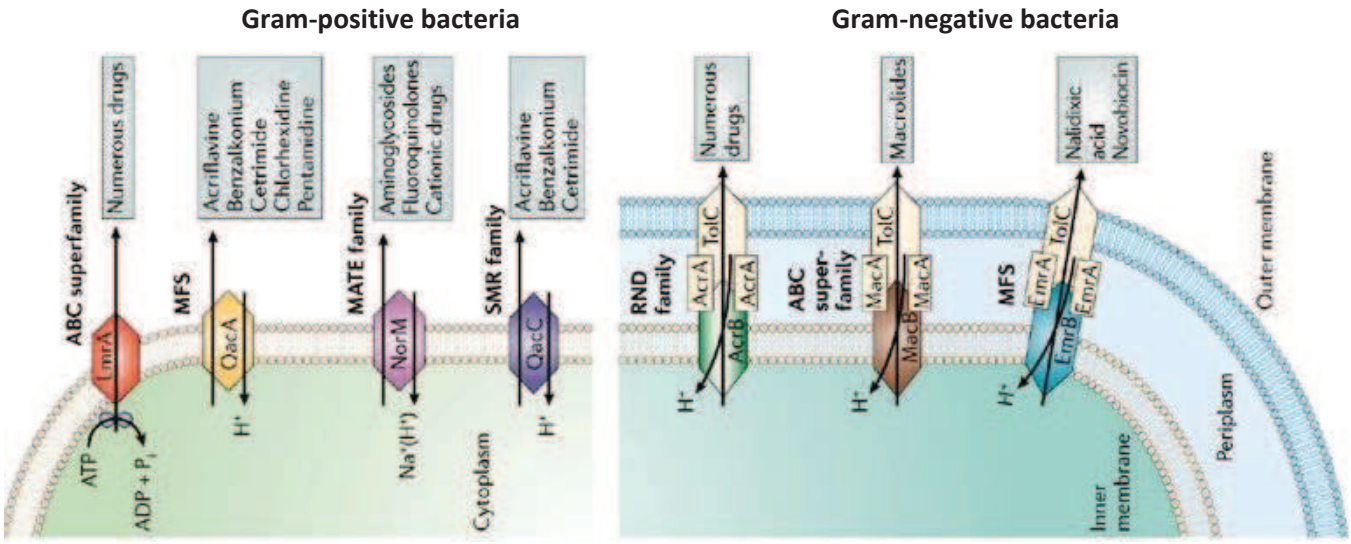


Figure 8. Multidrug-resistance efflux pumps in bacteria.
According to [24]

Organism	family	efflux pump
S.aureus	ABC	MsrA
	MFS	MdeA, NorA, NorB, Tet K-L, Tet38
S.pneumoniae	ABC	PatA and PatB
	MFS	MefA, MefE, PmrA, Tet K-L
H. influenza	MATE	HmrM
	MFS	TetB,K
	RND	AcrAB-TolC

E. coli	ABC	MacAB-TolC
	MATE	YdhE
	MFS	Bcr, Dep, YdhE Bcr Dep, ErmAB-TolC, Fsr, MdfA, SetA, Tet A-E Ycel, Ycel, YidY, YebQ
	RND	AcrAB-TolC, AcrAD-TolC, AcrEF-TolC, YegN
	SMR	ErmE
P. aeruginosa	MFS	Tet A, C, E; CmlA
	RND	MexAB-OprM, MexCD-OprJ, MexEF-OprN, MexJK-OprM, MexXY-OprM
E. aerogenes	MFS	CmlB
	RND	AcrAB-TolC, EefABC

Table 1. Major efflux pumps involved in human pathogens.
Adapted from [27]

The pumps-mediated resistance in human

In human, the ABC and solute carrier (SLC) superfamilies are implicated in drug disposition, drug–drug interactions, and variability in drug response and toxicity [28]. The ABC transporters's family has about 49 members mostly for maintaining low intracellular compounds concentration (efflux transporters) using energy derived from ATP hydrolysis to transport xenobiotics and toxic endogenous substrates. The second family includes much more members, about 300, mostly for maintaining high intracellular concentration (uptake transporters) creating gradients to co-transport and/or exchange of ions [29]. Both are expressed in important pharmacological barriers and in tissues involved in drug absorption, distribution, and elimination (see Figure 9 and appendix Figure 1). Thus, their ubiquitous expression and diverse affinity brings about many drug-to-drug interactions and pharmacokinetic alterations (reviewed in [30]).

The interplay between efflux and uptake of drugs at areas of bioavailability (intestine, liver, kidney, blood-brain barriers, etc...) makes them important in drug development. It has been reported that genetic polymorphisms in SLC family results in reduced accumulation of certain drugs. And for most ABC transporters single nucleotide polymorphisms (SNPs) may significantly influence their activity, substrate recognition, and regulation [31].

In the present manuscript, I have focused on the ABC transporters family.

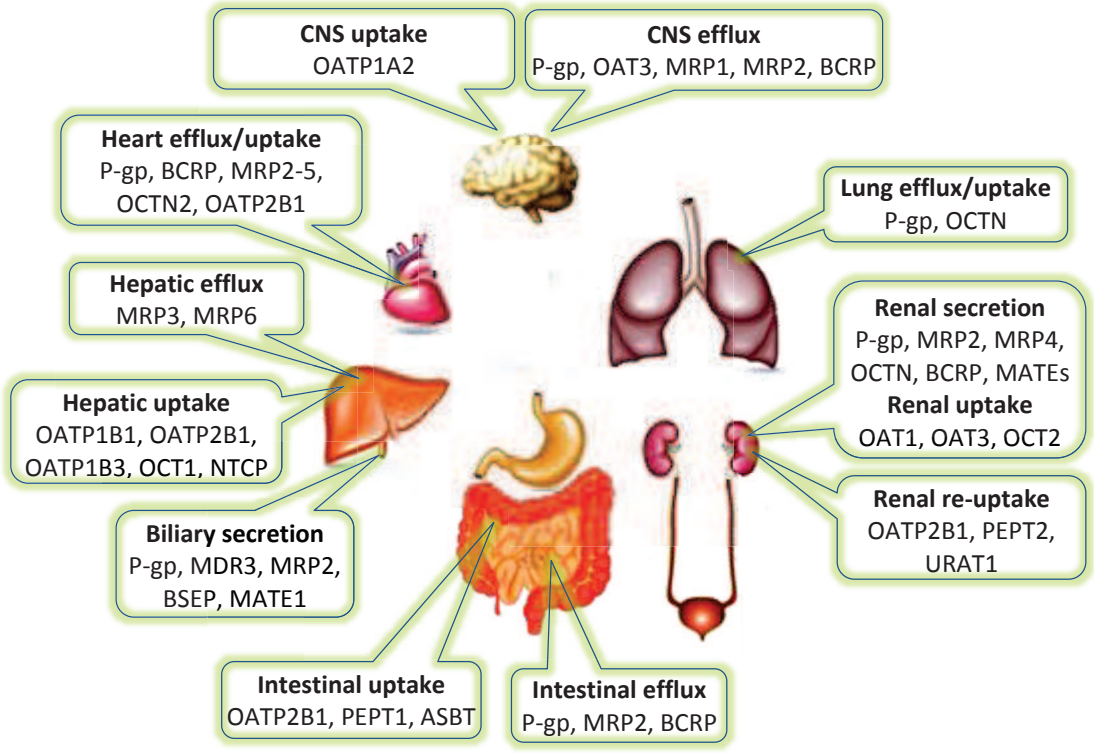


Figure 9. Non-exhaustive distribution of membrane proteins involved in drugs disposition.
P-gp: P-glycoprotein; BCRP: breast cancer resistance protein; MRP: multidrug resistance-associated protein; OCT: organic cation transporter; OCTN: organic cation/l-carnitine transporter; OAT: organic anion transporter; OATP: organic anion transporting polypeptide; PEPT: peptide transporter; ASBT: apical sodium-dependent bile transporter; MATE: multidrug and toxin extrusion protein; BSEP: bile salt export pump; URAT1: uric acid transporter 1
Adapted from [29]

And if ABC transporters are more than just drug efflux pumps?

In cancer, the question was underlined due to failure of clinical trials of P-gp inhibitors despite the clear *in vitro* and *in vivo* evidences that drug accumulation is increased by these molecules [32]. Many reasons may explain this paradox such as: low potency of P-gp inhibitors, not specific for P-gp, additional toxicities (due to alterations to the pharmacokinetics), wrong patient population (e.g., patient that didn't express P-gp) and possible SNPs of transporters. Nevertheless P-gp and other ABC transporters play a central role in MDR, several studies show that the likelihood of treatment failure is correlated with the degree of expression of ABC transporters. This is the case of leukemia, neuroblastoma, sarcomas and breast cancer [33]. Their expression in cancer cell not only is correlated with poor prognosis but has been shown to increase after chemotherapy [34] (detailed later). It has been suggested that P-gp and MRP1 could have a role in the stage of promotion and progression of cancer, as indicated by their expression levels in lymph node metastases of breast cancer patients [35]. Additionally, P-gp may regulate cell differentiation, proliferation, immune responses and programmed cell death [36]. Fletcher *et al.*, suggest that ABC transporters may contribute to cancer hallmarks [37] (self-sufficiency in growth signals, insensitivity to anti-growth signals, evasion of apoptosis, limitless replicative potential, sustained angiogenesis,

and tissue invasion and metastasis [38]); Bates and colleagues that P-gp may be a marker or surrogate of other more important resistance mechanisms [39].

In bacteria, other actions of these efflux pumps have been described. RND and MATE families are implicated in pathogenicity of bacteria. They may transport virulence determinants (adhesins, toxins or other proteins), allowing the bacterium to survive, to colonize and to infect human and animal cells. As examples, the E. *amylovora* MATE-family efflux pump NorM facilities the colonization of bacteria to the host plant [40]. The N. *gonorrhea* RND-family efflux pump MtrCDE is important for infection of the genito-urinary tract of female mice [41]. Piddock reported that when the gene coding for AcrB is disrupted, the S. *typhimurium* displays a reduced efficiency for adhesion, and host invasion in mouse macrophages or human embryonic intestinal cells [24]. Even more, some proteins belonging to ABC family are implicated in other processes. This is the case of the bacterial uvrA protein which is a component of nucleotide excision repair (NER), one of the primary pathways for removal of DNA damage [42].

For HIV, it has been reported that P-gp is expressed in lymphocytes which may affect the intracellular concentrations of ART drugs [43]. However, the relation between ABC transporters and HIV infection is even more complicated because most of the results are contradictory, possibly due to the selection of patients, cell types, or differences in the protocol of analysis. For example, some authors indicate a decrease in expression of P-gp on CD4+ lymphocytes in infected patients [44-47]; others show the opposite in T-lymphocytes and monocytes [48], and even an increase with disease progression [49, 50]. It appears that the regulation of P-gp expression is accompanied by a decrease in its activity in infected patients [44, 51]. Furthermore, the MRP1 expression in lymphocytes does not seem affected by the infection [45]. Besides, expression of these transporters in HIV infected patients seems to influence the infectivity and replication of HIV. While P-gp can inhibits the fusion of the virus with the host cell retarding the viral replication [52, 53], MRP1 seems to increase the viral infectiousness [53].

In conclusion the ABC transporters not only influence drug disposition and effectiveness in target cell, but also play a role more than just efflux pumps. In addition to their involvement in diverse cellular processes including, cell division, bacterial immunity, maintenance of osmotic homeostasis, drug resistance, antigen processing, pathogenesis and sporulation, cholesterol and lipid trafficking, and developmental stem cell biology [54]. Thus to counter MDR through efflux pumps, it is important to identify the main contributor(s) of this phenomenon.

ABC TRANSPORTERS

Dano *et al.*, reported for the first time in 1973 that drug resistance is mediated by an energy-dependent outward transport. Three years later the P-glycoprotein was the first human drug transporter identified in a pleiotropic resistance phenomenon [55]. The term "ABC transporter" was introduced later, in 1985, based on the presence of a highly conserved domain called ATP-Binding-Cassette that characterizes this superfamily [56-58]. The number of studies of these proteins have then increased significantly owing to their involvement in genetic diseases and MDR phenotype.

ABC transporters are present in organisms from all kingdoms of life from prokaryotes to mammals (Table 2). Some members of this superfamily are not involved in transport-related processes such as DNA repair (Rad50), translation (EF3) or regulatory functions (ABCE1, GCN20) [59-63]. The broad majority of them utilizes energy gained from ATP hydrolysis to transport endogenous or exogenous molecules across all cellular and organelle membranes (cell surface, mitochondrial, peroxisome, endoplasmic reticulum, Golgi apparatus, other vesicles excepting nuclear membranes which have pores) [64]. Depending on the direction of transport relative to the cytoplasm they can be categorized as importers (into the cell) or exporters (out of the cell). To date, it seems that importers are present exclusively in prokaryotic organisms [23]. However, a recent study shows that a eukaryotic ABC transporter, ABCA4, functions as importer [65]. There are 49 members of ABC transporters in humans [66], 52 in mouse, 56 in *Drosophila*, 58 in *C. elegans*, 29 in yeast [67] and about 80 in *E. coli* representing 5 % of their genome [23].

Yeasts (*Saccharomyces cerevisiae*)	**Bacteria**
Pdr5,10,11,12,15	LmrA *Lactococcus lactis*
Ste6	$HisQMP_2$ *Salmonella thyphimurium*
Snq2	$MalFGK_2$ *Salmonella thyphimurium and Escherichia coli*
Yor1	HlyB *Escherichia coli*
	OppF *Salmonella thyphimurium*
Parasites	**Vegetal**
Pfmdr *Plasmodium falciparum*	AtPGP1 *Arabidopsis thaliana*
LeMDR1 *Leishmania enriettii*	PGP1 *Arabidopsis thaliana*
LaMDR1 *Leishmania amazonensis*	AtMRP *Arabidopsis thaliana*
	NpABC1 *Nicotiana plumbaginifolia*
Champignons	**Other species**
CDR1,2,3 *Candida albicans*	Archaea, annelids, amoebae, sponges, paramecium ...
AfuMDR1 *Aspergillus fumigatus*	
AflMDR1 *Aspergillus flavus*	

Table 2. ABC proteins of different species. Adapted from: [68].

Overview of ABC transporters

In 2001, Dassa and Bouige carried out *in silico* phylogenetic studies of more than 600 ABC ATPases leading to the classification of ABC proteins into three major classes: class 1 is comprised of exporters with a transmembrane domain (TMD) linked to a nucleotide-binding domain (NBD). Class 2 contains all systems without TMD but with NBD, they are not transporters but are involved in various cellular functions (DNA repair, translation or regulatory functions), also known as ABC-ATPase subfamilies. Finally, class 3 contains all binding protein-dependent (BPD) importers where the TMD and NBD are borne on independent polypeptide chains and other less well characterized systems [69], two types below to this class: type I and II ABC importers.

A consolidated naming scheme for human and mouse subfamilies was developed by the HUGO organization (Human Genome Organization, http://www.genenames.org) which identified and classified seven subfamilies designated A to G followed by a number corresponding to each transporter [67, 70]:

1. *The subfamily A (ABC1)*: this subfamily has 13 members, some of the largest ABC proteins belong to this family up to 2100 amino acids. They are all full transporters (two TMDs and two NBDs expressed in a single polypeptide) and involved in lipid transport.

2. *The subfamily B (MDR/TAP):* is the most variable, it comprises four full transporters (B1, B4, B5 and B11) and seven half transporters (one TMD fused to one NBD in two or more different polypeptides; as B2, B3, B6-B10). While full transporters of this subfamily are localized in the plasma membrane (in the apical membrane compartment), the half-transporters are found in the membranes of various organelles [71]. ABCB1 (P-gp) was the first human ABC transporter cloned and characterized through its ability to confer a multidrug resistance phenotype to cancer cells [55]. Several of the B family members are known to confer MDR in cancer also (B4, B5 and B11).

3. *The subfamily C (MRP)*: only includes full transporters, but with different functions. ABCC1 is the primary transporter of glutathione and glucuronate conjugates, it is also one of the transporters responsible for efflux of cytotoxic drugs co-transported with glutathione [72]. The CFTR protein is a chloride ion channel that plays a role in all exocrine secretions. ABCC8 (SUR1) and ABCC9 (SUR2) bind sulfonylurea and regulate potassium channels involved in modulating insulin secretion. They don't have transport function identified to date. The rest of the subfamily is composed of MRP-related genes.

4. *The subfamily D (ALD)*: They are all half-transporters that work in homo and / or heterodimers to regulate the transport of fatty acids with very long chains.

5. *The subfamily E (OABP) and F (GCN20)*: These subfamilies are not known to be involved in a transport function, they only have NBDs (characteristic for ABC proteins). ABCE has only one member, ABCE1, which recognizes oligo-adenylate that is produced in response to infection by certain viruses. The ABCF subfamily includes three members and their genes appear to

be upregulated by tumour necrosis factor-α. It is believed that members of this subfamily might play a role in inflammatory processes [66].

6. *Finally, the subfamily G (WHITE)*: They are all half-transporters with a unique arrangement, the NBDs are located at the N-terminal and TMDs at the C-terminal "reverse half-transporters". Members of this subfamily function in homodimers (G1, G2 and G4) or heterodimers (G5 and G8). Since its cloning [73-75], ABCG2 has become one of the most studied ABC transporters and is considered today with ABCB1 and ABCC1, as one of the major transporters responsible for multiple drug resistance. However, several other MDR-linked ABC transporters have more recently emerged (indicated in bold the appendix Table 5).

Many human pathologies are caused by dysfunction or exacerbated function of an ABC transporter [67, 76, 77]. A majority of these diseases is recessive and due to a loss or reduction in activity, expression, folding, or trafficking of the proteins. For example, the Tangier disease is due to a gene mutation in ABCA1 leading to an orange color of the tonsils together with an enlargement of the liver and spleen and a significant reduction in the plasma concentration of HDL (high density lipoprotein). Any of dozens of mutations in the gene ABCC7 causes cystic fibrosis (also known as mucoviscidosis) which is characterized by chronic pulmonary obstruction due to a lack of fluidity of mucus secreted by the bronchial epithelium. ABCB1 and ABCB2 are critical for the proper functioning of the cellular immune response. Deficiency of ABCC2 causes an increased concentration of bilirubin glucuronosides in blood, known as Dubbin Johnson syndrome (DJS) in humans.

The appendix Table 5 describes cell or tissue expression, the main function and diseases caused by mutation in ABC genes. As it can be seen, the genes are dispersed in the human genome and their location is relatively ubiquitous in the body.

Architecture of ABC transporters

The basic structure that defines the members of this protein family is the combination of ATP-binding and transmembrane domains (NBD and TMD respectively), arranged in any possible combination. Some of them have additional transmembrane segments, extended NBDs or only have one NBD without TMD (Figure 10A). The TMD has several hydrophobic α–helices, poorly conserved in length or sequence and allows the binding of the substrates followed by a translocation across the lipid membrane (import or export). Contrarily, the hydrophilic cytoplasmic NBD contains several highly conserved motifs (Walker A and Walker B sequences, C-motif, the D-, Q-, and H-loops named according to conserved residues, and the conserved aromatic A residue: A-loop, Figure 10B) and uses the energy derived from ATP hydrolysis to transport substrates. As illustrated in the aligned sequences for each subfamily (Figure 10C) and for human and prokaryotic ABC transporters (appendix Table 6).

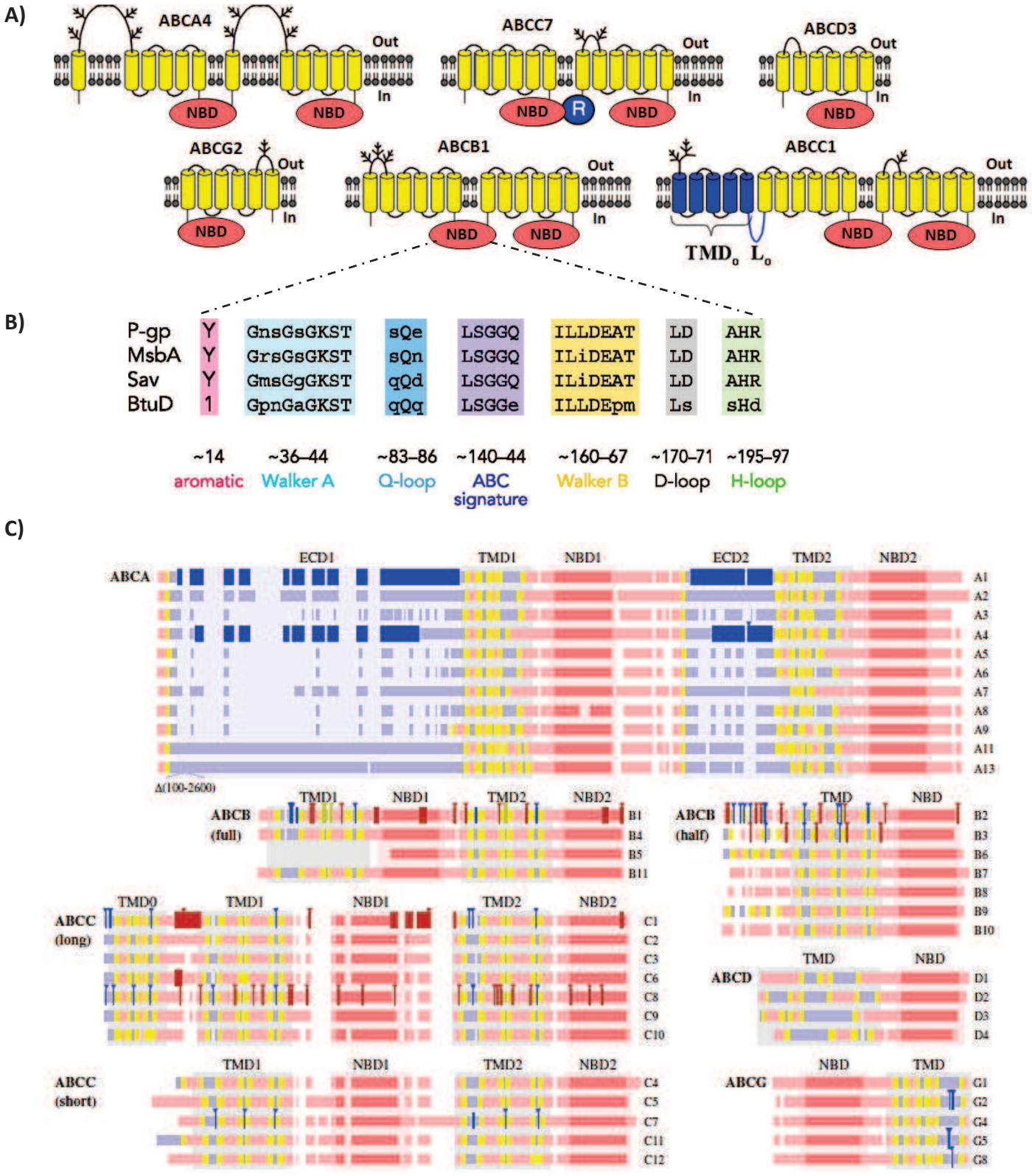


Figure 10. Predicted topologies of human ABC proteins.
A) Membrane topologies of some ABC transporters. **B)** Key motifs in the NBD. **C)** Aligned sequences for each subfamily. TMD: transmembrane domain, in yellow; ECD: extracellular domain; NBD: nucleotide binding domain, middle tone of red; R: regulator; darker thick line: experimentally established sequence regions; triangle above the sequence: experimentally established residues; light red and light blue: sequence regions predicted to be cytoplasmic or extracytosolic respectively.

According to [71, 78, 79]

Whereas full transporters are functional as monomers, half transported are organized as homo- or hetero-dimers, such as ABCG2 [80]. ABCG5 and ABCG8 function as heterodimers and do not have the ability to dimerize [81]. Higher oligomeric states have been described, potentially dimeric assemblies for full transporters such as P-gp [82-84] or MRP1 [85] and tetramers or dodecameric forms for half transporters such as BCRP [86, 87]. Considering the variety of techniques and experimental conditions (protein within natural membranes or reconstituted proteoliposomes, or lipid-detergent mixed micelles or solubilized in detergent) the interpretation and comparison of these data is very difficult.

Transmembrane domains (TMDs) and translocation

Crystal structures of full ABC transporters show than the alpha-helix forming TMDs are not simply bundles of regular α-helices, but also contain bends and kinks [88]. These structures have also revealed three main TMD folds:

i. One found in exporters Sav1866 (a bacterial ABC exporter from *Staphylococcus aureus* [89]) and MsbA (a lipid transporter of *E. coli* [90]), see Figure 11A. They are made of two membrane domains each formed by 6 alpha-helices (2X6). Two discrete extracellular loops (wings 1 and 2 in Figure 11A) are formed between TM1 and TM2 and TM3 and TM6. Intracellular loops ICL1 and ICL2 are made of two long α-helices and a short coupling helix (CH), also called Hinge [89]. They constitute an interfacial region that transmits conformational changes in response to substrate binding in the TMDs and ATP binding and hydrolysis in the NBDs.

ii. Another fold is present in the Molybdate importer ModBC [91], in the maltose importer $MalFGK_2$ [92] and in the methionine transporter MetI [93] (Figure 11B). It is made of 2x5 alpha-helices, with an arrangement distinct to that of Sav1866. These structures also reveal a closed gate near the external side of the membrane. A common gate architecture is suggested due the conservation of the primary sequences in the molybdate, sulfate and phosphate importers.

iii. A third fold is found in the vitamin B12 importer BtuCD [94], the metal-chelate-type transporter HI1471 from *Haemophilus influenza* [95] (Figure 11C). It is made of a set of 2x10 alpha-helices. Both proteins differ in the location of the gate, made of a loop between 2 TMs: it is either internal in the case of BtuC (between TM4 and TM5) or external for HI1471 (between TM5 and helix 5).

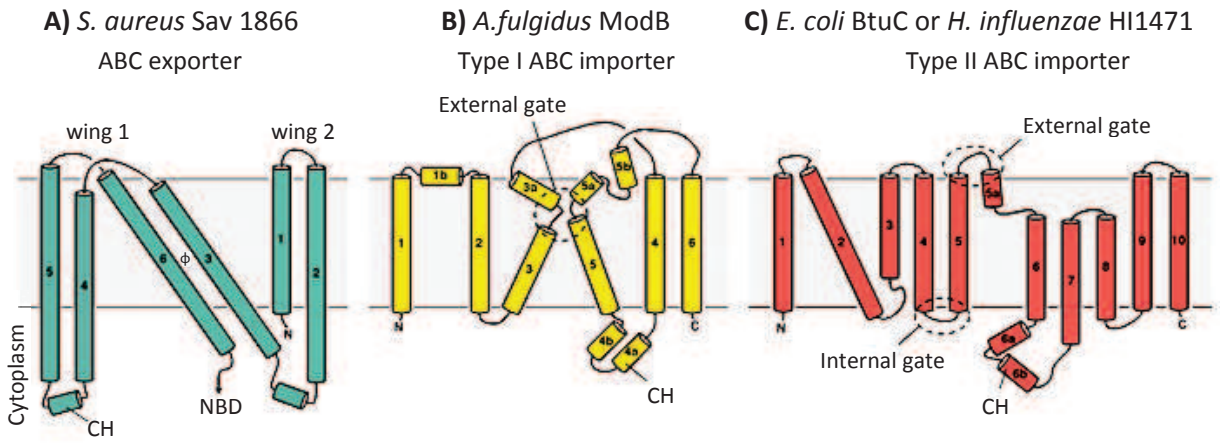


Figure 11. Topology of the transmembrane helices of a single TMD.

The TM helices are numbered consecutively, and additionally letters indicate short helices. Wings: long intracellular loops; CH: coupling helix; ϕ: the twofold symmetry relating TM1-3 to TM4-6 of Sav 1866.

According to: [96, 97]

A common structural feature to all these transporters is the presence of the coupling helices (CH). It has been proposed that these helix represent part of the transmission interface whereby binding and hydrolysis of ATP in NBDs is coupled to conformational transitions in TMDs that effects solute translocation. However, there is an important difference: exporters have four coupling helices connecting the two TMDs with the two NBDs and they are domain swapped, that is, the CH from one monomer interacts with the other monomer (or other moiety), Figure 12. While importers only have two CH and the domain swapped has not been observed, resulting in central inter-subunit gap lacking contacts between diagonally positioned subunits [96].

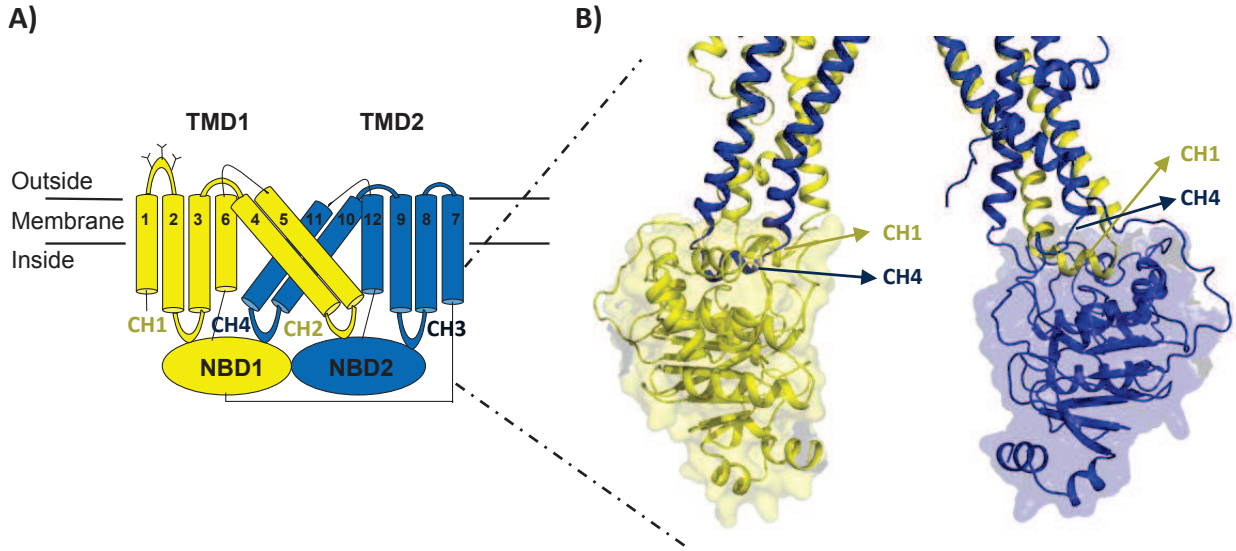


Figure 12. The coupling helices of P-gp.

A) Topological model of P-gp. B) Structure of P-gp (code PDB: 4KSC)**.** The two halves of P-gp are colored blue and yellow. The NBD is shown as a transparent solid.

CH: coupling helix; **PDB:** protein data bank

The TMDs form the substrate-binding sites, (SBS, developed later) which contribute to transport specificity. While importers require a binding protein that binds molecules with high affinity and deliver them to the TMDs (Figure 13A), the ABC exporter don't need such carrier. Molecules may: (i) approach from the cytoplasm side and be transported through a hydrophilic path formed by the transmembrane (TM) regions; this is the classical pumps model [98]; (ii) be extracted for the inner leaflet of the plasma membrane and pumped out directly to the external aqueous medium, according to the hydrophobic vacuum cleaner model [99]; or (iii) be flipped from the inner leaflet of the lipid bilayer, to either the outer leaflet of the plasma membrane or directly to the extracellular environment, the flippase model [100] (Figure 13B). It has been proposed that the physical exit of molecules is due to a change in binding affinity [101, 102].

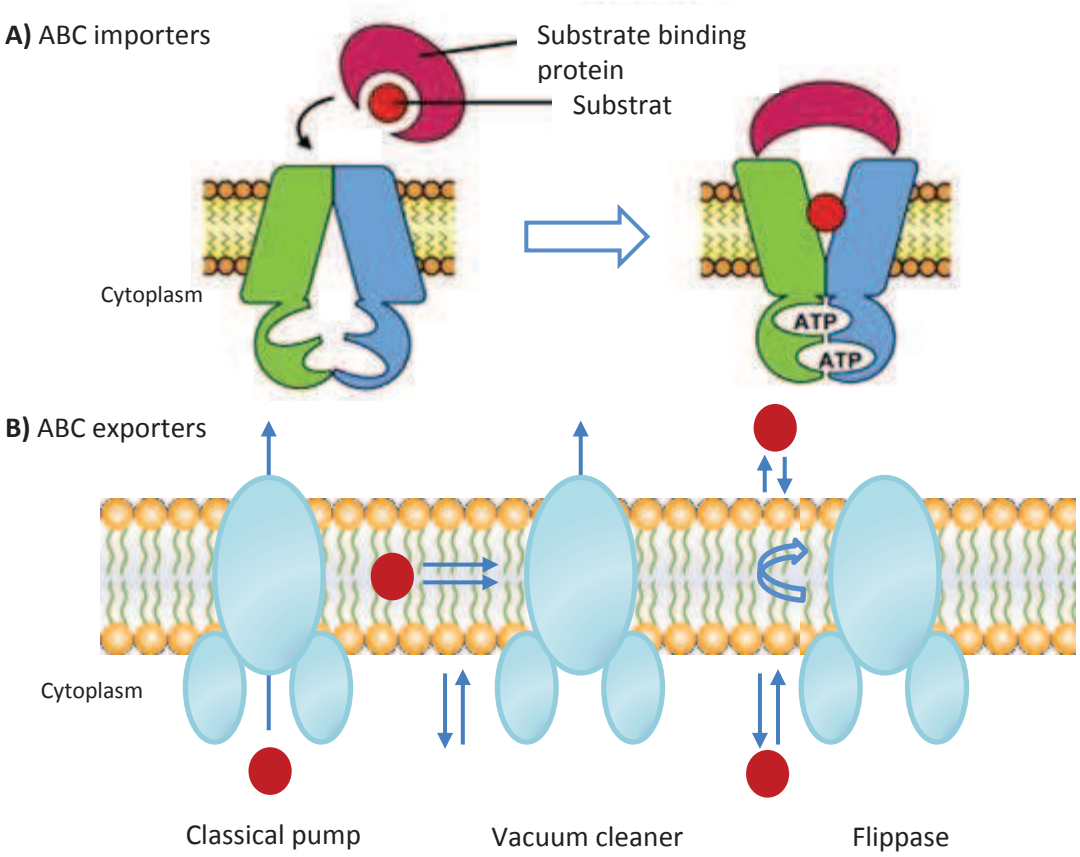


Figure 13. Substrates access in ABC transporters.
Adapted from: [103] and [104]

Nucleotide-binding domains and ATP catalysis

The NBDs of ABC transporters are always present as dimers and each monomer has roughly an L-shape with two lobes (Figure 14). The lobe I is composed of two subdomain: the ATP-binding core subdomain (similar to RecA- and F1-ATPase-like proteins) contains the Walker A and B motifs, and the β-subdomain (which binds the ribose and adenine moieties of ATP). The lobe II consists of a single α-helical subdomain which includes the ABC signature sequence or C-motif which is the hallmark of the ABC superfamily [105, 106]. Both lobes are linked by the Q-loop which also forms

part of the interface with the TMDs. A number of additional motifs and residues are involved in the ATP binding and hydrolysis processes which are briefly detailed in the following Table.

Motif	Sequence	Functions
A-loop	conserved aromatic (A) residue (usually tyrosine)	Base-stacking interaction between the aromatic ring of a tyrosine and the adenine moiety of ATP. Provides a contribution to the nucleotide-binding affinity of the NBDs
D-loop	EATSALD Has a conserved aspartate (D) residue	Involved in asymmetric inter-monomer interactions in the ATP- Mg^{2+}-bound state, thereby effecting communication between the NBD monomers
H-loop or "switch region"	conserved histidine (H) residue	The histidine interacts with the γ-phosphate group of ATP
Walker A or P-loop (glycine-rich phosphate-binding loop)	GXXGXGKS/T (where 'X' is any residue)	The lysine (K) residue binds the nucleotide through electrostatic interactions with the β- and α-phosphate of the triphosphate moiety
Q-loop	conserved glutamine (Q) residue	Links the ATP core (lobe I) and α-subdomains (lobe II). Through conformational changes may intervene in the communication between the NBD and TMD. In addition, the glutamine co-ordinates the Mg^{2+}
X-loop only conserved in ABC exporters	TXVGEXG (where 'X' is any residue)	Probably respond to ATP binding and hydrolysis and may transmit conformational changes in the intracellular loops
C-motif or ABC signature sequence	LSGGQ	Caps the N- terminus of a long helix that directs a positive charge dipole towards the γ- phosphate of ATP
Walker B	$\phi\phi\phi\phi$DE (where 'ϕ' is any aliphatic residue)	The aspartate (D) residue is hydrogen bonded to the catalytic Mg^{2+} ion

Table 3. Sequences and functions of the various conserved motifs of NBDs.
Adapted from to [89, 101, 107]

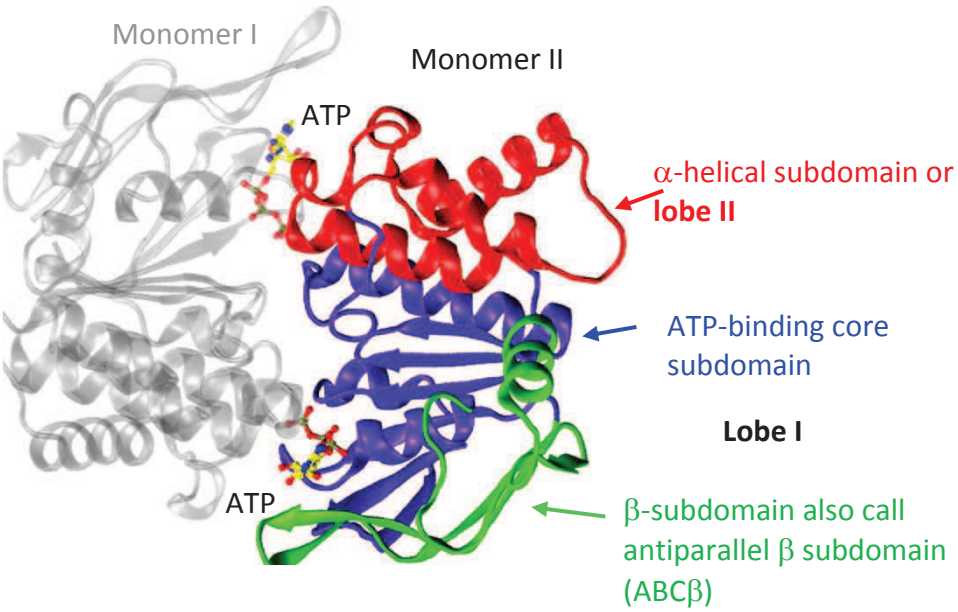


Figure 14. Three-dimensional structure of a nucleotide-binding domain.
Ribbon diagram of the MJ0796 (E171Q) ABC ATPase dimer (PBD code: 1L2T, [108]). ATP is shown in ball-and-stick form with carbon in yellow, nitrogen in blue, oxygen in red and phosphorus in tan. Adapted from [107]

It was advocated that a cooperative ATP binding causes NDB dimerization [108, 109] whereas nucleotide-free structures consistently show the NBDs physically separated (e.g. MalK [110]). Thus, NBDs are organized as dimers and two molecules of ATP are bound at their interface. The first NBD dimer was observed in HisP protein (subunit of the bacterial histidine permease complex) of *Salmonella typhimurium* with bound ATP but no Mg^{2+} and in a back to back organization (Figure 15A) [106]. Surprisingly, this structure showed an absence of interaction between the C-motifs (though to be part of the ATP-binding site) and both ATP molecules. Two years later, the MalK (together MalF/MalG domains and the BPD MalE, constitutes a functional maltose importer in Gram-negative bacteria) dimer from *Thermococcus litoralis* displayed NBDs in a head to head conformation (Figure 15B) [111], although crystallized with ADP, nucleotides could not be observed. Few months earlier, the Rad50cd 3D-structure from *Pyrococcus furiosus* complexed with AMPPNP (non-hydrolysable ATP-analog) or Mg^{2+}-ATP showed the NBD dimer in a head-to-tail conformation (Figure 15C) [112]. Among these three different arrangements, the head-to-tail is considered as the physiological conformation of most ABC transporters since found in several cases (Figure 15C). This conformation increases the stability of the nucleotide through interactions between the C-motif and the adenine ring pattern, the ribose and the triphosphate group of ATP. Thus two ATP-binding and hydrolysis sites are formed between the adjacent Walker A of one NBD and the C-motif of the other. Supplementary structures of isolated NBD dimers and in some of the complete ABC structures (Figure 15C), as well as biochemical and sequence data confirmed the relevance of the head-to-tail model [113].

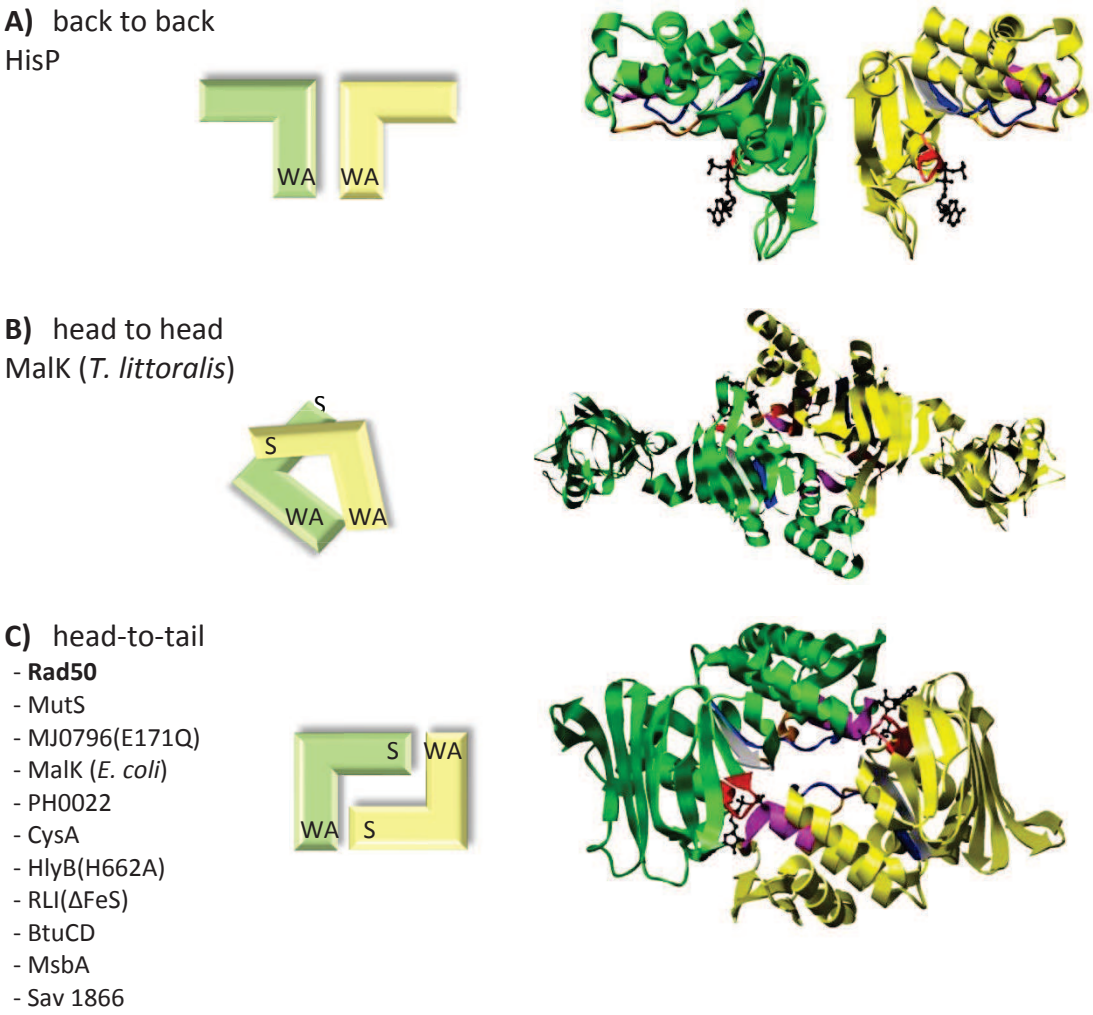


Figure 15. Schematic representation and structures of the NBD dimers.
A) Back to back model and HisP *Salmonella typhimurium* dimer structure (PBD code: 1BOU). **B)** Head to head model and MalK *Thermococcus littoralis* dimer structure (PBD code: 1G29). **C)** Head-to-tail model and Rad50 (Rad50cd) *Pyrococcus furiosus* dimer structure (PBD code: 1F2U).
S: ABC signature sequence; WA: Walker A motif.
In all cases: Walker A is in red, Q-loop in orange, ABC signature sequence in magenta, Walker B in dark blue, H-loop in gray and the ATP molecule in black.
Adapted from [114]

The dissociation of NBDs is achieved after that ATP hydrolysis breaks the interaction between the C-motif and the γ-phosphate moiety of ATP. The enzymatic hydrolysis of ATP, starts after its binding at the active site in the interface of NBDs 'sandwich' dimer. The oxygen atoms of the β- and γ-phosphates of ATP are stabilized by residues in the Walker A motif and coordinate with the cofactor ion Mg^{2+}, which also coordinates to the terminal aspartate residue of the Walker B motif through the attacking H_2O. Two main mechanisms for ATP hydrolysis have been proposed: the 'general-base catalysis' [62, 115, 116] and the 'substrate-assisted catalysis' [117-120]. In the former, the ATP hydrolysis proceeds through an acid-base reaction in which the conserved glutamate, immediately downstream of the Walker B motif, polarizes the water for a nucleophilic attack on the

ATP-phosphorus atom. While in the later, the H-loop histidine acts as the "linchpin or catalytic dyad" holding both, the glutamate and the water molecule. The composite catalytic site is depicted in the following figure.

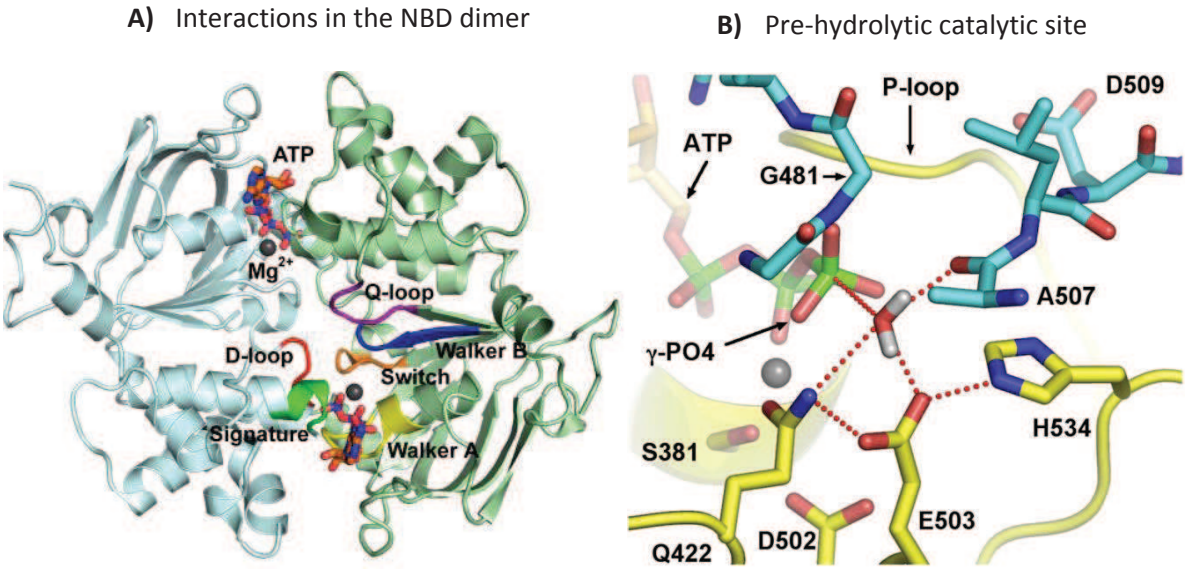


Figure 16. ATP-binding site at the interface of NBD dimer.

A) Structure of the NBD dimer of TAP1 (code PDB: 2IXF) showing the conserved motifs. NBDs are in pale blue and pale green, Walker A in yellow, Walker B in blue, switch or H-loop in gold, Q-loop in purple, the signature in green and D-loop in red. According to [103]

B) Pre-hydrolytic catalytic site, in which the nucleophilic water is positioned for an in-line attack on the ATP γ-phosphate. Oxygen atoms are red, phosphorus green, nitrogen blue, water hydrogen's light grey, and magnesium as a grey sphere. Carbon atoms of monomer A are yellow, and monomer B is cyan. Hydrogen bonds are depicted as dotted red lines. The representation was obtained by molecular dynamics simulations. According to: [120]

Mechanism of transport

Catalytic cycle of the NBD dimer

The following mechanisms have been proposed in the literature to explain the drug translocation process. They differ in several points: (i) the number, one or two, of ATP hydrolyzed to achieve a solute transport, (ii) the type of hydrolysis, sequential or alternate, for each binding site, (iii) the type of NBD interaction, permanent or temporary.

1. Monomer or dimer models: *processive clamp* [108, 121], *tweezers- like* [110] or *switch* models [122, 123]. They were originally derived from structural data showing two states of NBD dimer, either close for the ATP-bound or open for the nucleotide-free. Those models propose that each nucleotide in the dimer is hydrolyzed, followed by NBDs dissociation and ATP reloading. In the *processive clamp* model, the conformational changes leading to substrate translocation (known as "power stroke") is provided by the NBD dimer

dissociation. On the contrary, in the *switch* model the binding of ATP and not its hydrolysis provides the "power stroke" for the transport.

2. The *alternating site* model [124] was proposed after having observed by using vanadate (which mimics the transition state of the γ-phosphate after ATP hydrolysis) that one nucleotide per NBD dimer is hydrolyzed. The authors thus proposed that ATP hydrolysis occurs alternately on each NBD.

3. The *constant contact* model [125-127] is derived from the previous one. It proposes that NBDs are always in contact, opening alternately after ATP hydrolysis. Thus, the energy from hydrolysis at each site is harnessed separately for either substrate transport or the conformational change required for the transporter to reset.

In the following figure some of these models are illustrated.

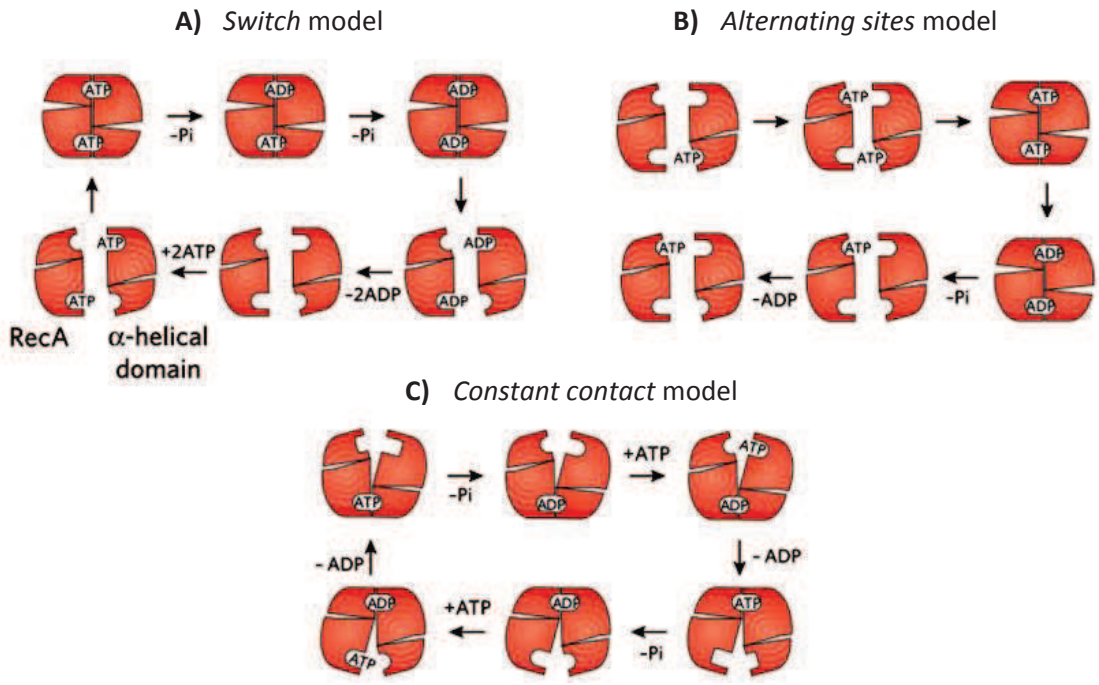


Figure 17. Scheme for the catalytic cycle of the NBD dimer.
The two subdomains in each NBD are indicated as RecA-like and α-helical domain.
Adapted from [128]

Solute translocation

Early in the sixties, Jardetsky proposed an *alternating access* model for the translocation of a solute across physiological membranes. According to this model, the pump displays three features (i) it contains a cavity large enough to accept substrates, (ii) it can adopt inward- and outward-facing configurations in which the cavity is alternately open to the cytoplasmic and periplasmic side of the plasma membrane respectively, and (iii) its binding site displays differential affinities in each

conformation [129]. Biochemical studies [98, 130, 131] together with crystal structures of several ABC transporters support the existence of these two conformations (detailed later). The key of the transport seems to be the 'closing' step of the NBD which reduces the distance between the coupling helices and triggers the TMD flipping from the inward-facing to the outward-facing conformations, a movement governed by drug and ATP binding, and ADP and inorganic phosphate (Pi) release. This suggests that the *switch* model and the *alternating access* model are compatible (Figure 18).

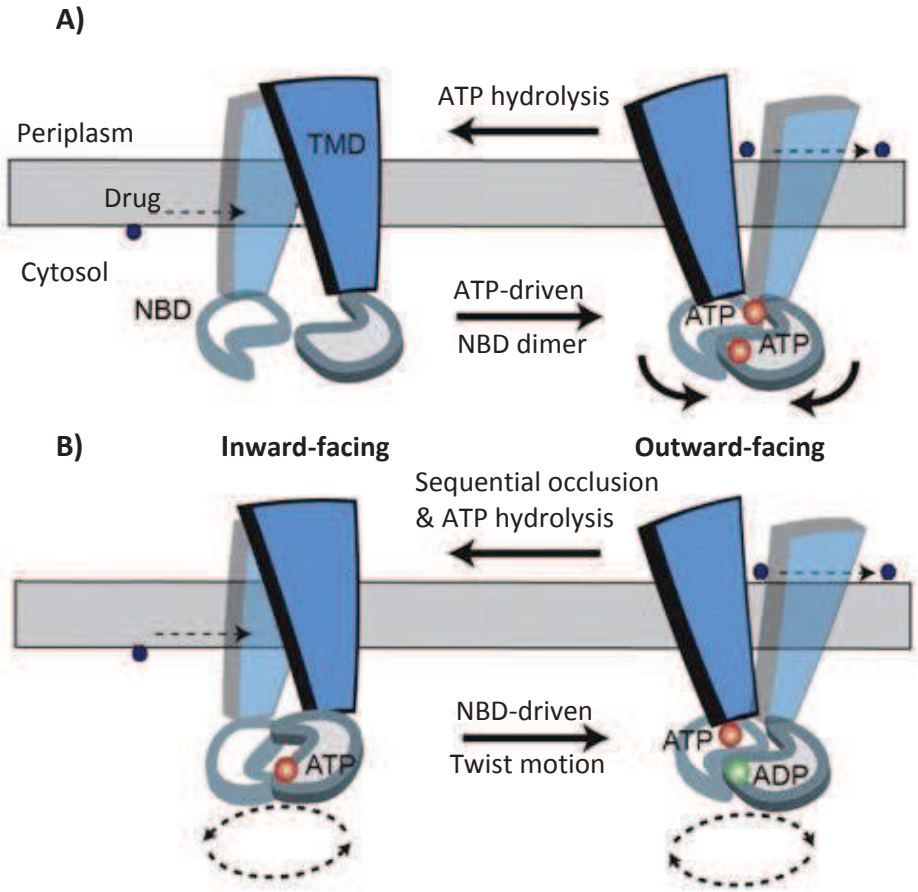


Figure 18. Conformational changes during the transport cycle.

A) The crux of the switch model is the ATP-dependent dimerization of NBDs progresses to pull the TMDs from an inward- to outward-facing conformation. **B)** In the constant-contact model, the inward- and outward-facing configurations are directed via subunit conformational changes and the twisting motion of NBDs alternating the ATP hydrolysis.

Adapted from [132]

The *alternating access* model proposes a common mechanism of transport for ABC importers and exporters [88, 96, 103]. In their outward-facing conformation, ABC importers load molecules from their binding proteins, whereas ABC exporters expel molecules out of the cell. The release of ADP and Pi allows the flip back to the inward-facing conformation in which importers release solutes into the cytoplasm whereas exporters recruit new ones (Figure 19 and Figure 20).

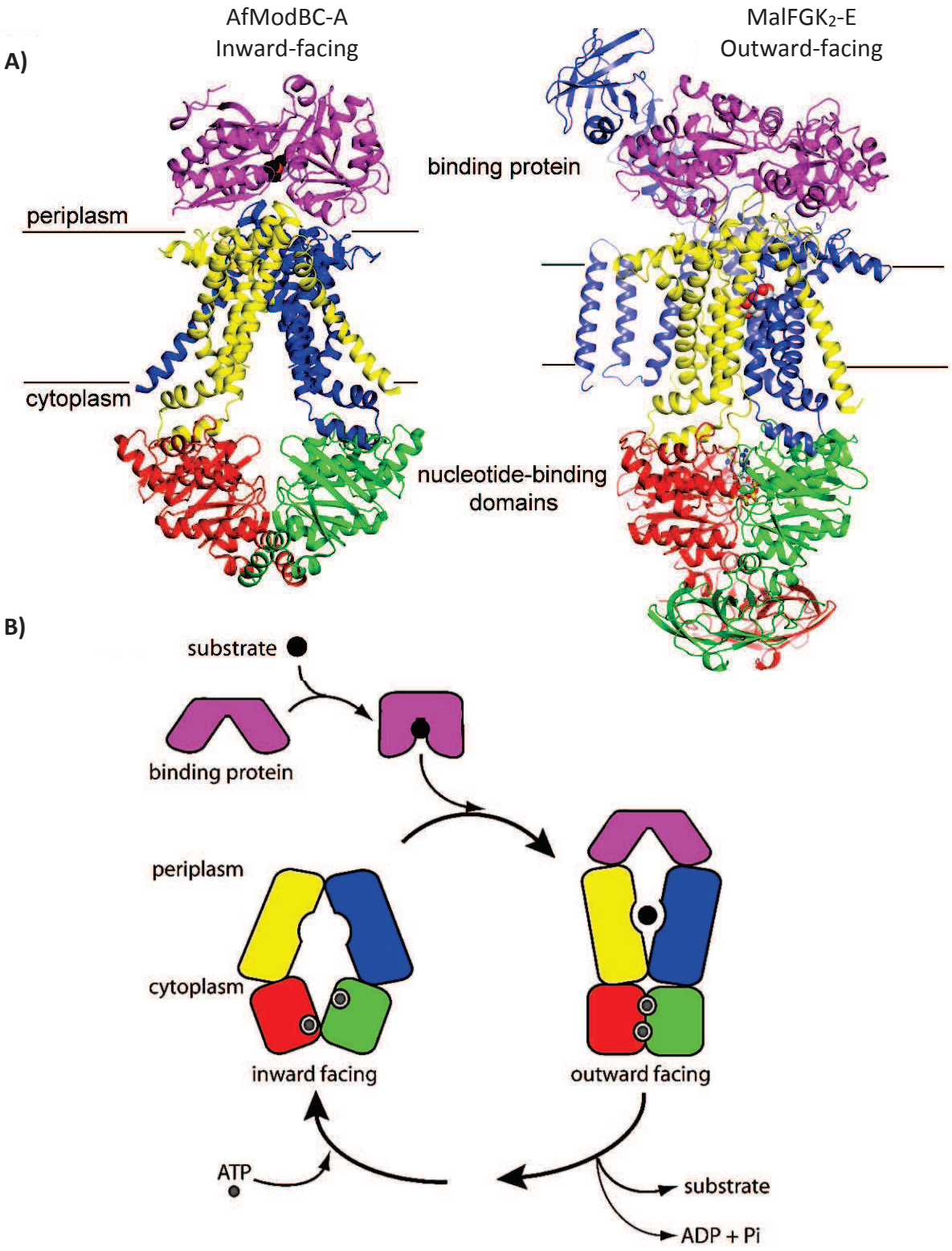


Figure 19. The transport general mechanism for ABC importers.
A) Structures of inward-facing molybdate/tungstate transporter AfModBC-A (PBD code: 2ONK) and the outward-facing maltose transporter $MalFGK_2$-E (PBD code: 2R6G). **B)** Alternating-access in ABC importers.

In purple: the binding protein; in blue and yellow: TMDs; in red and green: ATPase subunits. According to [133]

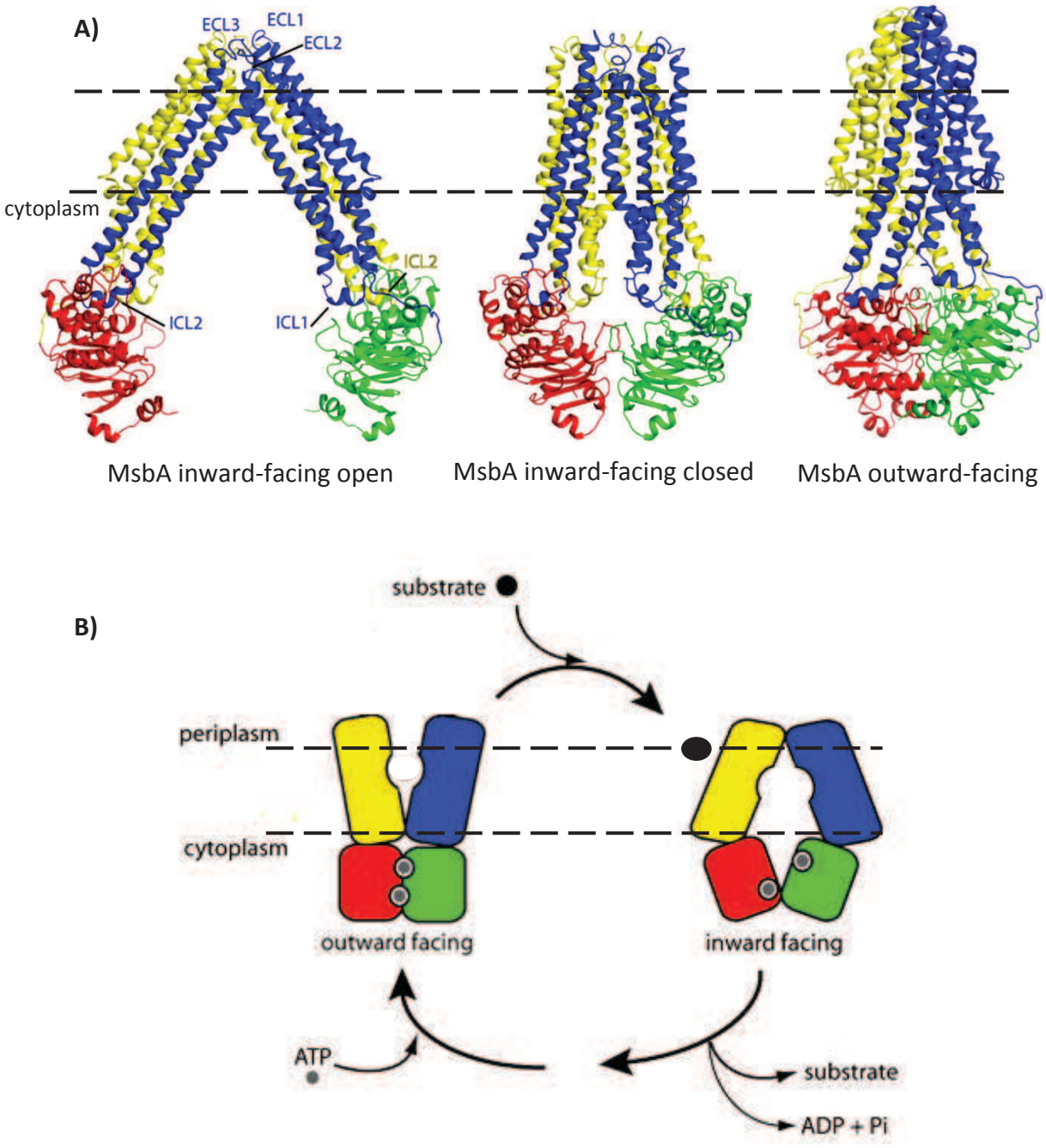


Figure 20. The transport general mechanism for ABC exporters.
A) Structures of MsbA in open inward-, closed inward- and outward-facing conformations [90]. **B)** Alternating-access in ABC exporters.
In blue and yellow: TMDs; in red and green: ATPase subunits.
Adapted from: [133]

X-ray crystals: a brief structural journey

There are several methods to determine the 3D structure of a protein, X-ray crystallography, nuclear magnetic resonance (NMR) and electron microscopy. The oldest one of the methods is the X-ray crystallography, with the first protein structure reported in 1958 [134]. This technique gives valuable information about biological function, interaction of proteins with substrates or inhibitors, etc. However, membrane proteins are more challenging to crystallize than soluble proteins because their insertion in lipid bilayers makes them insoluble in aqueous medium. They are also more difficult to express. As a result, less than 1 % of the 3D structures of membrane proteins have been deposited in the protein data bank (PDB) [135]. The main difficulty is that to be crystallized, the ABC transporters, like all membrane proteins, must be kept in solution with detergents. Thus, expression, purification and good-quality crystals of a membrane protein is often a bottleneck. And despite the valuable information that X-ray crystal structures can bring, we need substantial biochemical analyses to fully interpret these structures.

Structures of the NBD

The first crystallographic data for ABC transporters were obtained with the cytoplasmic domains, more soluble and therefore more easily to crystallize that the whole membrane protein. Thus, in 1998 the structure of the NBD HisP (histidine permease) binding the ATP has been described. At a resolution of 1.5 Å this hallmark structure gave the first overall view of a NBD monomer, and also provided detailed insight into the binding of ATP. Since then several NBD structures were resolved by high-resolution X-ray crystallography showing: two main conformations during the ATP hydrolysis cycle, open without any nucleotide and close with ATP bound. Several amino acids involved in the binding and hydrolysis of nucleotide have been identified, as well as the head-to-tail arrangement of the NBD–NBD dimer interface (as discussed above).

Details of the origin, function, PDB code and resolution for these structures are shown in Table 4.

	Name	organism	Substrate or Function	PDB code	Resolution Å, (ligand or conformation)	Ref.
ABC importers	HisP	*S. typhimurium*	histidine	1B0U	1.5 (ATP)	[106]
	MalK	*T. litoralis*	maltose	1G29	1.9 (PP_i)	[111]
		E. coli	maltose	1Q1E 1Q1B 1Q12	2.9 (apo, open); 2.8 (apo, semi-open); 2.6 (ATP, closed)	[110]
		P. horikoshii	maltose	1V43 1VCI	2.2 (apo, open conformation); 2.9 (ATP)	[136]
	BtuCD	*E. coli*	Vitamin B12	1L7V	3.2 (Cyclovanadate)	[94]
	GlcV	*S. solfataricus*	glucose	1OXS 1OXT 1OXU 1OXV 1OXX	1.65 (apo); 2.1 (apo); 2.1 ($ADP\text{-}Mg^{2+}$); 1.95 ($AMPPNP\text{-}Mg^{2+}$); 1.45 (I^-)	[137] [138]
	CysA	*A. acidocaldarius*	sulfate/thiosulfate	1Z47	2.0 (apo)	[139]
ABC Exporters	TAP1	*H. sapiens*	peptides	1JJ7	2.4 ($ADP\text{-}Mg^{2+}$)	[140]
	HlyB	*E. coli*	hemolysin	1MT0 1XEF	2.6 (apo) 2.5 ($ATP\ Mg^{2+}$)	[141] [142]
	NBD1 of MRP1	*H. sapiens*	multiple drugs	2CBZ	1.5 ($ATP\text{-}Mg^{2+}$)	[143]
ABC ATPase superfamily	MJ1267	*M. jannaschii*	branched-chain amino acid	1G6H 1GAJ	2.5 (apo); 1.6 ($ADP\text{-}Mg^{2+}$)	[144]
	MJ0796	*M. jannaschii*	Unknown	1F3O	2.7 ($ADP\text{-}Mg^{2+}$)	[145]
	MJ0796 (E171Q)	*M. jannaschii*	Unknown	1L2T	1.9 ($ATP\text{-}Na^+$)	[108]
Ion channel	NBD1 of CFTR (ABCC7)	Mouse, Human	Control de flow of H_2O, chloride and thiocyanate ions	1R0W 1Q3H 1R0Z 1R0X 1R0Y 1R10 1XMI 1XMJ	2.2 (apo); 2.5 (AMPPNP); 2.35 ($ATP\ Mg^{2+}$); 2.2 (ATP); 2.55 (ADP); 3.0 ($ATP\ Mg^{2+}$); 2.25 ($ATP\ Mg^{2+}$); 2.3 ($ATP\ Mg^{2+}$)	[146] [147]
Class 2 ABC transporters	Rad50	*P. furiosus*	DNA repair	1F2U 1F2T	1.6 (apo); 2.1 ($AMPPNP\text{-}Mg^{2+}$); 2.6 (ATP)	[112]
	MutS	*E. coli*	DNA repair	1E3M	2.2 (G-T mismatch DNA)	[148]
	MutS	*T. aquaticus*	DNA repair	1EWR 1EWQ	3.19 (apo); 2.2 (ADN)	[149]

Table 4. Crystal structures of ABC-NBDs

PDB: protein data bank; Apo: absence of nucleotide; ATP: adenosine-5'-triphosphate; ADP: adenosine-5'-diphosphate; AMPPNP: adenosine-5'(βγ-imido)triphosphate;

Adapted from: [150]

Structures of ABC transporters

The first structural data were obtained at low resolution by electron microscopy to human transporters: ABCC1 [151], ABCB1 [152-154] the bacterial transporter BmrA from *Bacillus subtilis* [155] and pdr5p from *Saccharomyces cerevisiae* [156]. These studies confirmed the general topology of ABC transporters, consisting of at least two TMD (each containing six transmembrane helices) and two NBDs. In 2002 the first full structure of an ABC transporter was determined by X-ray crystallography, 20 years after cloning the first ABC transporter: the histidine permease.

There are now several x-ray structures of both bacterial and eukaryotic ABC transporters (BtuCD, HI1470/1, HmuUV, ModBC, $MalFGK_2$, MetNI, Sav1866, MsbA, ABCB1, ABCB10 and TM287/288 heterodimer; detailed in the Table 5 and Table 6), solved at medium to low resolution (2.2-5.5 Å). All of them have been determined either with or without nucleotides bound. Three were resolved with either bound substrate (the maltose permease, Figure 22; and molybdate/tungsten transporter ModBC, Figure 19A) or inhibitors (murine ABCB1, Figure 64). Until today, none ABC exporter has yet been resolved in the presence of substrate transported.

Details of the origin, expression system, PDB code and resolution for these structures are shown in Table 5 and Table 6.

	Protein (substrate or function)	organism	Exp. system	PDB code	Res. (Å)	Nucleotide state	other ligands	Conf	Ref
Type II	BtuCD (vitamin B12)	*E. coli*	*E. coli*	1L7V	3.2	Apo or V_4O_{12}		OF	[94]
				2QI9	2.6	Apo	complex with SBP (BtuF)	IF-O	[157]
				4DBL	3.49	Apo		IF-O	[158]
				4FI3	3.47	AMPPNP		IF-C	[159]
	HI1470/1 (metal-chelate-type)	*H. influenza*	*E. coli*	2NQ2	2.4	Apo		IF-O	[95]
	HmuUV (heme)	*Y. pestis*	*E. coli*	4G1U	3.0	Apo		IF-O	[160]
Type I	ModBC (molybdate/ tungsten)	*A. fulgidus*	*E. coli*	2ONK	3.1	phosphate	complex with SBP ModA binder WO_4	IF-O	[91]
		M. acetivorans	*E. coli*	3D31	3.5	Apo	WO_4 bound in the regulatory domain	IF-O	[161]
	$MalFGK_2$ (maltose)	*E. coli*	*E. coli*	2R6G	2.8	ATP	complex with MBP	OF	[92]
				3FH6	4.5	Apo	absence of MBP	IF	[162]
				3PV0	3.1	Apo	complex with MBP maltose	Pre-T	[163]
				3PUY	3.1	AMPPNP			
				3PUZ	2.9	AMPPNP			
				3RLF	2.20	Mg-AMPPNP	complex with MBP; maltose	OF	[115]
				3PUX	2.30	ADP-BeF_3			
				3PUV	2.40	ADP-VO_4			
				3PUW	2.30	ADP-AlF_4			
	MetNI (methionine)	*E. coli*	*E. coli*	3DHW	3.7	Apo		IF-O	[93]
				3TUI	2.9	ADP	Selenium Methionine	IF-O	[164]
				3TUJ	4.0	Apo			

Table 5. Crystal structures of ABC importers.

PDB: protein data bank; Apo: absence of nucleotide; SBP: substrate binding proteins; MBP: maltose binding protein; ATP: adenosine-5'-triphosphate; ADP: adenosine-5'-diphosphate; AMPPNP: adenosine-5'(βγ-imido)triphosphate; Conf: conformation; OF: outward-facing; IF-O: inward-facing open; IF-C: inward-facing closed; Pre-T: pre-translocation (intermediate state between IF and OF); Ref: reference.

Adapted from: [23]

The structures of BtuCD (PDB code: 1L7V) and BtuCD-F (in complex with the binding protein BtuF, code PDB: 2QI9) show that the NBDs are essentially in the same open dimer configuration.

However, in presence of non-hydrolysable ATP-analogue, AMPPNP (PDB code: 4FI3), the NBDs form the expected closed sandwich dimer and the TMDs adopt a new conformation, Figure 21.

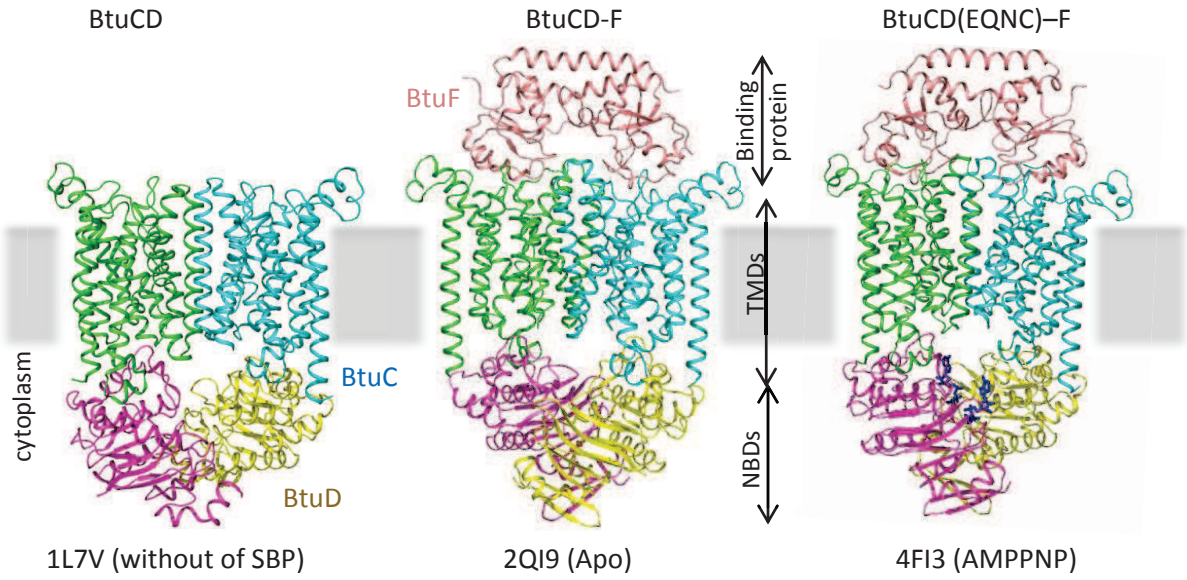


Figure 21. Different structures of vitamin B12 transporter (BtuCD–F).
The structures are shown in cartoon. BtuF is colored rose pale, BtuC subunits are in blue and green and BtuD subunits are in yellow and magenta. AMPPNP is shown in blue sticks.

The maltose transporter $MalFGK_2$ is the one of the best characterized in x-ray crystallography showing snapshots of each step of the cycle transport (Figure 22): either with his maltose binding protein (MBP) or without, with his substrate (maltose) in two different positions (in the MBP and in the TMDs) and with several non-hydrolysable ATP-analogues (AMPPNP alone or with Mg^{2+}, ADP in conjunction with phosphate BeF^-_3, VO^-_4, AlF^-_4) showing pre-translocation states and confirming that ABC transporters catalyze ATP hydrolysis via a general base mechanism. Maltose diffuses through the outer membrane via maltoporin into the periplasm. The maltose binding protein (MBP) binds maltose and interacts in the closed conformation with the resting state $MalFGK_2$. When MBP binds to the $MalFGK_2$ the NBDs are brought closer to each other and the transmembrane subunits are reorientated causing the MBP to open. This outward-facing conformation allows transfer of maltose from MBP to the transmembrane subunit binding site and simultaneously places ATP at the catalytic site for hydrolysis. Once ATP is hydrolyzed the subunits most likely reorient towards the cytoplasm and the substrate is released into the cell [115, 163].

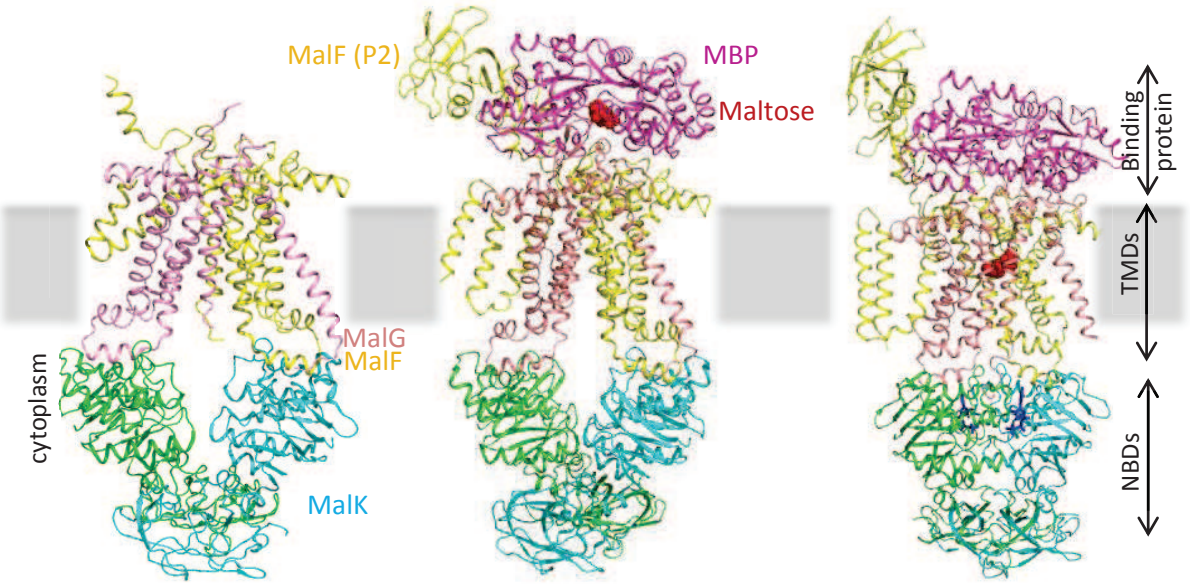


Figure 22. Different structures of maltose transporter ($MalFGK_2$).

The structures are shown in cartoon. Maltose binding protein (MBP) is colored magenta. MalF is colored in yellow. MalG is in rose pale. MalK dimer in green and blue. Maltose is shown in red spheres. AMPPNP is shown in blue sticks.

Protein (substrate or function)	organism	Exp. system	PDB code	Res. (Å)	Nucleotide state	other ligands	Conf	Ref
Sav1866 (multidrug transporter)	*S. aureus*	*E. coli*	2HYD	3	ADP		OF	[89]
			2ONJ	3.4	AMPPNP		OF	[165]
MsbA (lipid A)	*E. coli*	*E. coli BL21*	3B5W	5.3	Apo		IF-O	[90]
	V. cholerae		3B5X	5.5	Apo		IF-C	
	S. tiphymurium		3B5Z	4.2	ADP-VO_4		OF	
			3B5Y 3B60	4.5 3.7	AMPPNP		OF	
ABCB1 or P-gp (multidrug transporter)	*M. musculus*	*P. pastoris*	3G5U	3.8	Apo		IF-O	[166]
			3G60	4.4	Apo	QZ59-RRR		
			3G61	4.35	Apo	QZ59-SSS		
	C. elegans	*P. pastoris*	4F4C	3.4	Apo		IF-O	[167]
	M. musculus	*P. pastoris*	4KSB	3.8	Apo	Nanobody bound to 1st NBD	IF-O	[168]
			4KSC	4.0	Apo			
			4KSD	4.1	Apo			
	M. musculus Corrected 2009		4M1M	3.8	Apo		IF-O	[169]
			4M2S	4.4	Apo	QZ59-RRR		
			4M2T	4.35	Apo	QZ59-SSS		
ABCB10 (Heme biosynthesis)	*H. sapiens*	*insect cells*	4AYT	2.85	AMPPCP		IF-O	[170]
			4AYX	2.90	AMPPCP			
			4AYW	3.30	AMPPNP			
			3ZDQ	2.85	Apo			
TM287/288 heterodimer	*T. maritima*	*L. lactis*	3QF4	2.90	AMPPNP		IF-O	[171]

Table 6. Crystal structures of ABC exporters

PDB: protein data bank; Apo: absence of nucleotide; ADP: adenosine-5'-diphosphate; AMPPNP: adenosine-5'(βγ-imido)triphosphate; AMPPCP: β-γ-methyleneadenoside 5'-triphosphate; Conf: conformation; OF: outward-facing; IF-O: inward-facing open; IF-C: inward-facing closed; Ref: reference.

Adapted from: [23]

Structures Sav1866 (Figure 23) and MsbA (Figure 20A) in complex with ADP reveal a hydrophilic cavity open to the outside of the cell (outward facing conformation) and the formation of a closed dimer of NBD. It is remarkable that a single transporter, MsbA, can display conformations ranging from NBDs in contact until complete separation (PDB codes: 3B60, 3B5X, 3B5W).

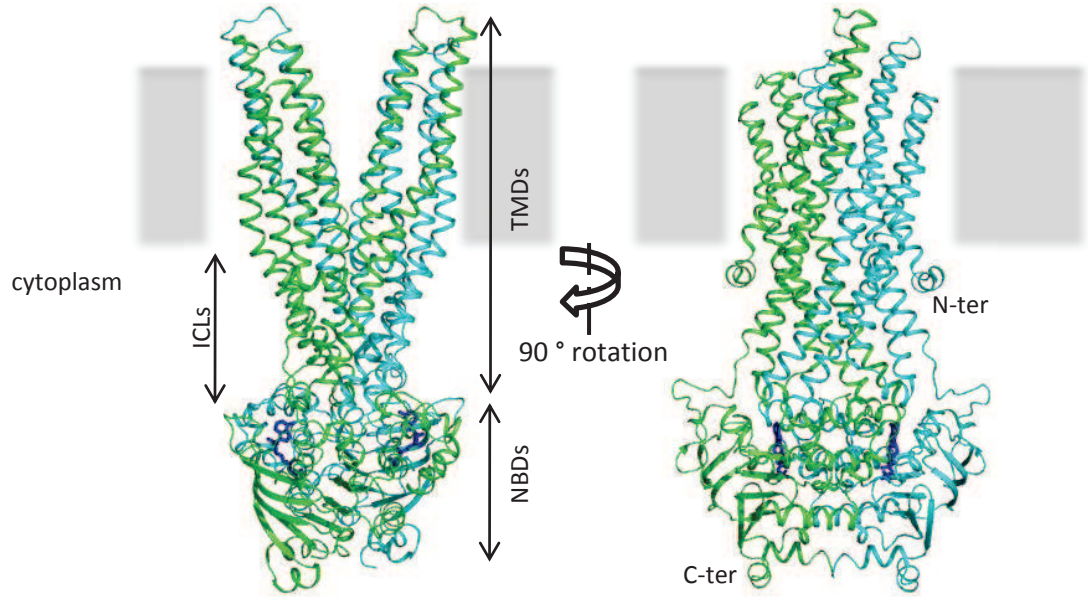


Figure 23. 3D structure of the homodimer Sav1866.
Each half of the transporter is showed in cartoon and colored in green and cyan. ATP molecules are shown in blue sticks.

The first eukaryotic ABC transporter structure is the murine ABCB1 which is 87 % identical in amino acid sequence to the human ABCB1. The protein was crystallized nucleotide-free in the apo form (Figure 24A) and in complex with two stereoisomers of the cyclic hexapeptide inhibitor QZ59 (Figure 24B). Thanks to the presence of 3 Selenium atoms in these compounds, one QZ59-RRR could be unambiguously located in one structure (3G60) and two molecules QZ59-SSS were identified in another one (3G61). Both structures were quite close, showing that the QZ59 compounds are located in the same cavity, the QZ59-SSS being positioned up and above the corresponding location of the QZ59-RRR (Figure 24A). This is the first time that two molecules were visualized on an ABC transporter, confirming earlier reports that two compounds can bind simultaneously to ABCB1 [172, 173]. The three essential features in these structures consist in a 6,000 $Å^3$ internal cavity in which drugs should bind and the presence of two portals in the inner leaflet part of the protein that allows the entry of hydrophobic drugs from the lipid bilayer, and a separation of the NBDs of ~30 Å. The conformational changes of these structures are displays in Figure 24C.

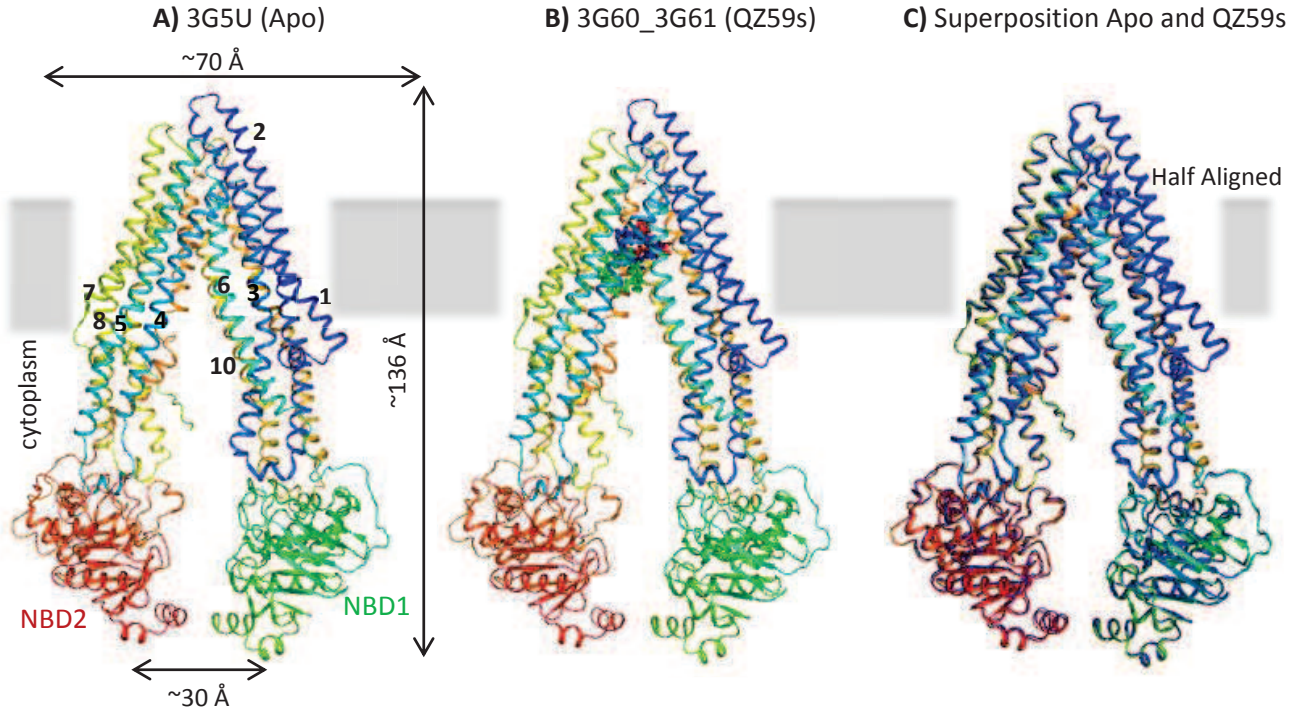


Figure 24. 2009 Chang's group structures.

A) Apo P-gp in inward-facing open conformation. **B)** P-gp in complex with two selenohexapeptide inhibitors in the central cavity. **C)** Conformational changes between these structures.

All the structures are show in cartoon colored by rainbow (N-terminal in blue, C-terminal in red). QZ59-RRR is colored in blue, the entire QZ59-SSS in green and half in red.

"Half Aligned": X-ray structures aligned by using residues in TMD1 and NBD1 (residues 33–209, 852–961, and 320–626)

This first molecular description of compounds bound to P-gp triggered a lively debate in the scientific community [174, 175] because of the low resolution, the unexpected separation of the NBDs (~ 30 Å) and the limited conformational changes between the empty form of the protein and those with QZ59 enantiomers (Figure 24C). Nevertheless: (i) the broad majority of structures of ABC transporter are in these range of resolution. (ii) The unexpected separation of NBDs is due to the concentration of ATP in cell (3–5 mM) far exceeds the binding affinity of ABC transporter for ATP (K_m ATP 0.3–1 mM), meaning a futile cycle, or locking the transporter in a non-productive conformation. But, as suggested by Linton and Zolnerciks locking at the maltose transport system, TMDs may prevent NBD:NBD closed dimer in presence of ATP and absence of substrate [176]. Furthermore, biochemical experiments give support to the existence of inward-facing and outward-facing conformations in different ABC exporters [130, 131, 177], , and a study of MsbA in electron paramagnetic resonance showed a significant separation of the NBDs (55–60 Å between the spin labels) which in the presence of ADP plus vanadate closes by ~30 Å [98]. (iii) The limited conformational changes is perhaps due to the inhibitors which cannot cause NBD dimerization.

In 2012 Chen's group published one structure of P-gp from *Cænorhabditis elegans* at 3.8 Å of resolution, Figure 25A. Which was also captured in a nucleotide-free in the apo inward-facing conformation but is even more open to the cytoplasm that mouse P-gp (NDB's distance of 46 Å for *C. elegans vs* 29 Å for mouse P-gp, Figure 25). In the structure of *C. elegans* P-gp, some transmembrane helices (TM1-2, TM6-8 and TM11, grey regions in Figure 25B) are consistent with the 2009 mouse P-gp structure. Others TM show a possible different conformation (TM9-10 and TM12, not easily comparable, green regions in Figure 25B). While, three TM (TM3-4 and TM5, magenta regions in Figure 25B) which can be compared displays significant differences [167].

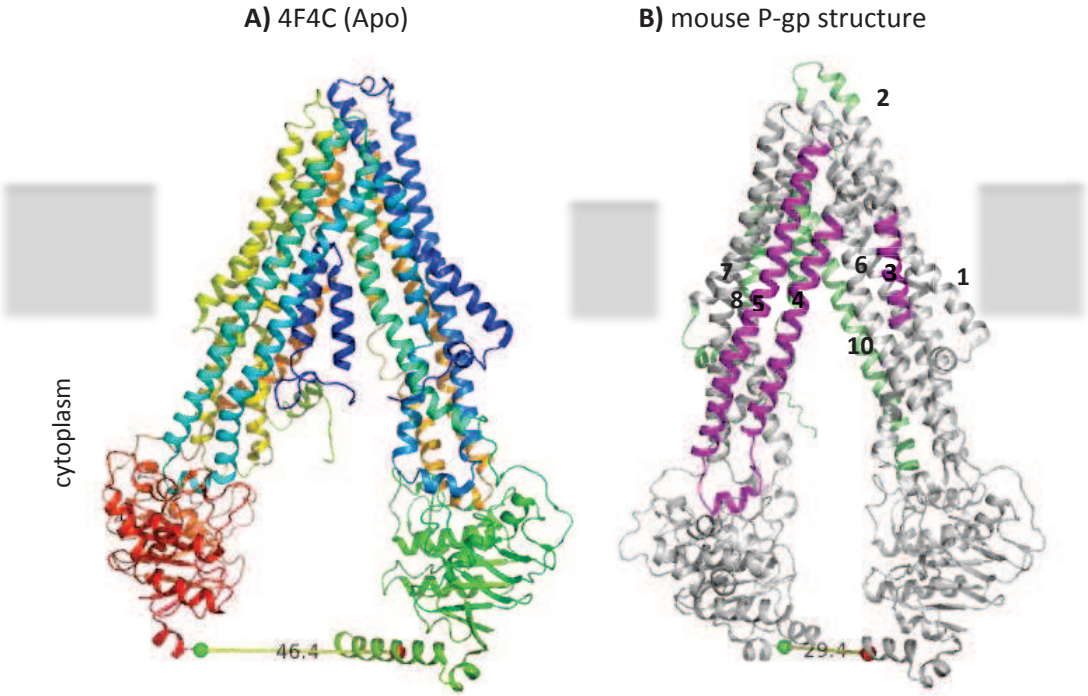


Figure 25. *C. elegans* and mouse P-gp structures.

A) Structure of *C. elegans* P-gp is show in cartoon colored by rainbow (N-terminal in blue, C-terminal in red). **B)** Compared zones between *C. elegans* and mouse P-gp. In magenta: regions containing register errors; green: regions where structures are not directly comparable; grey: regions where the structures are consistent.

The distance between C-term of NBD1 (residue 626 for mouse and 654 for *C. elegans* in red) to C-term of NBD2 (residue 1260 for mouse and 1306 for *C. elegans* in green) are indicated.

According to [167].

Recently, three new crystal structures of mouse P-gp were published [168], and the structures of mouse P-gp from 2009 were refined [169].

Several structures of the human mitochondrial ATP transporter ABCB10 are also known. All of them in an inward-facing conformation either in presence or absence of non-hydrolysable ATP-analogues (Figure 26A), with a few change in the distance of the NBDs (Figure 26B). These is unexpected and in contrast to other transporter structures which adopt an open-outwards conformation in complex with ATP.

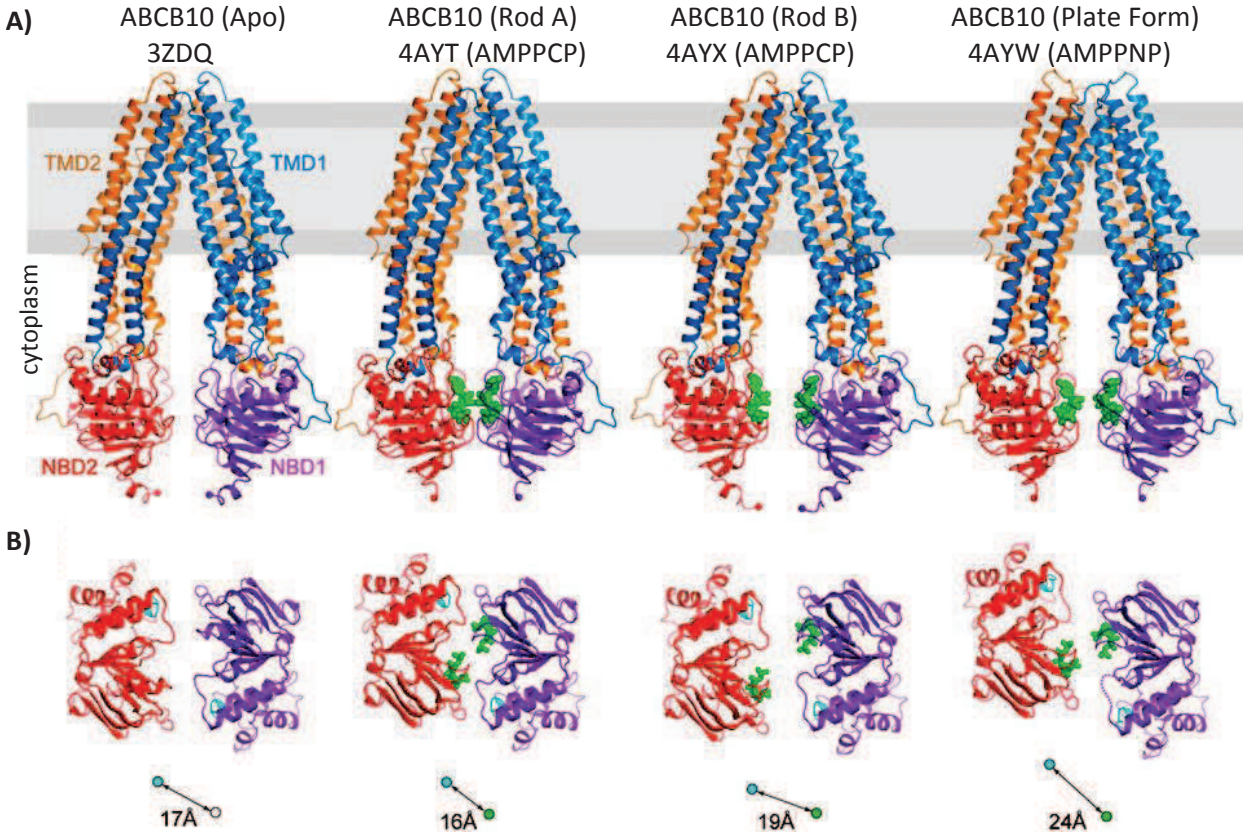


Figure 26. Different structures of human ABCB10 transporter.

A) Structures of human ABCB10 in the absence (apo) and in presence of bound nucleotide analogs. The structures have a single monomer in the asymmetric unit, the dimer was generated by a crystallographic twofold; nucleotides are in green

B) Separation of the NBDs of the structures shown in A.

Which was measured by the distance between the nucleotide γ-phosphate (green) and the C-α of the first glycine in the catalytic C loop of the adjacent NBD (cyan). In the Apo state the position of the γ-phosphate has been inferred by superposition of the AMPPCP complex.

According to: [170]

MDR PHENOTYPE- LINKED ABC TRANSPORTERS

As mentioned above, several transporters of the ABC superfamily are implicated in the MDR phenotype. The first discovered was the P-glycoprotein (P-gp or ABCB1) [55], the second was the Multidrug resistance-associated protein (MRP1 or ABCC1) [72] and the last was the breast cancer resistance protein (BCRP or ABCG2) [73-75]. The timeline of major discoveries of MDR-Linked Human ABC Transporters is displays in the appendix Figure 2.

P-glycoprotein (P-gp, ABCB1)

Characteristics

The best known and best characterized of the MDR ABC transporters is the P-gp. This protein is encoded by the *MDR1* gene in humans and by the *mdr1a* and *mdr1b* genes in rodents [178]. Generally, P-gp is an extremely stable protein with a half-life of 48-72 h for the human and approximately 18 h for the mouse isoforms [179]. The human protein is a single polypeptide of 1280 amino acids leading to a molecular mass of 170 kDa after post-translational modification. The two homologous halves (with ~ 43 % of homology between them) are connected to each other by a linker of ~80 residues (633~709), which can be phosphorylated by protein A and C kinases at four serine or threonine residues in this region (S661, S667, S671, and S683). P-gp needs to be glycosylated at three asparagine residues (N91, N94, and N99) for a proper trafficking to the plasma membrane. However absence of either glycosylation or phosphorylation does not affect its transport activity [180]. Only its molecular weight is affected in the absence of glycosylation, 140 kDa. P-gp contains 11 Trp residues distributed throughout the protein, with 8 in the N-terminal half and 3 in the C-terminal half [181], which provide a source of intrinsic protein fluorescence. In addition, there are 7 Cys residues in P-gp, including one in each NBD, which are useful for labeling P-gp with specific Cys-reactive probes [182].

P-gp has a specificity to transport large hydrophobic and amphiphilic molecules ranging between 250 Da, for smaller molecules (e.g., morphine 285 Da and phenytoin 252 Da) and can go up to about 1250 Da to larger molecules (e.g., actinomycin D 1255 Da and PSC833 1214 Da). These compounds often contain aromatic rings and a positive charge at physiological pH [183], and can be divided in five groups [184]:

- weak lipophilic bases (e.g., doxorubicin)
- lipophilic cations (e.g., rhodamine-123)
- the neutral polycyclic compounds (e.g., aldosterone)
- amphiphilic molecules (e.g., Triton X-100)
- hydrophobic peptides (e.g., valinomycin)

One of the current challenges is to determine the characteristics of the specific recognition of transport substrates of P-gp. In this regard, many structure-activity studies have been conducted to

identify common physico-chemical characteristics to all these molecules and to identify potential chemical consensus motifs for all transport substrates of P-gp.

Physiological expression

As mentioned above, this protein is widely expressed throughout the body, in excreting organs (intestine, liver and kidney) and physiological barriers (such as the placenta, brain or testis), Figure 27. Several studies confirmed its participation in detoxification and its protective role against potentially toxic endogenous and exogenous substances (environmental poisons, drugs) [183].

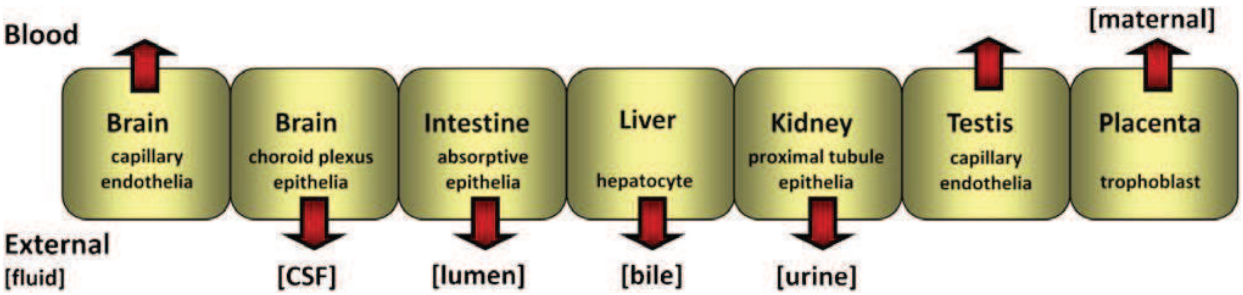


Figure 27. P-gp expression and direction of net transport at different cellular barriers.
CSF: cerebrospinal fluid. According to [185].

Intracellular localization and possible traffic

Confocal microscopy studies reveal that P-gp is localized in the plasma membrane and in other intra-cellular compartments, such as endoplasmic reticulum (ER), Golgi, several endosomes and lysosomes (reviewed in [186]). Its localization in the mitochondria was reported but this observation has not been confirmed. A 150 kDa intermediate P-gp protein is synthesized in the ER but if misfolded it is degraded by the proteasome. After folding in the ER, the protein moves to the Golgi for further glycosylation as a 170 kDa mature protein. P-gp is then delivered to the plasma membrane in different pathways, such as: vesicles and the intracellular endosomal system (e.g., Rab5 protein). Finally, P-gp is recycled by Rab GTPases (e.g., Rab11a protein). The Figure 28, shows the sites of P-gp synthesis (ER), modification (Golgi), trafficking and recycling (endosomes), and degradation (lysosome and proteasome).

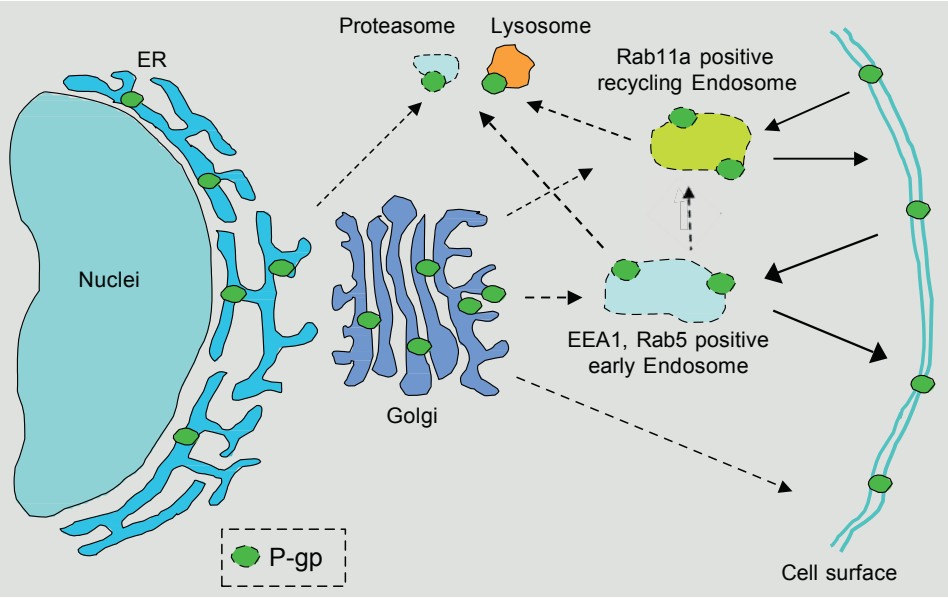


Figure 28. Cellular localization of P-gp and possible traffic and cycling routes. According to [186].

Regulation of P-gp expression

The induction of P-gp is not fully understood (reviewed in [187]). But, up to now several factors are known to promote (e.g., toxic compounds), to control (e.g., the pregnane X receptor, PXR) or to modulate (e.g., vitamin D receptor and thyroid hormone receptor) the *MDR1* gene transcription. Stress stimuli (e.g., oxygen, pH) may also up-regulate the *MDR1* gene. Furthermore, other non-substrates of P-gp may enhanced its expression through a common transcription factor. For example, PXR also regulates the transcription of the CYP3A subfamily (e.g., cytochrome P450, known for its involvement in drug resistance), and xenobiotics that stimulate cytochrome P450 also induce the expression of P-gp.

Involvement in MDR tumors

P-gp is highly overexpressed in the majority of MDR tumor cells of leukemic and solid tumors (reviewed in [188-190]). Certain tumors, intrinsically drug resistant, have high levels of P-gp expression (e.g., colon, kidney, pancreas, and liver carcinoma). While others, initially with low basic levels of P-gp expression (such as hematological malignancies) generally display an increases after chemotherapy.

For acute myelogenous leukemia (AML), numerous studies have found a very good correlation between P-gp expression levels and poor prognosis and response to chemotherapy. For example, P-gp expression was found in 30 % of the patients at the initial diagnosis. For patients with disease relapse, overexpression of P-gp was detected in more than 50 % of the patients. The degree of P-gp expression in tumors from these patients was comparable to the P-gp expression seen in MDR resistant tumor cells selected through *in vitro* assays.

Multidrug resistance-associated protein (MRP1, ABCC1)

Characteristics

MRP1 was the first representative of the MRP (Multidrug Resistance-associated Protein) subfamily. As P-gp, MRP1 has three glycosylation sites, two located near the N-terminus of residues 19 and 23 and a third located in the Asn1006, which seem to be not necessary for the transport activity [191]. About its phosphorylation, there are very few studies. The additional domain (TMD0) contains five transmembrane helices, and an extracellular loop with the N-terminus located on the outside of the cell.

This protein has the specificity of transport of anionic substrates, typically conjugated to glutathione (GSH), sulfate or glucuronate. Its preferential substrate is leukotriene C_4 or LTC_4, but it also transports GSH itself [192]. Thus, MRP1 uses several modes of transport involving GSH (Figure 29).

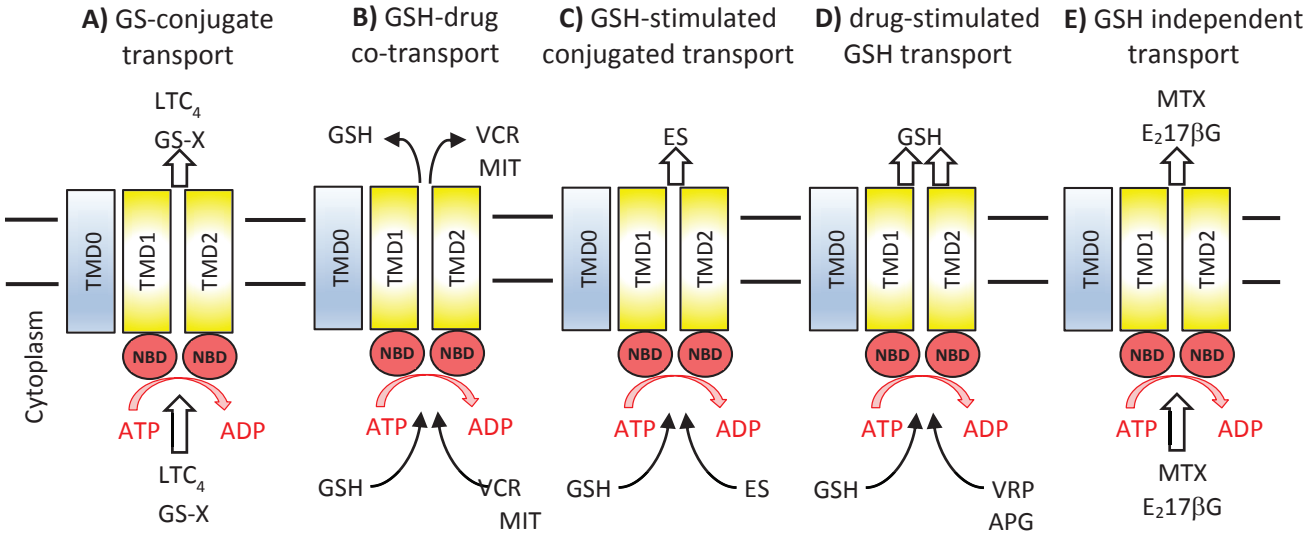


Figure 29. Representation of different modes of transport by MRP1 that involve GSH.
LTC_4: leukotriene C_4; GS-X: GSH conjugate; VCR: vincristine; MIT: mitoxantrone; ES: estrone sulfate; APG: apigenin; VRP: verapamil; MTX: methotrexate; $E_2$17βG: estradiol glucuronide.
Adapted from [193].

Physiological expression

MRP1 expression is found in the adrenal gland, breast, lung, kidney, heart, blood-brain barrier and endocrine tissue (reviewed in [194]). Studies in mrp1 -/- knockout mice have indicated that MRP1 is a transporter functioning as a resistance factor for some mucosal tissues and bone marrow. Furthermore, MRP1 is involved in the physiological barrier for the protection of the testis and the choroids. In contrast to P-gp, MRP1 is expressed on the basolateral site of epithelial cells. Thus, MRP1 transports its substrates away from the luminal surface (e.g. intestinal lumen, bile duct) into the tissue beneath it.

Due to its strong affinity for glutathione and glucuronate conjugates, both products from the phase II cellular detoxification of hydrophobic compounds, MRP1 holds an important role in the physiological detoxification system.

Involvement in MDR tumors

As for P-gp, higher or elevated levels of MRP1 are detected in many different tumor types, including solid tumors (non-small-cell lung cancer, gastrointestinal carcinoma, melanoma, neuroblastoma and cancers of the breast, ovary and prostate), and hematological malignancies including leukemias (reviewed in [188-190]). Some studies have linked the presence of MRP1 to poor treatment outcome, but no comprehensive picture of its role in clinical MDR has emerged yet.

Breast cancer resistance protein (BCRP, ABCG2)

Characteristics

ABCG2 is the only member of the G- subfamily involved in MDR. But, like others member of this subfamily, ABCG2 is a half transporter with intra-and extracellular loops. The third extracellular loop, called ECL3 (about 60 to 80 residues), constitutes a porphyrin-binding domain which is strategically positioned to release the bound porphyrin to extracellular partners such as the human serum albumin [195]. ECL3 is glycosylated in the N 596 position, but as P-gp and MRP1 this glycosylation is not required either for the expression or the function of ABCG2 [196]. ABCG2 also has seven cysteine residues, three of which are located at ECL3 (C592, C603 and C608). Cysteine 603 has been described to be involved in dimerization of the transporter, while cysteine 592 and 608 form an intramolecular disulfide bridge potentially involved in addressing and transport function of ABCG2 [197].

Like MRP1 (but unlike P-gp) ABCG2 appears to transport both positively and negatively charged drugs, including protoporphyrin IX [198], uric acid [199] and sulfated estrogens [200]. The most important mutation of ABCG2 is the arginine residue at position 482, R482T and R482G. Since they are capable of transport certain substrates unsupported by the wild type, such as doxorubicin, daunorubicin and rhodamine 123 [201, 202]. However none of these mutants have been found in the human population [203]. It has been described that other mutations such as E446, N557 and H630, also alter the substrate specificity of ABCG2 [204].

Physiological expression

Similarly to P-gp, ABCG2 is widely expressed throughout the body, but with the highest expression in the placenta. This findings suggests that ABCG2 plays a protective role for the fetus, preventing potentially harmful substances from entering the uterus [205]. Moreover, ABCG2 expression is found in many other organs and tissues including the liver, ovaries, colon, kidneys and brain microvessels [206]. Tissue distribution studies have demonstrated an extensive overlap between ABCG2 and P-gp, suggesting that ABCG2 plays a similar physiological role in providing protection from xenobiotics. Indeed, ABCG2 regulates the distribution of xenobiotics to various

tissues like the liver, intestine and placenta [207]. Interestingly, ABCG2 has been found to be highly expressed in stem cell subpopulations, such as hematopoetic stem cells (CD-) [208] or embryonic kidney cells [209](Figure 30). These findings imply that ABCG2 might be involved in the general regulation and protection of stem cells.

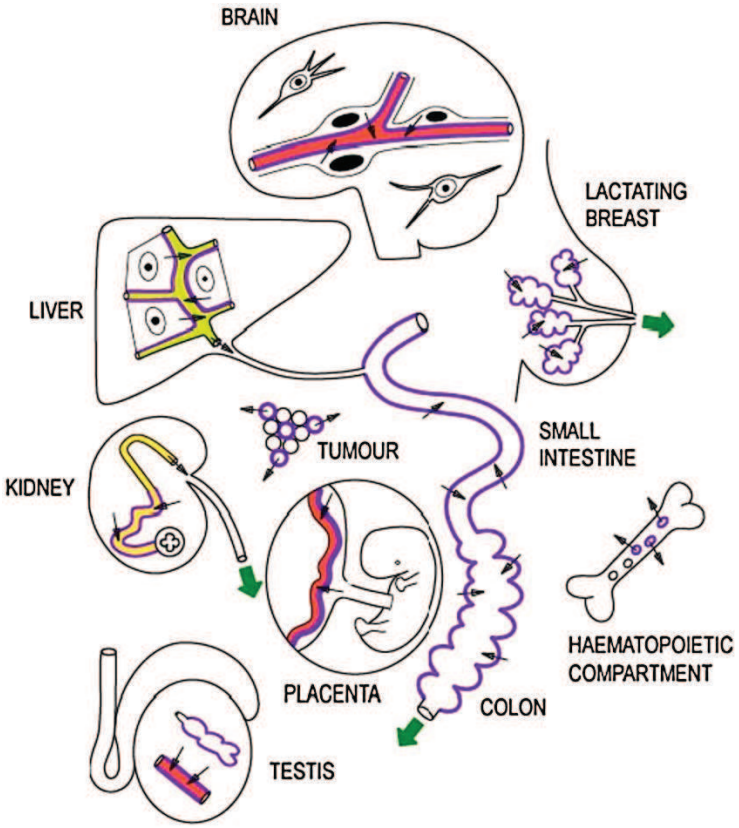


Figure 30. ABCG2 expression and direction of net transport throughout the body According to [210].

Involvement in MDR tumors

Immunohistochemical investigations of various tumors have revealed that ABCG2 is expressed in over 40 % of solid tumors (reviewed in [188-190]). ABCG2 overexpression was detected in tumor specimens obtained from colon, endometrium, lung and melanoma. Thus, ABCG2 has a substantial involvement in the development of MDR tumors.

ABCG2 is believed to play a role in recurrent tumors. High ABCG2 expression was found to be correlated with disease relapse and poor prognosis in acute myeloid leukemia. Furthermore, ABCG2 expression was verified in premature stem cells of various different tissues. These findings suggest that ABCG2 might be involved in the protection of cancer stem cell subpopulations where it is believed to offer survival advantages, particularly under hypoxic conditions.

The wide spectrum of transport substrates and activity modulators

Molecules transported by ABC transporters are not, strictly speaking, substrates because they do not undergo a molecular transformation by the enzyme. However as the term is widely used I will keep it here. It should be noted that some substrates also behave as inhibitors, or induce the expression of MDR transporters. This is the case of HIV protease inhibitors. Several direct or indirect methods are used to study the drug transport: cell-based assays, vesicular transport, drug-stimulated ATPase activity, and fluorescent dye transport (Figure 31). Each of them has shortcomings, and a combination of approaches is often needed to unambiguously identify substrates and inhibitors [211]. The interactions of drugs obtained from such *in vitro* studies can be then validated *in vivo* (genetic knockout, gene deficient and chemical knockout animals) and the pharmacokinetic and pharmacodynamics properties can be estimated. However, it is often difficult to compare these data and then extrapolate them to a clinical context. As an example, inhibitory potencies (IC_{50}) cannot be compared between different studies, because these values depend on the cell system used, the substrate used, and also the assays conditions [212], thus the IC_{50} of tariquidar has a variation of 60-fold [213].

Since these methods are expensive and time-consuming, *in silico* methods for predicting substrates or inhibitors become more and more performant [214-216]. In the absence of structural data, the ligand-based approach is used, based on a quantitative structure-activity relationships, QSAR, which associate physical, chemical and structural properties of molecules with their biological activity to predict their efficacy. When structural data are available the structure-based method is preferred, such as molecular docking, to predict atomic details. However both have some restrictions: the first approach needs that the biological data comes from the same study and have a number of compounds high enough to gain in accuracy; the second method needs high resolution structures to allow a good prediction.

The following figure summarizes the properties of some *in vitro* assays with their advantages and disadvantages.

Cell-based assay systems

Use either cells stably (parental cell line and cell line transfected with the MDR-ABC transporter) or transiently overexpressing MDR-ABC transporter (parental cell line and cell sectioned using a known cytotoxic drug to overexpress a particular ABC transporter)

Advantages: suitable for the assessment of kinetic parameters of the transport and can be used in high-throughput mode

Disadvantages: varying expression levels of the proteins may greatly influence results. The cell overexpressing MDR-ABC transporter by drug selection may have inherent differences (relative to the parental cell line) that are not readily identified or controlled, making interpretation of results more difficult.

Although same cell lines, the results obtained in different laboratories may not be comparable to each other due to variability in the culture time, passage number, culture conditions as well as the differences in experimental set up.

Cytotoxicity/ Chemosensitivity assay/ drug resistance

Cell are incubated for X h with in increasing concentrations of the tested drug. Cell viability is quantified and the EC_{50} (half-maximal effective concentration) value may be determined.

Substrates: EC_{50} parental cell line < EC_{50} ABC transporter cell line.

Inhibitors: In present of a known substrates the EC_{50} parental cell line ~ EC_{50} ABC transporter cell line.

Disadvantages: the EC_{50} values are depending of the cell lines used.

Cellular accumulation/ Fluorescence dye transport

Cell are incubated for X min with in increasing concentrations of the tested compounds with intrinsic fluorescence or a radioactive label and their intracellular amounts is quantified (e.g. with a cytometer).

Substrates: ABC transporter cell line show lower accumulation that parental cell line.

Inhibitors: ABC transporter cell line show a similar accumulation that parental cell line. The IC_{50} (half-maximal inhibitory concentration) value may be determined.

Disadvantages: cellular sequestration, 'membrane leakage', intracellular binding, and dependence of fluorescence on intra-cellular milieu of the MDR-ABC substrates complicate quantification. Few compounds have intrinsic fluorescence or are available in radio-actively labeled form.

Transcellular ('vectorial') transport/ Monolayer transport assays

Cells are seeded on a membrane surface (e.g. in a transwell), and at 100% confluency, the test compound is added into the apical or the basolateral solution. The apical to basolateral (A–B) and basolateral to apical (B–A) permeability is determined, and a ratio of the two transport rates, reflect the function of the MDR-ABC transporter.

Substrate: The ratio $R_{B-A/A-B}$ >2 for P-gp or ABCG2 (apically localized) and reversed in case MRP1 (basolateral located).

Advantages: direct transport activity of the MDR-ABC transporter.

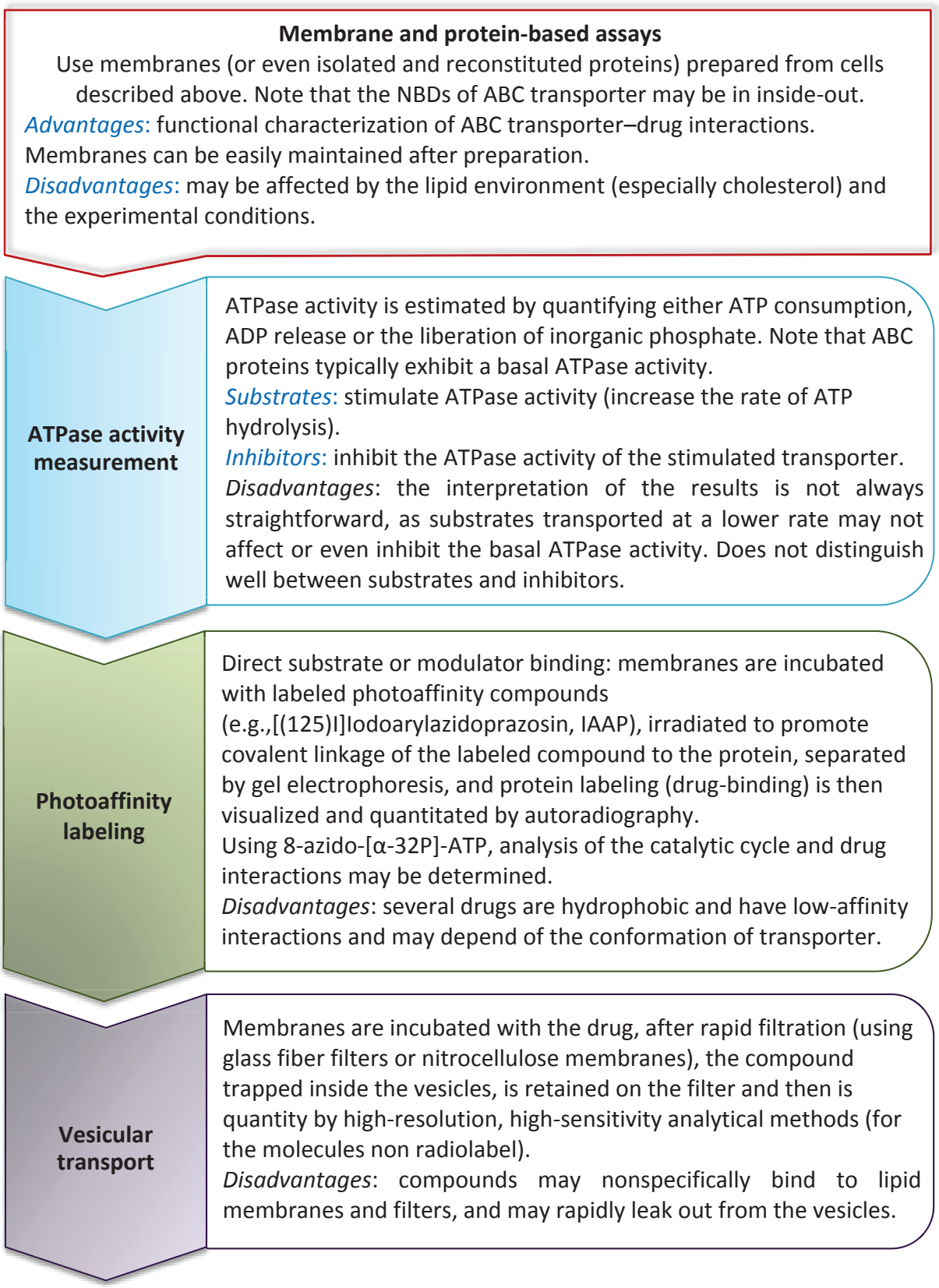


Figure 31. Description of some *in vitro* techniques used to study drug interactions with MDR-ABC transporters.
Adapted from: [217, 218]

All three MDR-ABC transporters display an overlapping of drug specificity and it is quite rare to find substrates or inhibitors specific to a single ABC transporter (Figure 32 and Figure 33).

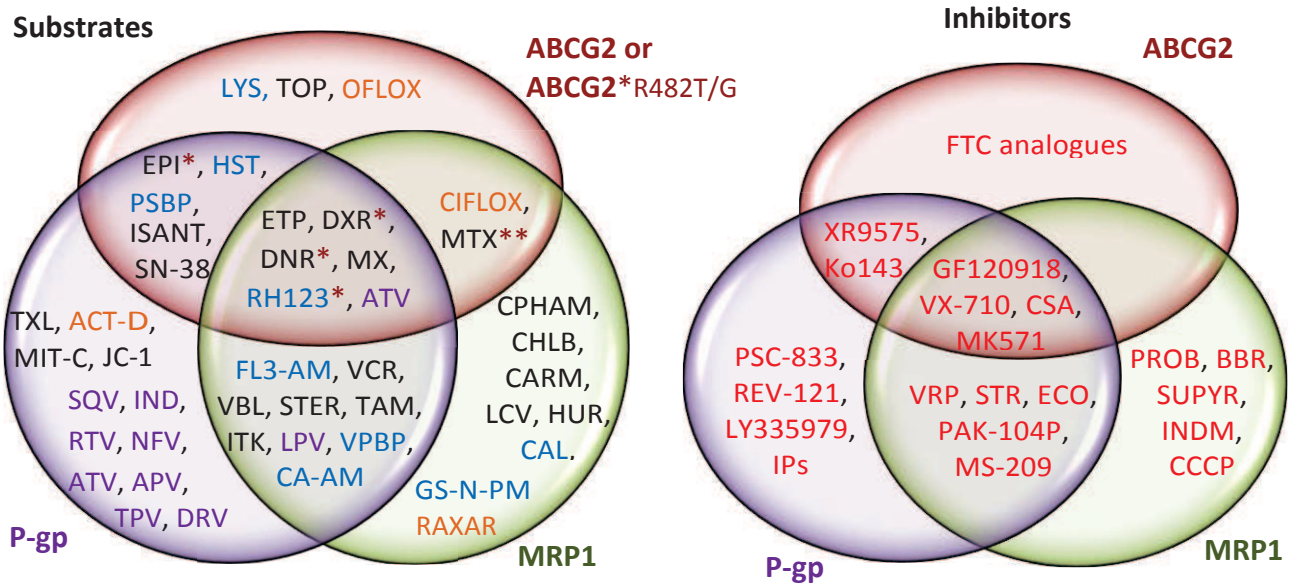


Figure 32. Venn-diagram for selected substrates and inhibitors of MDR-ABC transporters.

In black: anti-cancer drugs; In blue: fluorescence compound using for the detection of function; In orange: antibiotics; In violet: HIV protease inhibitors. In red: inhibitors.

*= molecules transported by the R482G and R482T protein variants form of ABCG2.

**= molecules transported by the wild-type ABCG2

VCR: vincristine; VBL: vinblastine; ETP: etoposide; STER: steroids; EPI: epirubicin; SN-38: Irinotecan; TAM: tamoxifen; ITK: tyrosine kinase inhibitors; DXR: doxorubicin; DNR: daunorubicin; MX: mitoxantrone; TOP: topotecan; BISANT: bisanthrone; COL: colchicine; MIT-C: mitomycin C; MTX: methotrexate; CPHAM: cyclophosphamide; CHLB: chlorambucil; CARM: carmustine; LCV: leucovorin; HUR: hydroxyurea; TXL: taxol;

CA-AM: calcein-AM; CAL: calcein; FL3-AM: fluo-3 AM; RH123: rhodamine 123; HST: Hoechst 33342; GS-N-PM: N-pyrenemaleimide glutathione conjugate; VPBP: BODIPY verapamil; PSBP: BODIPY prazosine; LYS: LysoTracker;

CIFLOX: ciprofloxacin; ACT-D: actinomycin D; RAXAR: grepafloxacin; OFLOX: ofloxacin;

SQV: saquinavir ; IND: indinavir; RTV: ritonavir; NFV: nelfinavir; LPV: lopinavir; ATV: atazanavir; APV: amprenavir; TPV: tipranavir; DRV: darunavir;

VRP: verapamil; STR: staurosporine; ECO: econazole; PAK-104P: pyridine analogue; MS-209: dofequidar fumarate; GF120918: elacridar; VX-710: biricodar; CSA: cyclosporin A ; PSC-833: valspodar; REV-121: reversin 121; LY335979: zosuquidar; IPs: HIV protease inhibitors; XR9575: tariquidar; FTC: fumitremorgin C; Ko143: fumitremorgin analogue; PROB: probenecide; BBR: benzbromarone; SUPYR: sulfinpyrazone; INDM: indomethacin; MK571: leukotriene receptor antagonist; CCCP: chlorocarbonyl cyanide phenylhydrazone

Adapted from: [16, 189, 203, 219]

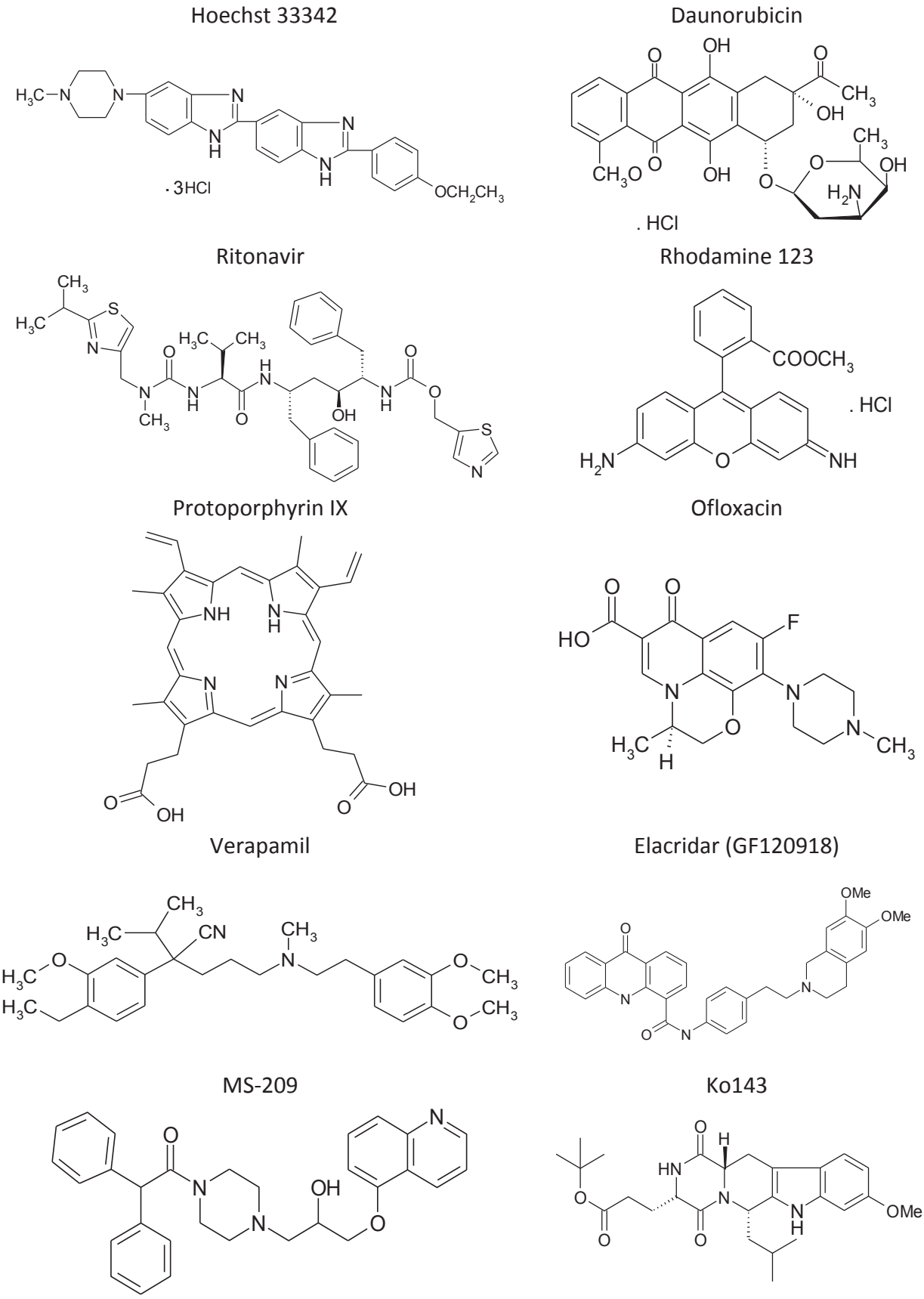


Figure 33. Structures of some molecules that are substrates or inhibitors of MDR-ABC transporters.

Substrate-binding sites (SBS): where and how many?

Until today, we still miss crystal structures of ABC exporters in complex with substrates. The best existing information comes from the mouse P-gp structure in complex with stereoisomers of a cyclic peptide inhibitor (QZ59s), which may represent a ligand-binding state of the protein. Several biochemical methods such as photoaffinity labelling, cysteine or arginine scanning, ATPase activity were used to localize the regions of P-gp interacting with drugs. The use of photoaffinity drug analogs has shown that substrate binding occurs principally in both halves of the TMD region of P-gp (reviewed in [189, 220]).

Two hypotheses have been proposed about drug binding: it occurs either in the same large pocket binding substrates and inhibitors [221] or through several specific sites that can interact [222-226]. Shapiro and Ling proposed the existence of three distinct binding sites (DBS) for the P-gp (Figure 34): termed H-, R- and P-sites. They proposed that these sites work cooperatively, meaning that the binding of a drug in one site promotes or inhibits the transport of another drug in a second site. The H-site binds Hoechst 33342 and quercetin, the R-site preferentially binds rhodamine 123 and anthracyclins such as daunorubicin and doxorubicin, while the P-site binds prazosine and progesterone [227, 228]. The authors observed that below 2 µM a drug binding to the R- or H- site activates the transport of drugs binding to the other site and above 2 µM competes with it. Using polarized monolayers of MDCK cells transfected with human MDR1, Tang *et al.*, found that rhodamine 123 stimulated Hoechst 33342 efflux but contrary to the observations of Shapiro and Ling, Hoechst 33342 inhibited rhodamine 123 efflux [229]. However, other studies referred to a positive or negative cooperatively between different drugs [230], corroborating the theory of Shapiro and colleagues. These conflicting observations may be due to the different environment of P-gp in each case, but they do indicate that the allosteric interactions in the P-gp drug binding pocket are likely much more complex than the original R-site and H-site model had suggested.

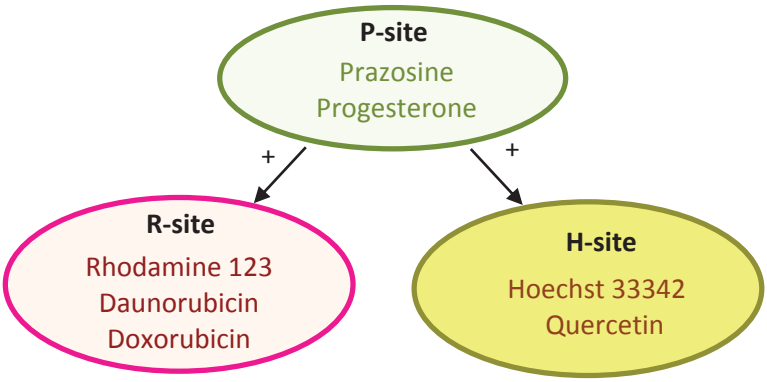


Figure 34. Diagram of the three binding sites on P-gp proposed by Shapiro and Ling. The binding of a drug to the P-site stimulates transport of an R or H- site drug ("+").

By classifying drugs that interact with P-gp either as substrates or modulators, Martin and coworkers developed a model of P-gp's transport, in which they postulate a minimum of four drug

binding sites (Figure 35). These binding sites switch between high and low affinity binding and belong to two categories of either transport or regulatory function [226].

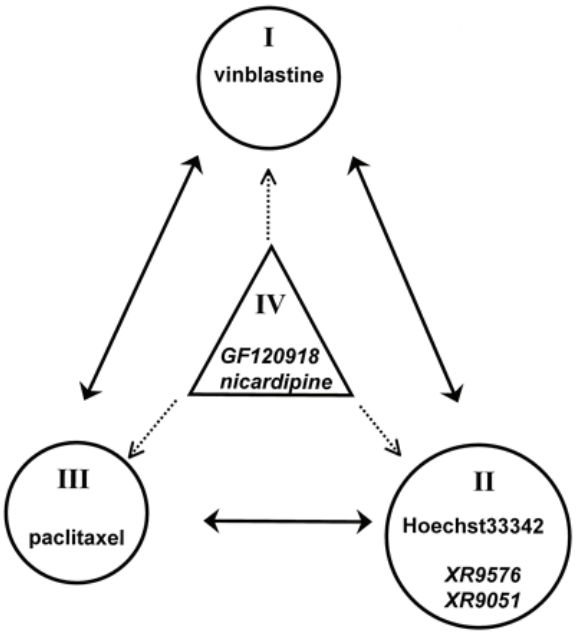


Figure 35. Classification of four drug binding sites on P-gp.
Arrows indicate communication between the four sites.
According to [226].

Another scenario has been proposed by Pajeva and co-workers, a single large pocket containing discrete sub-sites to accommodate different types of substrates [231]. Eckford and Sharom, propose that these sub-sites are formed by multiple hydrophilic electron donor and acceptor groups, charged groups, and aromatic amino acids [189]. The binding pocket is thought to be mobile, however, when binding of the substrate occurs, the binding pocket stabilizes. The R-, H- and P- sites could be different sub-sites of a single binding site for drugs.

Studies carried out by Loo and Clarke using a library of 252 single cysteine mutants of P-gp in the TMD region together with a thiol-reactive analog of a drug substrates, methanethio-sulfonate (MTS) verapamil and rhodamine, confirmed that several transmembrane segments (TMs) contribute to the same binding site (reviewed in [78]). The experiences of "cross-linking" of Cys residues show that the assembly of TMs of P-gp is altered when the substrate binds, and differently from one substrate to another. Thus a substrate can create its own binding site in the common drug-binding pocket by using a combination of residues from different TMs, this mechanism is called "induced fit" [232]. Residues I340 (TM6), A841 (TM9) and L975-V981-V982 (TM12), when replaced by a cysteine residue were indeed protected from the MTS-rhodamine modification by rhodamine B [233] making that these residues may participate or be close to the R-site. The position of the H-site is more controversial. Fluorescence resonance energy transfer experiments showed that Hoechst 33342 is in close proximity with the inner (cytoplasmic) leaflet [234], while the use of a pharmacophore pattern for the same dye suggested that F303, I306, Y307 and Y310 in the outer leaflet of TM5 may be involved in Hoechst 33342 binding [235].

For MRP1 and ABCG2 the DBS appear also to reside within the TM segments. The same strategy of site-directed mutagenesis and crosslinking with photoactivatable substrate analogs showed that TM10-11 and TM16-17 of MRP1 (corresponding to TM5-6 and TM11-12 of P-gp) are implicated in drug binding [189]. The DBSs of ABCG2 may function in a similar way to that of P-gp and Clark *et al.*, shown three distinct sites on each monomer ABCG2^{R482G} (broader substrate range) two of them are allosterically linked while the third does not seem to interact with the other two [236], Figure 36.

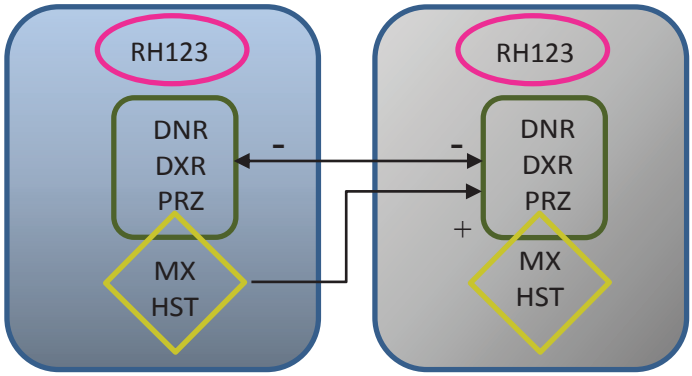


Figure 36. Diagram of the three binding sites on ABCG2^{R42G} proposed by Clarke *et al.*
One dimer of ABCG2 is in blue and the other in gray. There is 2 × 3 sites represented here.
RH123: rhodamine 123; DNR: daunorubicin; DXR: doxorubicin; PRZ: prazosine;
MX: mitoxantrone ; HST: Hoechst 33342; -: negative cooperatively; +: positive cooperatively

Ways to fight against MDR due to ABC transporters

Several approaches have been developed to restore the drug sensitivity of cells expressing MDR ABC pumps: (i) develop drugs (e.g., non-substrates, prodrugs) or use delivery systems (e.g., drug encapsulation) to avoid interaction with MDR-ABC transporters, called "Evade" [8]; (ii) "Exploit" features of multidrug-resistant cells (e.g., ''collateral sensitivity''); (iii) develops agents that interfere with either the function (e.g., inhibitors) or expression (e.g., inhibiting either the DNA transcription or translation of the mRNA) of transporter proteins, named "Engage". The later approach is the oldest one and the most studied. Here I'll briefly detail each approach (illustrated in Figure 37) with emphasis in the inhibitors development.

(i) Evade

Developing compounds that are not substrates for the MDR-ABC transporters is a challenging area for pharmaceutical companies due to the broad substrate spectrum of MDR-ABC transporters. In early stages of drug discovery, the new molecule is checked whether it shows an interaction with P-gp, ABCG2 or MRPs transporters or not [28, 30]. As example, the epothilones, novel agents targeting microtubules, with mechanism of action similar to paclitaxel but not recognized by P-gp [8]. Nevertheless, this does not guarantee that the drug is not transported because ABC transporters are able to adapt to a drug. So there is a chance that even if the drug

candidate is not exported during this test, it may be the case clinically with an adaptation of these pumps [217].

An alternative approach is to improve certain factors of the drug (e.g., more lipophilic, prodrugs) or its formulation (e.g., microspheres, nanoparticles and liposomes) in a way that the uptake rate exceeds the rate of efflux of ABC transporters, thereby efficient drug concentration in the cell is achieved. The amsacrine (antileukaemia drug) or N-[2-(dimethylamino)ethyl]acridine-4-carboxamide (experimental anticancer agent) are examples of lipophilic drugs that are thought to be substrates for ABC transporters but whose uptake rates are rapid [237]. Val-quinidine, a prodrug analogue of quinidine, was reported successful in evading P-gp transport [238]. Paclitaxel-ceramide, an apoptosis modulator with nanoparticle beads is used to increase bioavailability [239].

Recently other way to convert substrate into potent inhibitor has been reported by Chmielewski group, a monomer substrate drug of P-gp is converted into a dimer with potential inhibitor effect by occupying multiple binding sites within the transporter and once entry into cells would revert to their monomeric forms in the reducing environment of the cytosol, thus, delivering the therapy [240]. Until now, it has been applied for a molecule in the treatment of HIV, the Abacavir (reverse transcriptase inhibitor), but this strategy could be applied for any substrate drug not only of P-gp but also ABCG2 or MRP1.

(ii) Exploit

An alternative strategy to circumvent and exploit the overexpression of ABC transporter in cancer cell could be use compounds that kill selectively cells that express these specific transporters, a phenomenon termed collateral sensitivity (CS) (review in [241, 242]). It seem that the pathway that resistant cells develop to resist towards one agent, in connection to the activity of overexpressed ABC transporters, is accompanied by the development of hypersensitivity towards a second agent, creating an “Achilles’ heel”. This strategy has been particularly studied for the cells overexpressing P-gp and MRP1. Thus, verapamil, one of the first P-gp inhibitors, was one of the first compounds to show high collateral sensitivity towards P-gp-overexpressing Chinese hamster ovary cells. This hypersensitivity is due to a futile cycle of transport carrying heavy consumption of ATP by P-gp. The cell is then emptied of ATP necessary for the physiology function, establishing an intense ATP synthesis which gives the formation of a significant amount of reactive oxygen species (ROS) promoting signaling pathway apoptosis [241]. Verapamil also causes CS against cell overexpressing MRP1, but with another mechanism, since it stimulates glutathione (GSH) efflux which precludes cell response to ROS generation and induce caspase-dependent apoptosis [243]. Surprisingly, only the S-verapamil is responsible for the death of cells overexpressing MRP1 while the R-verapamil does not initiate GSH efflux or cell death, but rather inhibits MRP1 [244]. Energetic sensitivities and perturbation in the plasma membrane have been also proposed to explain CS, besides the already mentioned (ROS generation and excretion of essential endogenous substrates). But many mechanisms remain unknown and it is very likely that this induced death may be through several pathways that differ from one compound to another.

(iii) Engage

Ideally, the simplest and most direct way to address the MDR phenotype is blocking (either the drug binding or efflux, the NBD or ATP activity, or some residues involve in transport) ABC transporters by inhibitors, also known as modulators, chemosensitizers or reversal agents, depending on their type of action (irreversible, reversible, competitive or noncompetitive) [245]. So, they are co-administrated with the substrate drug allowing its accumulation in cells. An ideal inhibitor is an inhibitor having the following characteristics: be effective at low concentration, be specific to the ABC transporter, having no pharmacokinetic interaction with the co-administrated drug, and be not toxic. The last point is a little bit tricky, because as described above the ABC transporters have essential physiological roles and their inhibition in normal tissues leaves unprotected the healthy cells, thereby increasing drug toxicity. However specific inhibitors for a single transporter can overcome this obstacle maintain the protection through other ABC transporters.

Many teams have worked in this direction since 1980 and hundreds of compounds have been identified and synthetized. They may divided into three generations (review in [246, 247]): the first generation of inhibitors (e.g., verapamil, cyclosporine A, tamoxifen, and several calmodulin antagonists) are often substrates and at high concentration compete with the drug for the efflux. In addition they showed interaction with other transporters and enzyme systems, resulting in unpredictable pharmacokinetic and side effects unacceptable. The second (e.g., dexverapamil, valspodar, and biricodar) are commonly less toxic and more potent than their predecessors, but they inhibit the activity of the cytochrome P450 altering the pharmacokinetics of the anticancer drug, limiting the use of these agents. The third generation inhibitors (e.g., zosuquidar, tariquidar, elacridar,) was developed using structure-activity relationship and combinatorial chemistry approaches, they overcome the early disadvantages but clinical data of these agents are disappointing due to the limited clinical benefit (Table 7). Finally other inhibitors have been recently obtained by diverse strategies: compounds extracted from natural origins and their derivatives; peptides or antibodies and dual activity agents; surfactants and lipids.

P-gp inhibitor	Common Name	Types(s) of cancer	Clinical benefit	ClinicalTrials.gov Identifier(s)
PSC 833	Valspodar	AML	No	NCT00004217; NCT00005823
XR 9576	Tariquidar	Solid tumors	Limited	NCT00020514; NCT00069160
LY 335979	Zosuquidar	AML	No	NCT00046930
R 101933	Laniquidar	Breast cancer	Not known	NCT00028873
MS 209	Dofequidar	Solid tumors	Not known	NCT00004886
BMS 217380-01; DPPE	Tesmilifene	Breast cancer	Limited	NCT00364754
CBT-1	-	Solid tumors	Limited	NCT00972205

Table 7. Related clinical trials of some P-gp inhibitors.
According to [248].

Possible reasons that explains the negative results of clinical trials have been mainly attributed to the intolerable side effects, associated to the increase of anticancer drug and or inhibitor toxicity. This increase is due to the non-specific inhibition of P-gp, interfering with anticancer drug excretion. In the order to improve this strategy we should: (i) develop an *in vitro* diagnostic test that identified that MDR-ABC transporters are the main contributors to the resistance; (ii) make sure that the patient don't have SNPs, because it has been show that they affect the binding of the substrate or the inhibitor on the transporter [249]; and (iii) develop a better pre-clinical models to study the pharmacokinetic of the inhibitor alone and in combination with the possible therapy agent.

Another possible use for inhibitors of ABC drug transporters could be like coenhancers or booster drugs, increasing the oral bioavailability and plasma concentrations limited by the poorly absorbed drugs. This approach has been used against AIDS giving ritonavir at low dose as a pharmacologic booster to others protease inhibitors by inhibiting P-gp and or cytochrome P450 [250].

Other approaches that have developed including modification at the plasma membrane to prevent P-gp overexpression (e.g., fatty acid-polyethylene glycol fatty acid diesters) and regulation of the expression of ABC transporters by pointing either the transcription [251]: using phosphorothioate-modified antisense oligonucleotides (target MED-1 sequence); LANCL2 gene (suppress the activity of the P-gp promoter); K2-5F repressor (against the P-gp promoter SP1/EGR1/WT1), etc. Or the translation: using antisense oligonucleotides (binds the complementary mRNA forming a duplex and thus prevent its translation); hammerhead ribozymes (hybridize to a particular sequence and catalyze a chemical reaction at this position), and siRNA (small RNAs complementary to their target mRNAs). Nevertheless clinical studies using these alternative approaches are still lacking to establish their efficacy in cancer patients.

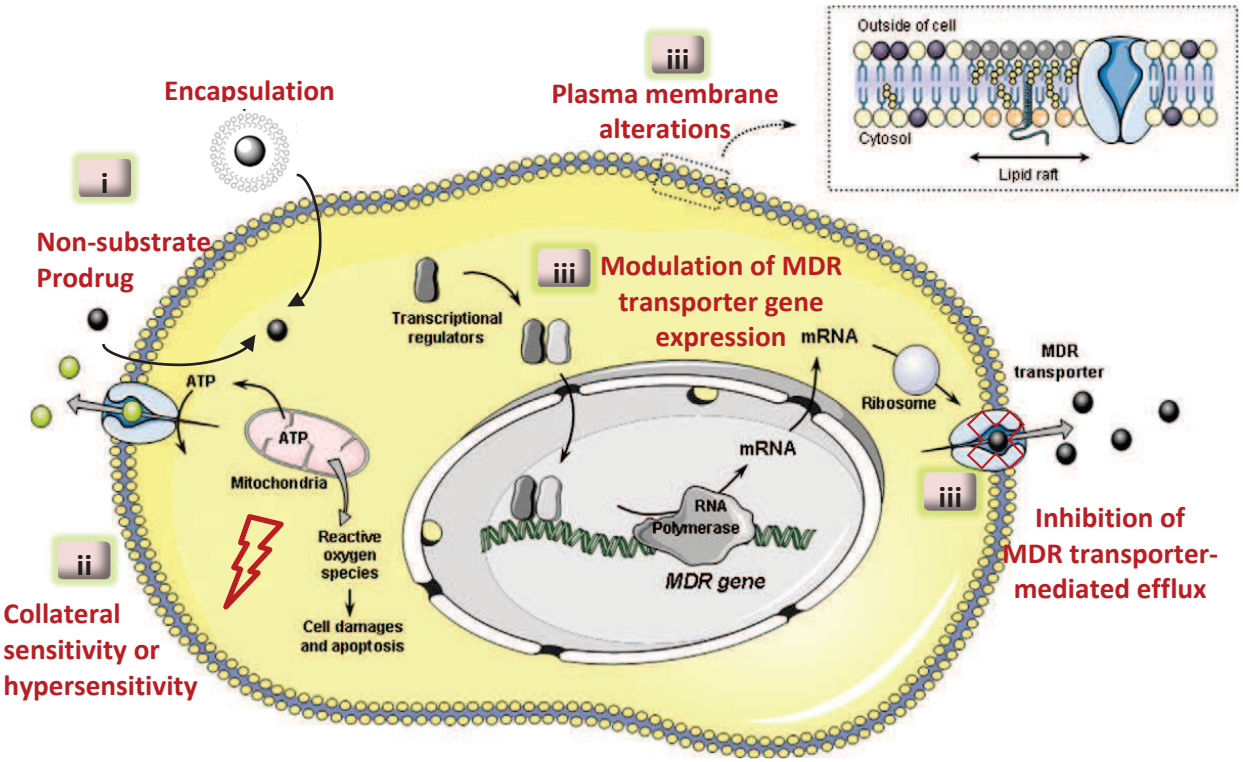


Figure 37. Schematic methods used to avoid MDR mediated by ABC transporters.
(i) Evade (Non-substrate, prodrugs or encapsulation);
(ii) Exploit (collateral sensitivity or hypersensitivity);
(iii) Engage (Inhibitors, gene silencing of ABC transporters and alteration in the plasma membrane).
Adapted from [252]

MATERIALS AND METHODS

This section presents first, the materials (products, cells) and methods used (flow cytometry, monitor HIV-1 protease activity and functional transport assays). Then the key theoretical foundations (in crystallography, docking, enzyme kinetics and statistical analysis) of the present work.

MATERIALS

Compounds studied

Product	Supplier	Reference
indinavir	Merck	noncommercial
doxycycline	Fluka -Sigma Aldrich	44577
quaternary hemiaminal-based HIV protease inhibitor	Université de Lyon – ENS (noncommercial)	#36 to #43
GF120918 (elacridar)	GlaxoSmithKline	noncommercial
Nb592		noncommercial
QZ59-RRR and QZ59-SSS cyclic hexapeptides	Tao *et al.*[253]	noncommercial
Hoechst 33342	Fluka- Sigma Aldrich	14533
Mitoxantrone	Sigma Aldrich	M6545
Daunorubicin		D8809
Doxorubicin		D1515
Rhodamine		R8004

Table 8. List of products with their respective supplier and reference.

Routine cell culture

Cells were handle in a class II laminar flow hood (Safeflow1.2, Bio Air Instruments, France) under aseptic conditions. Cells were grown by serial passages in a cell growth medium (CGM), in flasks of 75 cm^2 cell growth area and maintained in a humidified incubator at 37°C with 5 % CO_2 in air.

Human T-cell line SupT1 ("The human T-cell lymphoblastic lymphoma cell line") were transfected with Gal4 (Figure 38A) or PR/Gal4 (Figure 38B): HIV-1 wild-type protease fused within the Gal4 DNA-binding and transactivation domains, kindly provided by Dr. R. Wolkowicz (SDSU, USA) [254]. CGM: RPMI 1640 media (PAA laboratories) supplemented with 10 % fetal bovine serum (PAA laboratories), and 1 % v/v of mixture of penicillin/streptomycin (PAA Laboratories).

NIH3T3 cell line (Figure 38C) was established from NIH Swiss mouse embryo cultures, the 3T3 refers to the cell transfer and inoculation protocol meaning "3-day transfer, inoculum 3 x 10^5 cells". NIH3T3-P-gp drug-resistant cell line (Figure 38D) was co-transfected by calcium phosphate precipitation with the *mdr1* gene with a glycine at position 185, contained in the vector pHaMDR1/A, then selected by 60 ng/ml colchicine (Sigma-Aldrich®) [255]. CGM: Dulbecco's modified Eagle's medium (PAA laboratories), 10 % fetal bovine serum (PAA laboratories) and 1 % v/v of mixture of penicillin/streptomycin (PAA Laboratories).

Human embryonic kidney (HEK293) cell lines were from the ATCC. HEK293-pcDNA3 (Figure 38E) and HEK293-ABCG2 (Figure 38F) were co-transfected by Sira Macalou (PhD in the laboratory) using nucleofection method with the empty pcDNA3.1 (Invitrogen) vector and with the same pcDNA3.1 vector containing the gene encoding the wild-type strain of ABCG2 (arginine at position 482) respectively. Then cell were selected by 0.75 mg/ml of geneticin G418 (PAA laboratories). CGM: Dulbecco's modified Eagle's medium (PAA laboratories), 10 % fetal bovine serum (PAA laboratories) and 1 % v/v of mixture of penicillin/streptomycin (PAA Laboratories).

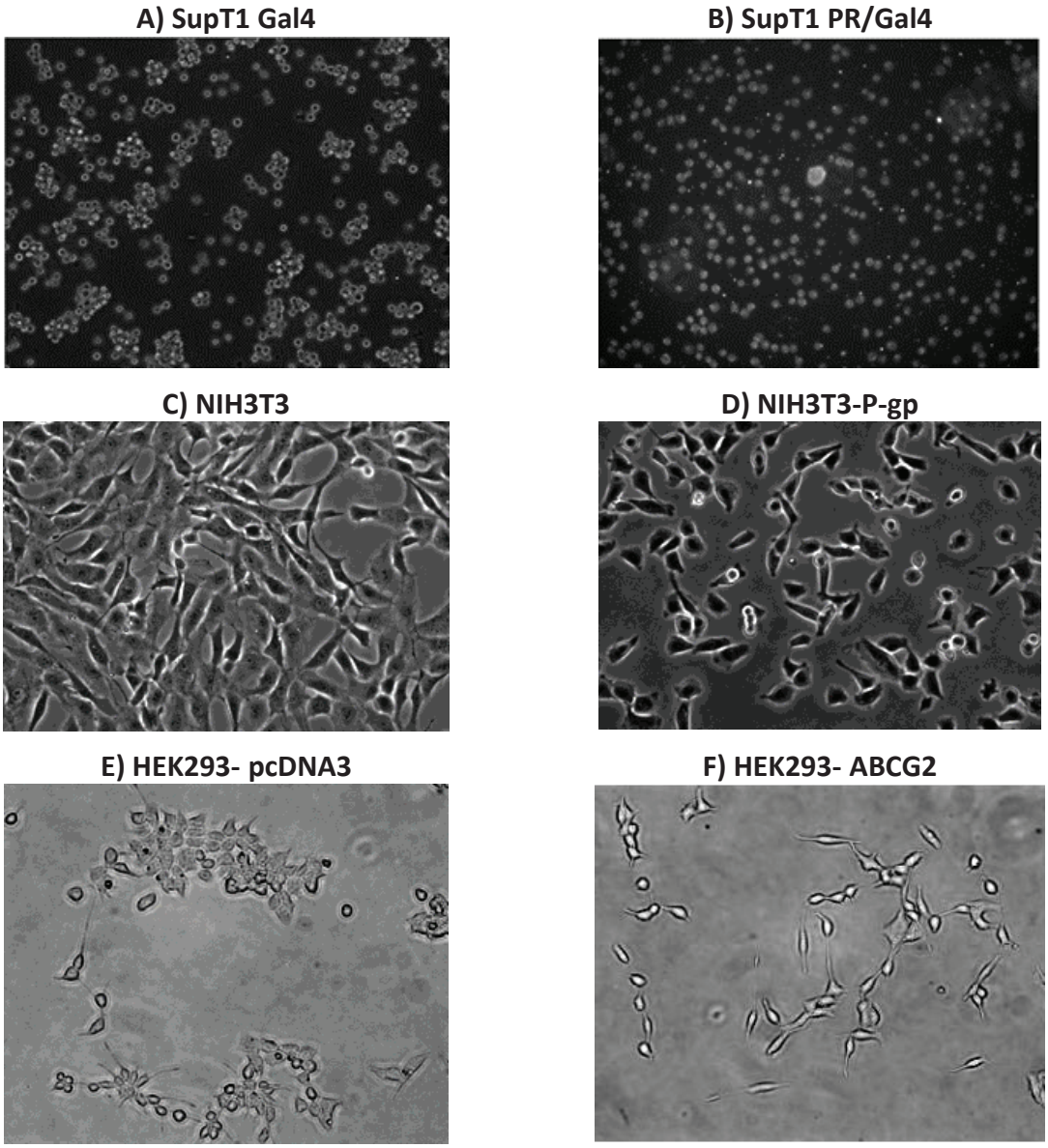


Figure 38. Pictures of the cell lines.
Optical microscope in contrast phase, magnification 20 ×.
Pictures C-F were taken by Arnaud Ophélie (PhD in the laboratory)

For the test that requires a number of cell per well, the cell counter (CellCounter) was used to count the cells automatically. For this, the cells were put in the presence of trypan blue (Sigma Aldrich) which is not excluded in dead cells, thus it is possible to differentiate dead cells (blue) and living cells (translucent).

METHODS

Flow cytometry

Is a powerful method to measure multiple properties of single cells or particles within heterogeneous populations through a fluid in front of a laser beam. With a maximal speed of about 10 m/s and hydrodynamic focusing (the fluid force cells to pass through a small tunnel), the cells (typically between 0.2-50 micrometers) pass one by one and refract or scatter light at all angles. Light scatter (which is proportional to the cell size) is quantified by a detector (mostly photomultiplier) that converts intensity into voltage and three parameters may be collected:

- Forward scatter (FSC) or low-angle light scatter: the photodiode is placed in line with the light beam and gives information about the size of cells.

- Side scatter (SSC) or larger-angles light scatter: mirror is placed perpendicularly (90 degrees from the laser's path) to light beam and gives information about granularity and structural complexity inside the cell.

- Fluorescent light (FL): the emission pass through the same path as the side scatter signal and it is directed through a series of filters and mirrors, so that particular wavelength ranges are delivered to the appropriate detector or photomultipliers tube (PMT).

A flow cytometer consists in the following subsystems: **Fluidic subsystem:** which brings the sample to the laser beam and takes away the waste. **Optical subsystem:** that provides the light source for scatter and fluorescence (Laser λ: 488, 355 nm, etc.), dichroic mirrors (incidence ray is at 45° and wavelengths are reflected) and filters (absorb specific wavelength and transmit others) to select the wavelength; for example a Long pass (LP) filter of 525 (LP525) mean that starting wavelengths of 525 will pass. **Electronic and computer subsystems:** that convert the detected light in electronic signal, further processed and sent to the computer, (Figure 39).

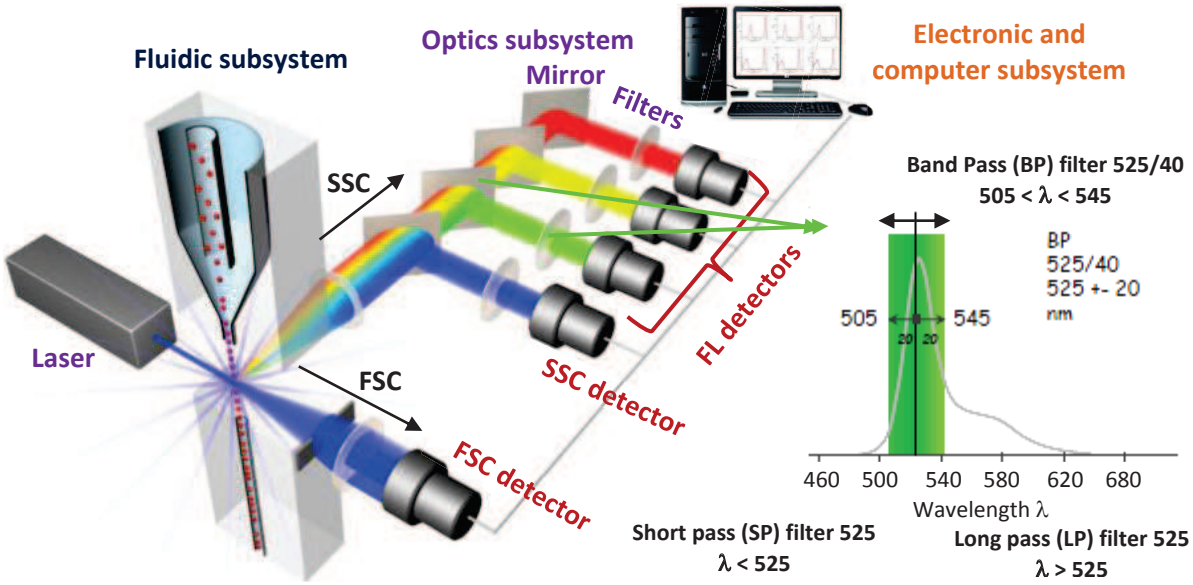


Figure 39. Flow cytometer subsystem and parameters measured by it.
Adapted from: http://probes.invitrogen.com/resources/education/tutorials/4Intro_Flow/player.html

Two cytometers were used: FACS Calibur cytometer or FACS LSR II from BD Biosciences. Data were collected with CellQuest Pro 4.0 or FACSDiva 6.1.2 softwares and then exported to FlowJo (TreeStar) for analysis. First step: choose the population of interest using density plot (SSC versus FSC) and Gates to define subgroups and restrict the analysis on interesting cells signals (Figure 40).

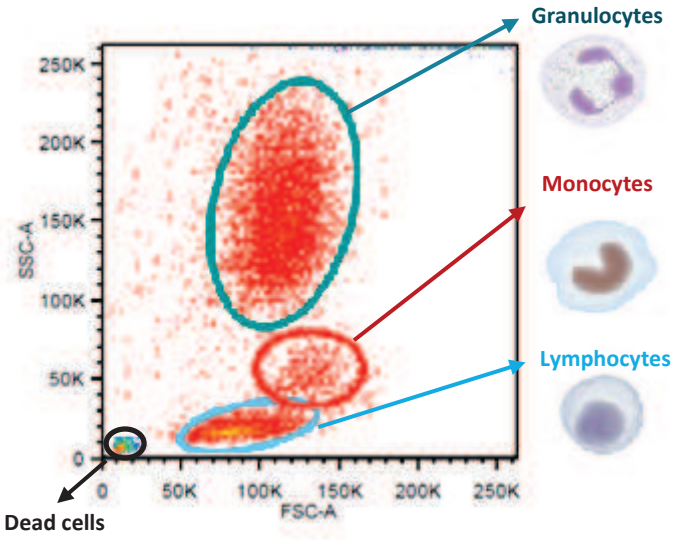


Figure 40. Choose the population of interest.

Second step: remove doublets, when two cells go along to the laser, the height of the electrical impulse of their size (FSC-H) or their granularity (SSC-H) will be the same but the width of the pulse (FSC-W and SSC-W) will be doubled (Figure 41A). About 90 % of doublets are eliminate by the first plot (Figure 41B) and the second allows to eliminate about 5 % of the remaining pairs (Figure 41C).

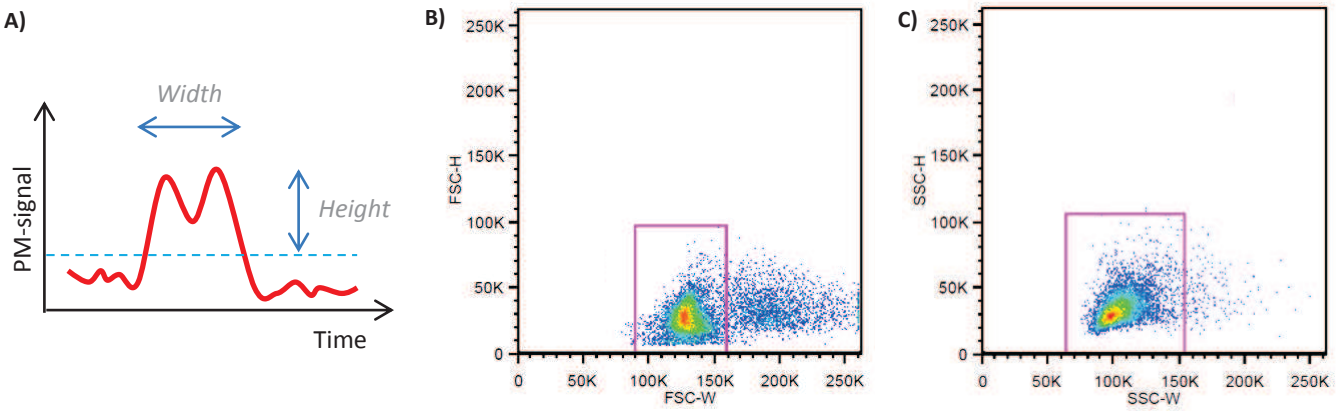


Figure 41. Doublet elimination.
A) Representation of the electrical impulses. B) Elimination through Forward SCattering (=size).
C) Elimination through Side SCattering (=granulosity).

Once these doublets eliminated, the intensity of intracellular fluorescence is measured on the last subpopulation and represented by the number of cells (Figure 42). The next methods use the flow cytometry.

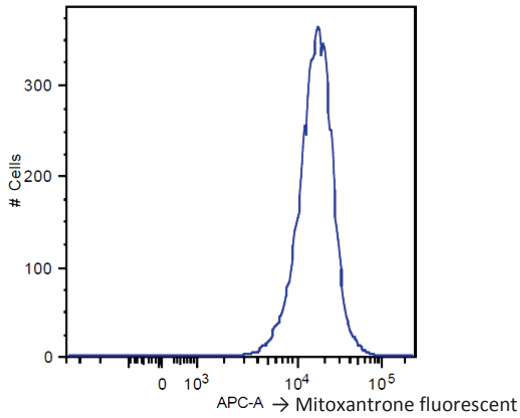


Figure 42. Representation of the number of cells as a function of the mitoxantrone fluorescence intensity.

HIV-1 protease activity monitoring

This assay monitors the catalytic activity of HIV-1 wild-type protease (PR) in T-cells, natural targets of HIV, using the prototypic transcription factor Gal4, which consists of the N-terminal DNA-binding domain (DBD) and the C-terminal trans-activation domain (TAD) (detailed protocol in [254, 256]). The assay is based upon (1) introduction of PR in between the two Gal4 domains, to obtain a PR/Gal4 fusion protein, (2) utilization of the enhanced Green Fluorescent Protein, acting as a biosensor of HIV-1 PR activity and (3) all under the control of a tetracycline- or doxycycline-inducible promoter. The DBD of Gal4 recognizes and binds to the Gal4 responsive upstream activation sequences (UAS), which allows the TAD to activate transcription of the downstream reporter gene, the enhanced Green Fluorescent Protein (eGFP). The DBD and TAD must act in conjunction, and neither domain can act independently as a transcription factor. Thus, in the presence of doxycycline

(Doxy) and a "bad" inhibitor, PR/Gal4 is expressed; then PR has autocatalytic activity and cleaves itself from the Gal4 domains, resulting in the inability of the fusion protein to induce GFP expression. In contrast, in present of Doxy and a potential inhibitor, the PR activity is blocked and the two domains of Gal4 remain intact, resulting in GFP expression, the more the HIV-1 protease activity is inhibited, the more GFP is expressed. The assay is illustrated in Figure 43.

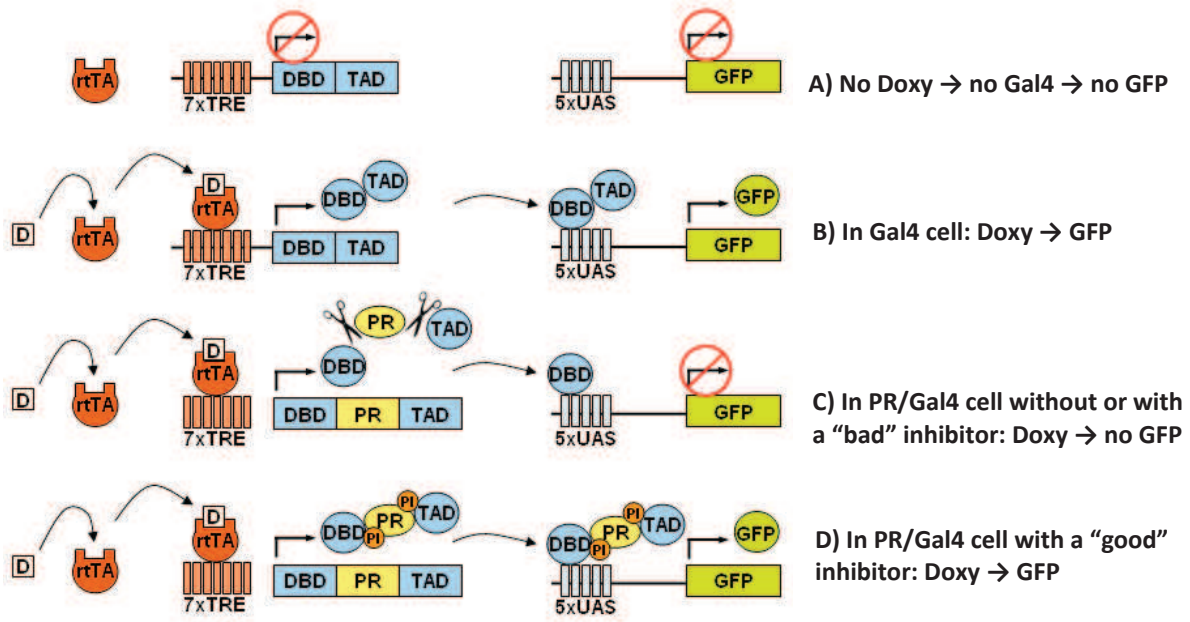


Figure 43. Biosensor of HIV-1 PR activity.
D or Doxy: doxycycline; rtTA: reverse tetracycline transactivator; TRE: tet-responsive element; DBD: N-terminal DNA-binding domain; TAD: C-terminal trans- activation domain; UAS: upstream activation sequence; PR = HIV-1 wild-type protease; GFP: green fluorescent protein; PI: protease inhibitor
According to [254, 256].

The protocol is as follows: In a 96 well plate, 50,000 cells per well (both control, Gal4, and tested, PR/Gal4, cell line) are incubated for 10 min in growth medium condition and then inhibitors were added at varying concentrations and incubated for at least 10 min. Cells are activates with 1 µg/ml doxycycline. After 48 h, cells are washed with Phosphate Buffer Saline (PBS, PAA laboratories), and the expression of GFP is quantified by flow cytometry carried out with a FACS Calibur cytometer. Excitation and emission are set up to 488 and 530 nm, respectively. Each test series contains (Figure 44): (i) a control (Gal4) with and without Doxy to verify the expression of GFP; (ii) a reference inhibitor (indinavir) as positive control for PR/Gal4 cell and to check there is no interference in the witness (Gal4).

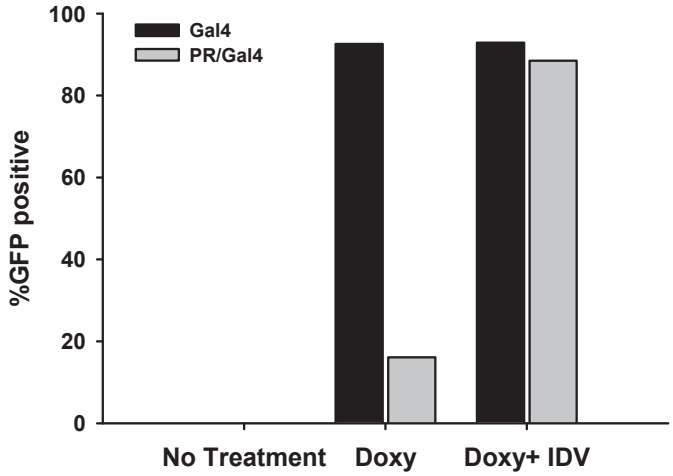


Figure 44. HIV-1 PR activity assay controls.
IDV: indinavir at 10 µM. Doxy: doxycycline at 1 µg/ml.

One disadvantage of this experiment is that hydrophobic compounds will have a poor cell penetration. So they will not reach the HIV-1 protease, which is located in the cytoplasm, what appear to be a lack of efficacy of the inhibitor (false negative).

Drug transport assays

As described below (Figure 45), the activity of efflux pumps can be measured by flow cytometry. For this, the substrate must be fluorescent. If the substrate is expelled, the intracellular fluorescence lowers. While if it is not transported the intracellular fluorescence remain high (Figure 45).

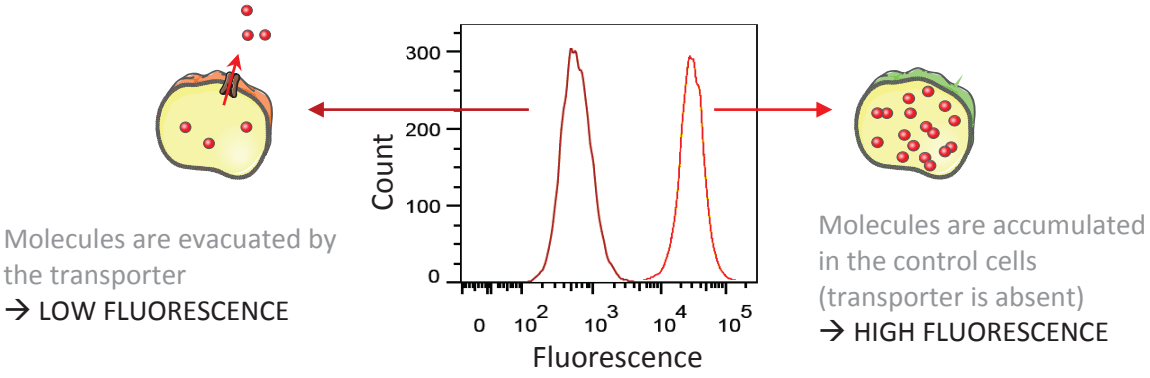


Figure 45. Indirect measure of efflux pumps activities.

The protocol is as follows: In a 24 wells plate, 50,000 cells per well (NIH3T3 or NIH3T3-P-gp) or 10 times more (HEK293-pcDNA3 or HEK293-ABCG2) are seeded for 24 h. After removing the medium, 100 µL complete medium ± inhibitor plus 100 µL complete medium ± substrate are added to cell and 30 min later cells are washed with PBS, trypsinised, and their intracellular amounts is

quantified with the correct cytometer. Each test series contains (i) a background noise due to the auto-fluorescence of cells, corresponding to both cell without substrate and in the case of inhibitory activity, with the inhibitor to overcome the potential intrinsic fluorescence of the inhibitor; (ii) a control of maximum accumulation, corresponding to the control cells with the substrate (100 % of accumulation); (iii) the test, which overexpressed cell are incubated with the substrate ± inhibitor (Figure 46).

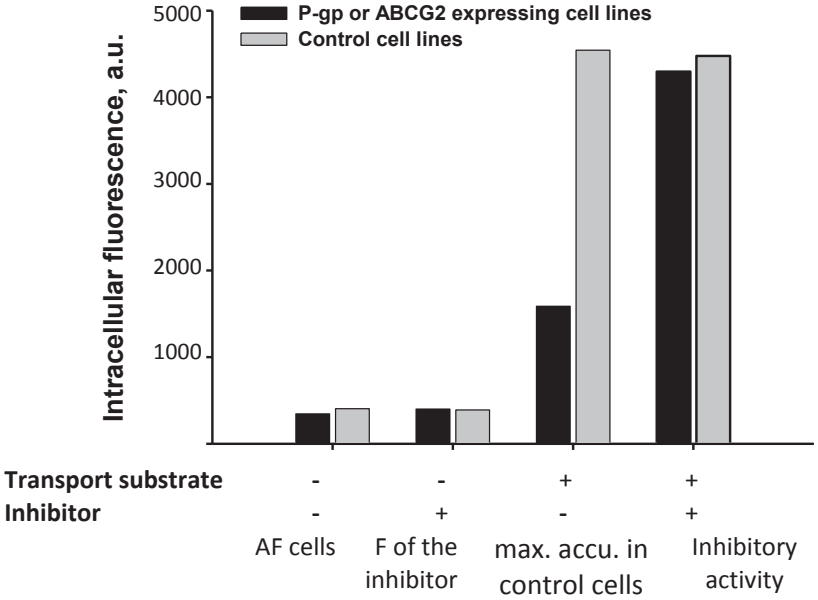


Figure 46. Efflux pumps activity assay controls.
AF: auto-fluorescence; F: fluorescence; max. accu.: maximum accumulation

The percentage of accumulation or inhibition is calculated as follows:

$$accumulaton, \% = \frac{FLA-FLB}{FLC-FLB} x100 \qquad inhibition, \% = \frac{(FLA_i-FLA)}{(FLC_i-FLA)} x100 \quad [257]$$

Where, FLA: fluorescence in P-gp or ABCG2 expressing cell lines with the transport substrate; FLB: auto-fluorescence of cells or background; FLC: fluorescence in control cell lines with the transport substrate; FLA_i: fluorescence in P-gp or ABCG2 expressing cell lines in the present of the transport substrate with the inhibitor; FLC_i: fluorescence in control cell lines in the present of the transport substrate with the inhibitor.

Antiproliferative assay by MTT

The objective of this experiment is to determine the effective concentration (EC_{50}) define as the half-maximal concentration of a given molecule that inhibits 50 % of cell proliferation. For that in a 96 well plate, 2,500 cells of NIH3T3 or NIH3T3-P-gp or 4 times more of HEK293-pcDNA3 or HEK293-ABCG2 are seeded per well for 24 h. One hundred microliters of increasing concentrations of the tested compound is added and incubated for 72 h. Cell viability was evaluated using an MTT colorimetric assay [258] as follows: at the end of the incubation time 0.5 mg/ml of MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) is added and incubated for 4 h. Then the

yellow MTT is reduced by mitochondrial succinate dehydrogenase in living cells to formazan blue-violet color (Figure 47).

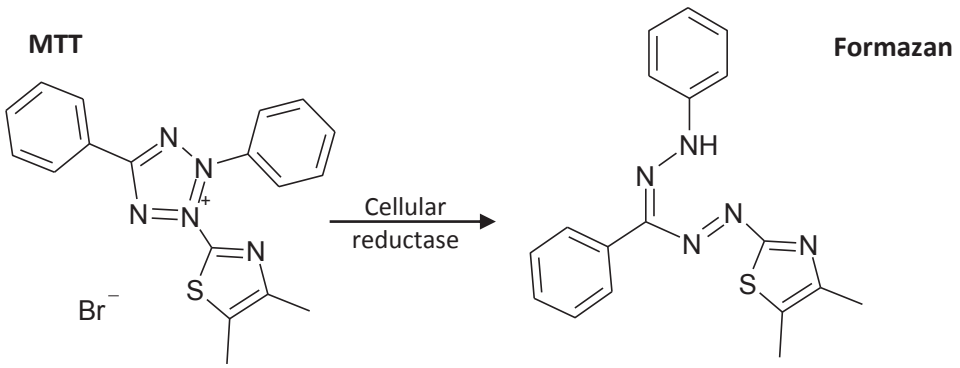


Figure 47. Reduction of MTT to formazan.

Then the medium is remove and formazan crystals are solubilized in 100 µL of 50:50 solution of dimethyl sulfoxide (DMSO) and ethanol, and then incubated for 10 min. The intensity of the resulting color is proportional to the number of living cells. The absorbance is measured at 570 nm (or optical density at 570 nm, OD_{570nm}) using a spectrophotometric plate reader. The percentage of cell survival and each well is determined by the following calculation:

$$survival\ cell, \% = \frac{A_x}{A_o} x100$$

Where, A_x: is the OD_{570nm} at concentration *x* of the test compound; A_0: is the average OD_{570nm} of the cells in the absence of the investigated compound.

Workflow of crystallographic study of P-glycoprotein

Gene-optimized of mouse P-gp

Opti-P-gp was obtained by removing the three N-linked glycosylation sites, Asn-83, Asn-87, Asn-90 which were all mutated to glutamine, adding a hexahistidine tag at the C-terminus to facilitate purification and optimizing the codon usage for highly expressed genes in *P. pastoris* cell line GS115 [259].

Growth of P. pastoris cells in the fermenter

A colony or glycerol stock is used to inoculate 10 ml of medium growth yeast (MGY: 1.34 % w/v yeast nitrogen base without amino acid (Difco), 1 % v/v glycerol, 0.4 mg/L Biotin) and grown up to an optical density (OD) of 2-4 in a shaker for 24 hours at 28 °C. One milliliter of this culture is then used to inoculate 500 lL of MGY in a 1 L flask for 24 hours (grow up to OD 2-4 in 28 °C shaker). Then this 500 ml was inoculated into 10 L of culture medium (MGY, 11.2 g/L K_2SO_4, 7.25 g/L $MgSO_4 \cdot 7H_2O$, 3.75 g/L KOH, 0.675 g/L $CaSO_4$, 4 % v/v glycerol, green metals for Invitrogen and antifoam) in a bioreactor (Bioflow 415; New BrunswickScientific) at 28 °C and pH 3.5 was maintained using 14 %

w/v ammonium hydroxide. Fermentation proceeded in two phases. The first one was 24 hours after inoculation, the "glycerol-fed phase", in which cells were fed with 300 ml of glycerol. The second phase was the "methanol-fed phase", in which protein expression was induced by addition of methanol (3.6 ml/h per liter of culture volume) overnight. Cells were centrifuged at 3,063 x g, 10 min, 4 ˚C and stored at - 80 °C. The yield of production is about 2.5 kg / fermenter.

Membrane preparation

Cells were suspended 1:1 mass per volume in lysis buffer (200 mM Tris-Cl pH 8.0, 50 mM NaCl, 15 % v/v glycerol, plus Sigma protease inhibitors cocktail) for 1 hour. The suspension was passed a single time through a cell disrupter (TS-Series; Constant Systems) at 40,000 psi. The cell wall and debris were removed by centrifugation (3,500 × g, 35 min, 4 °C), and crude plasma membranes were obtained by centrifugation at 35,267 × g for 2–3 h at 4 °C, dissolved 1:1 in lysis buffer (~ 150 g) and stored at - 80 °C.

Detergent extraction of P-gp

50 g of membranes containing P-gp were resuspended in 200 ml of solubilization buffer (100 mM NaCl, 15 % glycerol, 20 mM Tris-HCl at pH 8.0, 9 % v/v Triton X-100 and Sigma protease inhibitors cocktail) and solubilized at least 2 hours on the stir plate at 4 ^{o}C. Insoluble material was removed by centrifugation at 38,400 × g, 4 °C for 30–60 min.

Ni^{2+}-NTA column chromatography

The supernatant from the solubilization was poured at 7 ml/min (keeping the pressure below 0.7 MPa) over a metal resin (Ni-NTA Superflow; Qiagen) equilibrated with buffer 1 (100 mM NaCl, 15 % glycerol, 20 mM Tris-HCl at pH 8.0, 4.5 % v/v Triton X-100, 0.5 mM tris(2-carboxyethyl)phosphine (TCEP), 0.04 % sodium cholate, 20 mM imidazole, and Sigma protease inhibitors cocktail) in FPLC (AKTA; GE Life Sciences). Immobilized P-gp was washed two times, the first with 250 ml of buffer 1 and the second with 300 ml of buffer 2 (100 mM NaCl, 20 mM Hepes pH 8.0, 0.2 mM TCEP, 0.04 % sodium cholate, 20 mM imidazole, and 0.0675% β-dodecyl maltoside (βDDM)). Protein was eluted with 100 ml of elution buffer (100 mM NaCl, 20 mM Hepes pH 7.5, 0.2 mM TCEP, 0.04 % sodium cholate, 200 mM imidazole, and 0.0675 % βDDM).

A second Ni^{2+}-NTA column was carry out to improve P-gp homogeneity. Thus, the eluted protein was diluted 1:10 in dilution buffer (100 mM NaCl, 20 mM Hepes pH 7.5, 0.2 mM TCEP, 0.04 % sodium cholate, 0 mM imidazole, and 0.0675 % βDDM) and bound to a new Ni-NTA column equilibrated with buffer 2. The column was washed with 300 ml of buffer 2 and eluted with 100 ml of elution buffer.

Gel filtration chromatography

After elution, the protein was concentrated (Centricon YM-50 or YM-100; Millipore), spun at 32,0424 × g, 4 °C for 30–60 min, and subjected to gel filtration (GF) chromatography (Superdex200

16/60; GE Healthcare) at 1 ml/min, 4 °C. The GF column was equilibrated with a buffer containing 100 mM NaCl, 20 mM Hepes pH 7.5, 0.2 mM TCEP, 0.01 % sodium cholate and 0.035 % βDDM.

Crystallization

After gel filtration P-gp is subjected to reductive methylation. Briefly, freshly made borane and formaldehyde were added to 10 ml of protein solution at final concentrations of 50 mM and 100 mM, respectively, and incubated in the dark for 2 h at 4 °C with gentle shaking. The reaction was quenched with the addition of ice cold 2.5 mM glycine and incubated for 30 min at 4 °C. Methylated P-gp was then concentrated to 1 ml (Centricon YM-50 or YM-100; Millipore 4) and subsequently diluted with 9 ml of quench buffer (100 mM NaCl, 20 mM Hepes pH 7.5, 0.2 mM TCEP, 0.01 % sodium cholate and 0.035 % βDDM). The concentration/dilution step was repeated two times, leading P-gp at 8–12 mg/ml. In some cases, 2 mM of ligand was added to the concentrated protein and the mixture was keep overnight in the dark with gentle shaking at 4 °C. Then, the concentration and dilution step was repeated two times and the protein solution was concentrated until it research the initial volume.

P-gp crystals were grown by using the sitting drop method at 4 °C, combining protein and mother liquor at 1:1 volume per volume. The different mother liquors were 0.1 M Hepes (pH 7–8), 50 mM lithium sulfate, 10 mM EDTA, and 24–29.5 % (w/v) PEG 600 Da at 4 °C. See Table 9 and Table 10.

% PEG 600 Da	**24%**	**24.5%**	**25%**	**25.5%**	**26%**	**26.5%**	**27%**	**27.5%**	**28%**	**28.5%**	**29%**	**29.5%**
ml PEG600 at 40%	24	24.5	25	25.5	26	26.5	27	27.5	28	28.5	29	29.5
ml L_iSO_4 at 0.5M	4	4	4	4	4	4	4	4	4	4	4	4
ml EDTA at 0.5M	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
ml H_2O	8	7.5	7	6.5	6	5.5	5	4.5	4	3.5	3	2.5

Table 9. Preparation of solution A.
Final volume is 32 ml. For L_iSO_4 and EDTA their final concentration was 50 mM and 10 mM respectively.

		24 %	24.5%	25%	25.5%	26%	26.5%	27%	27.5%	28%	28.5%	29%	29.5%
pH		1	2	3	4	5	6	1	2	3	4	5	6
8.4	A												
8.2	B												
8	C												
7.8	D												
7.6	A												
7.4	B												
7.2	C												
7	D												

Table 10. Crystallization conditions tested for mouse P-gp.
Mother liquor: 550 µL of solution A at *X* concentration of PEG (red arrow) + 60 µL of 1M HEPES at *X* pH (blue arrow), left at 4 ° C at least 1 hour.
Dashed lines indicate the different plates used (4 plates of 24 well).

Molecular docking

This computational method aims at predicting the binding modes of the ligand and receptor. Thus, the main interest is the possibility to screen thousands of compounds in a short time (virtual- or *in silico*- screening), accelerating the drug discovery. Furthermore, this approaches allow to explore binding which are not observed by structural methods. Here, basic concepts and several methodologies are briefly presented, reviewed in [260-264].

Two main components are involved in docking: the search algorithm and the scoring function (summarized in Table 11 and illustrated in Figure 48). The first generates all possible binding modes between receptor and ligand using several approximations: rigid (receptor and ligand are rigid), pseudo-flexible (receptor is rigid and ligand is flexible) and flexible (both are flexible). The second use a number of mathematical methods to predict the binding affinity and rank different docked ligands in order to discriminate the active compounds. The choice of a method has an impact on the computation time and means. It have been shown that scoring functions (entropy and desolvation effects) are the Achilles heel of molecular docking. Other issues may be the ligand protonation, tautomerism, and stereoisomerism.

The first step in molecular docking is to obtain a structure for the receptor (X-ray crystallography, NMR or modeling) and check its quality. Afterwards the receptor and ligand must be prepared for the docking (remove water molecules, add hydrogens, stabilizing charges, filling the missing residues, generating the side chains, minimized, etc.), and parameters are given in the docking software, which offers one or more modes of potential interactions that can then be used in a multitude of ways.

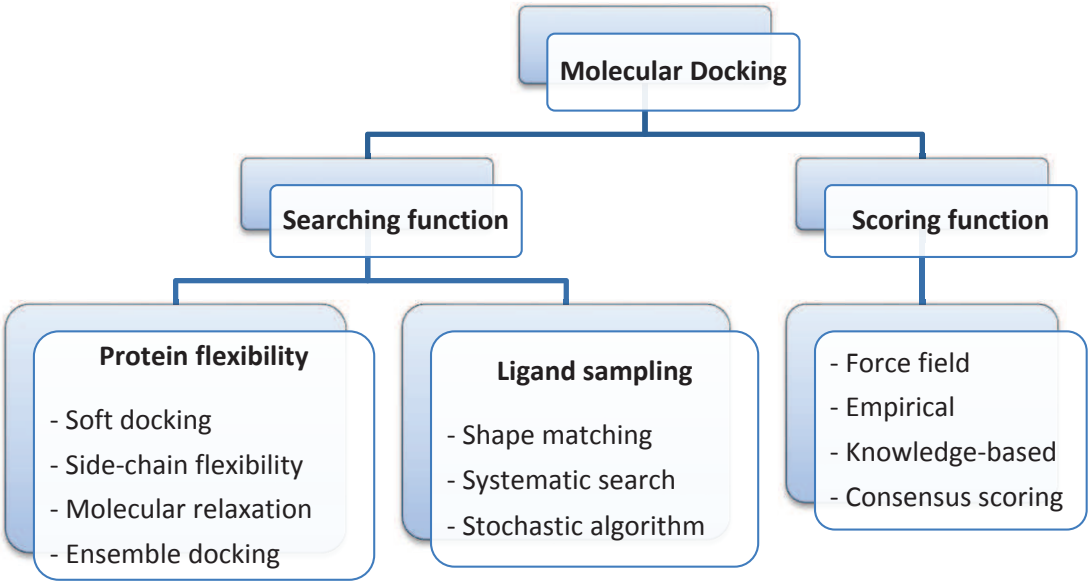


Figure 48. Workflow used in molecular docking.
Adapted from: [262]

	Method	Principle	Examples programs that use the method
Protein flexibility	Soft docking	Considers protein flexibility implicitly. Allow small conformational changes.	
	Side-chain flexibility	While the protein backbone is kept rigid, the side-chain conformations are sampled.	
	Molecular relaxation	Use first a rigid docking to see possible clashes and then relaxed or minimized the protein backbone and side-chain atoms using Monte Carlo (MC) or Molecular Dynamic simulations (MDS).	- AutoDock 4 (MC)
	Docking of multiple protein structures	Use different protein structures to represent several possible conformational changes.	- FlexE
Ligand sampling	Shape Matching	Considers that the molecular surface of the placed ligand must complement the molecular surface of the binding site on the protein.	- DOCK - FRED - Surflex
	Systematic Search	Explore all degrees of freedom of the ligand. Three types: Exhaustive Search (ES, rotates all possible rotatable bonds of the ligand), Fragmentation (F, ligand is divided into different rigid parts or fragment and then the binding conformation is incrementally grown by placing either one fragment at a time or all the fragments in the binding site) and Conformational Ensemble (CE, use of libraries of pre-generated conformations).	- Glide (ES) - FRED (ES) - DOCK (F) - FlexX (F) - FLOG (CE) - DOCK3.5 (CE)
	Stochastic Algorithms	The ligand is considered as a whole. Involve random changes to modify the position of the ligand (translation and rotation) as well as torsion angles in order to generate different conformations. The manly methods are Monte Carlo (used Boltzmann probability function), Genetic Algorithms (GA, based on concepts borrowed from Darwinian evolution) and Tabu search (acceptance depends on the previously explored areas in the conformational space of the ligand).	- MOE-Dock (MC,TS) - GOLD (GA) - AutoDock (GA) - AutoDock-Vina (MC-+ search methods) - PRO_LEADS (TS)

Scoring Functions	Force Field Scoring Functions	Binding free energy is a sum of independent molecular mechanics force fields potentials, such as Coulomb, van der Waals, and hydrogen bonding. Solvation and entropy contributions can also be considered.	- D-Score - G-Score - GOLD - AutoDock - DOCK
	Empirical Scoring Functions	Binding energy score of a complex is calculated by summing up a set of weighted empirical energy terms such as VDW energy, electrostatic energy, hydrogen bonding energy, desolvation term, entropy term, hydrophobicity term, etc.	- LUDI - SYBYL/F-Score - ChemScore - SCORE - Fresno - X-SCORE
	Knowledge-Based Scoring Functions	Designed to reproduce experimental structures rather than binding energies. Protein–ligand complexes are modelled using relatively simple atomic interaction-pair potentials.	- PMF (potential of mean force) - DrugScore - SMoG
	Consensus Scoring	A docked pose is rescoring with different scoring functions or a combination thereof.	- Cscore (SYBYL) - MultiScore - X-Cscore

Table 11. Methods used for protein-ligand docking.
Adapted from: [262]

The scripts for AutoDock-Vina is displayed in appendix Figure 3.

Enzyme kinetics models and statistical analysis

Enzymes catalyze the vast majority of chemical transformations inside cells. They accelerate the rate of chemical reaction with a high selectivity. They are regulated by positive and negative feedback systems, allowing a precise control over the rate of reaction. The main interest to know the enzyme's mechanism is to control and manipulate the course of metabolic events, e.g., for a successful design of drugs. Here several models for enzyme kinetics (to determine its maximum reaction velocity and its binding affinities for substrates and inhibitors) and their evaluation are presented (reviewed in [265-269]).

The first thing to know about enzyme reaction is that it is saturable, meaning that velocity increases linearly as the substrate increases only to a certain extent, reaching a maximal velocity which is independent of substrate concentration. The first equation to explain this saturation in rate was in 1902 by Victor Henri, where substrate binding and product release are reversible. 11 years later a more thoroughly established equation, based on experimental observations, was determinated by Maud Menten and Leonor Michaelis (known as Michaelis−Menten or sometimes Henri−Michaelis−Menten equation). In this model the substrate binding is reversible while the product release is not, as follows:

Reaction scheme

$$E + S \underset{k_1}{\overset{k_0}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$

Where, E: free enzyme; S: substrate;
ES: enzyme-substrate complex;
k_0: second order rate constant;
k_1 and k_2: first order rate constants

Follows the mass law

$$\begin{cases} dS/dt = -k_0 [E][S] + k_1 [ES] \\ dE/dt = -k_0 [E][S] + (k_1 + k_2)[ES] \\ dES/dt = k_0 [E][S] - (k_1 + k_2)[ES] \\ dP/dt = k_2 [ES] \end{cases}$$

Suggested by their experimental studies the substrate S is in instantaneous equilibrium with the complex ES (k_0, $k_1 >> k_2$), thus $dS/dt = 0$; $k_0 [E][S] = k_1 [ES]$ (∗). Since $[E]_T$, the total enzyme concentration is equal to $[E] + [ES]$, where [E] is the free enzyme. Substituting $[E]$ by $[E]_T - [ES]$ in the equation (∗) and solving for [ES] we obtain: $ES = \frac{[E]_T [S]}{\frac{k_1}{k_0} + [S]}$. Replacing this in the rate of product P,

$$v = \frac{dP}{dt} = k_2 [ES] = V_{max} \frac{[S]}{K_S + [S]}$$, where $V_{max} = k_2 [E]_T$ and $K_S = k_1/k_0$.

Based on the same reaction mechanism in 1925 Briggs and Haldane suggested an alternative hypothesis (the quasi-steady state approximation): if the enzyme is present in "catalytic" amounts ([S] >> [E]), then, very shortly after mixing E and S, a steady state is established in which the concentration of ES remains essentially constant with time, thus: $\frac{dE}{dt} = 0$ and $k_0 [E][S] = (k_1 + k_2)[ES]$. Replacing $[E]$ by $[E]_T - [ES]$ in this equation we obtain: $[ES] = \frac{k_0 [E]_T [S]}{k_0 [S] + (k_1 + k_2)} = \frac{[E]_T [S]}{[S] + \frac{(k_1 + k_2)}{k_0}}$, thus the

rate of product P, $v = \frac{dP}{dt} = k_2[ES] = \frac{k_2[E]_T[S]}{[S]+\frac{(k_1+k_2)}{k_0}} = V_{max}\frac{[S]}{[S]+K_m} = \frac{V_m S}{K_m+S}$ eq.1, where V_{max} = V_m= $k_2[E]_T$ and $K_m = (k_1 + k_2)/k_0$, known as Michalis-Menten constant, which provides an indication of the binding strength of that enzyme to its substrate. k_2 is also known as catalytic constant, K_{cat}, = $V_m/[E]_T$ and give a measure of the effectiveness of the catalysis by the enzyme on the substrate.

Only the K_m has a slightly different meaning, however when k_0, $k_1 >> k_2$ we have $K_m \rightarrow K_S$. This equation says that the velocity of an enzyme-catalyzed reaction is defined by two constants (K_m and V_m) and the concentration of the substrate at that moment. It describes the known rectangular hyperbola, in which velocity increases linearly as the substrate increases, but when $S >> K_m$ ($v = V_m$) is no longer dependent on S. Note that when $v = V_m/2$, $S = K_m$. Thus K_m is defined by the substrate concentration that gives a velocity equal to one-half the maximal velocity.

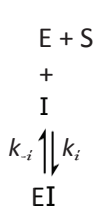
In general, linear transformation of equation 1 is used to define the kinetics parameters (K_m and V_m) by simply drawing a straight line instead the direct plot (v vs. [S]), summarized in Table 12.

Plot of	Potted variables	Slope	*y*-axis intercept	*x*-axis intercept	Disadvantages
Lineweaver-Burk (kwon as double reciprocal plot)	$1/v$ vs. 1/[S]	K_m/V_m	$1/V_m$	$-1/K_m$	Small errors on v are enlarged when reciprocals are taken and the slope is affected.
Hanes Wolf	[S]/v vs. [S]	$1/V_m$	K_m/V_m	$-K_m$	At low v values (near the y-axis) the slope is slightly affected.
Eadie-Hofstee	v vs. v/[S]	$-K_m$	V_m	K_m/V_m	Both are subject to experimental error.
Eadie-Scatchard	v/[S] vs. v	$-1/K_m$	V_m / K_m	V_m	

Table 12. Common linear transformations of equation (1).
According to [265].

Modulation of Enzyme Activity

As mentioned above, enzymes are selective, however depending on the nature of the binding site and the affinity of the substrate, several compounds (e.g. substrate analog) may be accepted by the enzyme and alter its catalysis. Thus, some molecules may bind to the active site and prevent the substrate binding, having an antagonist effect (competitive inhibition); or bind exclusively either to the free enzyme in another place than the active site (non-competitive inhibition) or only to the substrate-enzyme complex (uncompetitive inhibition); or both: free enzyme or enzyme-substrate complex (mixed inhibition). Only the equation of competitive inhibition will be demonstrated, the others follow the same type of demonstration.



$$dES/dt = k_0 [E] [S] - (k_1 + k_2) [ES] \quad \text{eq.2}$$

$$dEI/dt = k_i [E] [I] - k_{-i} [EI] \quad \text{eq.3}$$

$$dP/dt = k_2 [ES] \quad \text{eq.4}$$

$$[E]_T = [E] + [ES] + [EI] \quad \text{eq.5}$$

From the previous assumption of steady-state (dES/dt = 0) and solving for [ES], the eq. 2 can be written: $[ES] = \frac{k_0[E][S]}{k_1 + k_2} = \frac{[E][S]}{K_m}$. Assuming that E + I $\rightleftharpoons$ EI reaches rapid equilibrium (dEI/dt = 0) and solving for [EI], the eq. 3 can be written: $[EI] = \frac{k_i[E][I]}{k_{-i}} = \frac{[E][I]}{K_I}$; where $K_I = k_{-i}/k_i$. Replacing [ES] and [EI] and solving for [E], the eq. 5 can be written: $[E] = \frac{K_I K_m [E]_T}{(K_I K_m + K_I[S] + K_m[I])}$. Finally, the rate of product formation is obtained replacing *[ES]* and *[E]*: $v = \frac{dP}{dt} = k_2[ES] = k_2 \frac{[E][S]}{K_m} = \frac{k_2 K_I [E]_T [S]}{(K_I K_m + K_I[S] + K_m[I])} = \frac{V_{max}[S]}{\left(K_m + [S] + \frac{K_m[I]}{K_I}\right)}$ *or* $v = \frac{V_m S}{K_m\left(1 + \frac{I}{K_I}\right) + S}$

The Table 13 displays the enzyme reaction scheme, rate equation, *v vs [S]* and the Lineweaver-Burk representation (1/*[S]* vs 1/*v*) of each inhibition model.

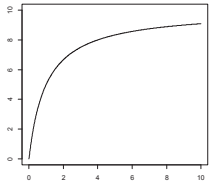
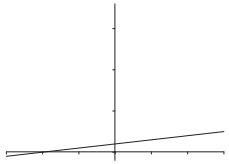
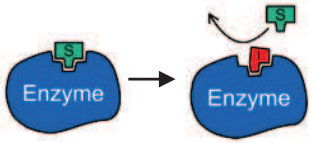
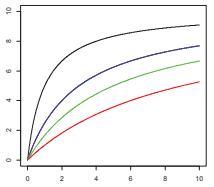
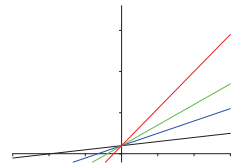
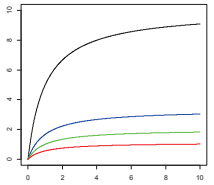
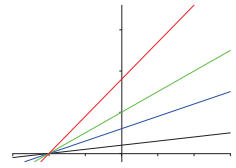
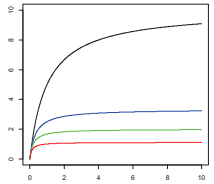
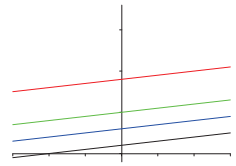
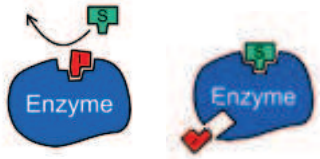
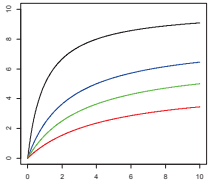
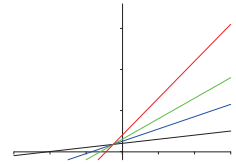
Inhibition models	Illustration	Enzyme reaction scheme	Rate equation	v vs $[S]$	Lineweaver-Burk
None		$E + S \underset{k_1}{\overset{k_0}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$	$v = \frac{V_m S}{K_m + S}$		
Competitive		$E + S \underset{k_1}{\overset{k_0}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$ $+$ I $\Vert\downarrow K_I$ EI	$v = \frac{V_m S}{K_m\left(1 + \frac{I}{K_I}\right) + S}$		
Non-competitive		$E + S \underset{k_1}{\overset{k_0}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$ $+$ $+$ I I $\Vert\downarrow K_I$ $\Vert\downarrow$ EI ESI	$v = \frac{V_m \Big/ \left(1 + \frac{I}{K_I}\right) S}{K_m + S}$		
Uncompetitive		$E + S \underset{k_1}{\overset{k_0}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$ $+$ I $\Vert\downarrow K_I$ ESI	$v = \frac{V_m \Big/ \left(1 + \frac{I}{K_I}\right) S}{K_m \Big/ \left(1 + \frac{I}{K_I}\right) + S}$		
Mixed		$E + S \underset{k_1}{\overset{k_0}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$ $+$ $+$ I I $\Vert\downarrow K_{i1}$ $\Vert\downarrow K_{i2}$ EI ESI	$v = \frac{V_m \Big/ \left(1 + \frac{I}{K_{I2}}\right) S}{K_m \frac{\left(1 + \frac{I}{K_{i1}}\right)}{\left(1 + \frac{I}{K_{i2}}\right)} + S}$		

Table 13. Standard inhibition models.

[I]=0; [I]=1; [I]=2; [I]=4.

Substrate and product may also regulate the activity of the enzymes. Substrate inhibition (Figure 49, plot 5) occurs when second substrate acts as inhibitor through a different site or by changing its binding mode. In other cases, the activity of enzymes is controlled through regulatory site(s) in either a positive (allosteric activator) or negative (allosteric inhibitor) cooperative way (Figure 49, plots 2 and 4 respectively). Several equations and models describe this cooperative behavior, among them the most common are: the Hill equation [270], reversible Hill equation [271], Adar equation [272], Monod, Wyman and Changeux model (MWC) [273] and the Koshland, Nemethy and Filmer model (KNF) [274].

The Hill equation is $v = V_m S^h/(K_m{}^h + S^h)$ and its coefficient, h, may indicate the degree and direction of cooperativity: for a classical Michaelis–Menten enzyme, h is equal to 1; for enzymes with positive cooperativity, h is >1; and enzymes with negative cooperativity, h is <1.

The MWC and KNF models share the concept of two enzyme states: a tense one where substrate binding is weak (T-state) and a relaxed where substrate binding is strong (R-state). While the first suggest that the two states exist in the absence of the regulatory ligands, the later suggest that is the ligand that induce a conformational switch in the enzyme [275].

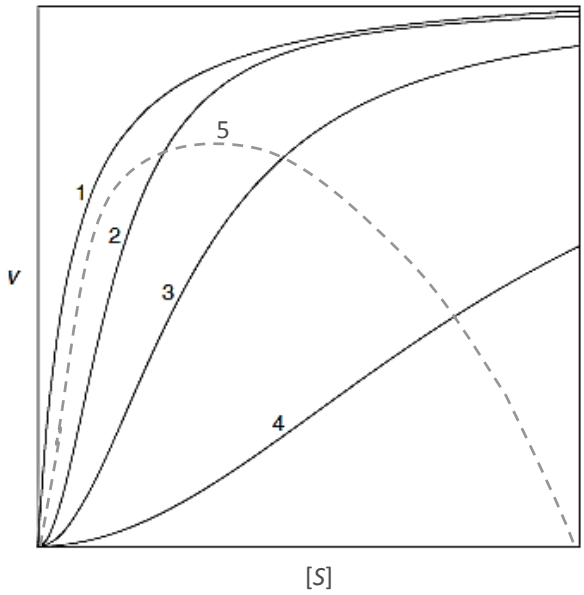


Figure 49. Evolution of the substrate concentration in an enzyme-catalyzed reaction. Plot 1: normal Michaelis–Menten-type enzyme kinetics; plot 2: presence of an allosteric activator; plot 3: positive cooperative binding; plot 4: presence of an allosteric inhibitor; plot 5: inhibition by excess substrate.
Adapted from [267].

Enzyme inhibitions models used in chapter III

All previous cases and the new ones can be described by the initial Michalis-Menten equation written as follow:

$$v = \frac{V_m^{app} S^{h_{app}}}{K_m^{app\,h_{app}} + S^{h_{app}}}$$

Where V_m^{app}, $K_m^{app\,h_{app}}$ and h_{app} will change according to each case. For example, in the case of non-inhibition (Michaelis–Menten equation), $h_{app} = 1;\ V_m^{app} = V_m; K_m^{app} = K_m$. In a competitive inhibition, the V_m is constant and the K_m increases with the concentration of inhibitor, thus, $h_{app} = 1;\ V_m^{app} = V_m; K_m^{app} = K_m\left(1 + \frac{I}{K_I}\right)$.

Table 14 shows the classical inhibition models in this way, as well as the variation of V_m^{app} and K_m^{app} against $[I]$.

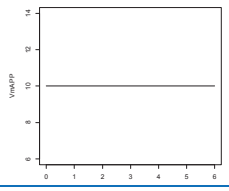
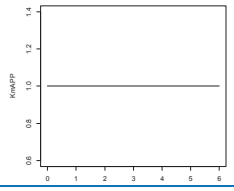
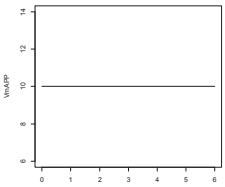
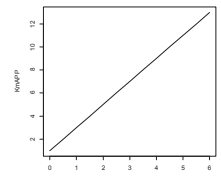
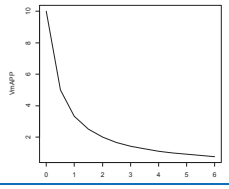
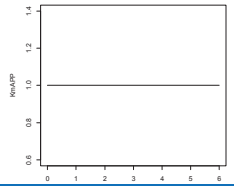
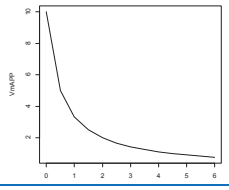
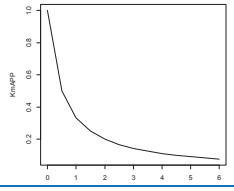
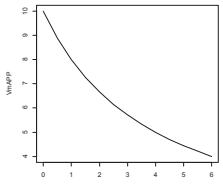
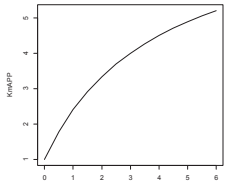
Inhibition Type	V_m^{app}	K_m^{app}	h_{app}	V_m^{app} *vs* $[I]$	K_m^{app} *vs* $[I]$
None, Michalis-Menten (eq. 2.1)	V_m	K_m	1		
competitive inhibition (CI) (eq. 3.1)	V_m	$K_m\left(1 + \frac{I}{K_I}\right)$	1		
non-competitive inhibition (NCI) (eq. 4.1)	$\frac{V_m}{\left(1 + \frac{I}{K_I}\right)}$	K_m	1		
uncompetitive inhibition (UI) (eq. 5.1)	$\frac{V_m}{\left(1 + \frac{I}{K_I}\right)}$	$\frac{K_m}{\left(1 + \frac{I}{K_I}\right)}$	1		
mixed inhibition (MI) (eq. 6.1)	$\frac{V_m}{\left(1 + \frac{I}{K_{i2}}\right)}$	$K_m\frac{\left(1 + \frac{I}{K_{i1}}\right)}{\left(1 + \frac{I}{K_{i2}}\right)}$	1		

Table 14. Influence of inhibitor on enzymatic constants

Our data needed more complex scenarios, such as: cooperativity, efflux activation and inhibition by the substrate, generally modulated by the inhibitor. We addressed these phenomenon by introducing into the above equations the following parameters affecting either the V_m or K_m: the

Hill coefficient (h), the activation constant ($1 + I/K_A$) and the inhibition by the substrate ($1 + S/K_{SI}$), displayed in Table 15.

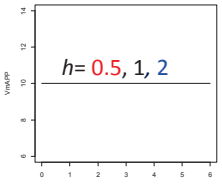
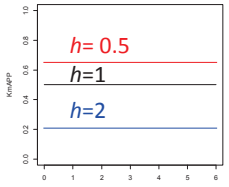
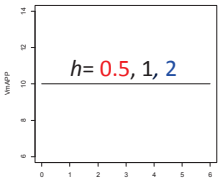
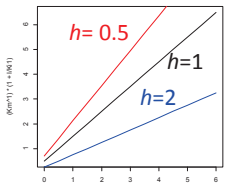
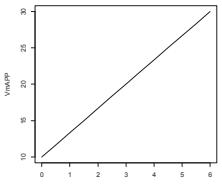
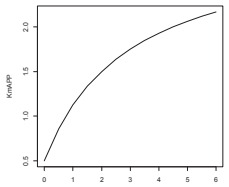
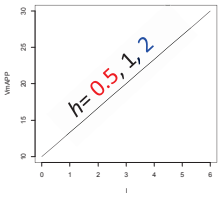
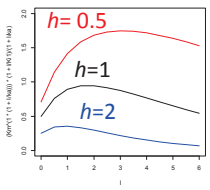
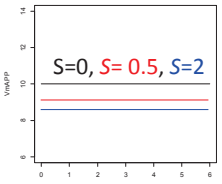
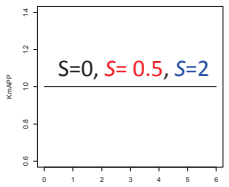
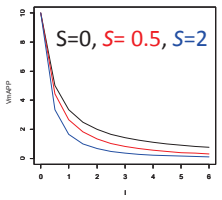
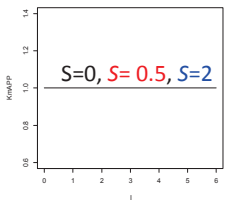
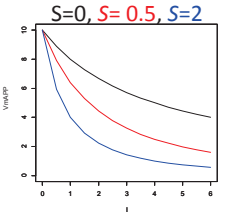
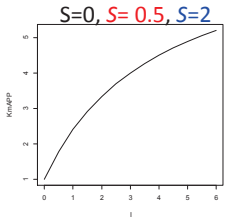
Inhibition Type	V_m^{app}	K_m^{app}	h_{app}	V_m^{app} *vs* $[I]$	K_m^{app} *vs* $[I]$
Cooperativity + none (eq. 2.2)	V_m	${K_m}^{h_{app}}$	h		
Cooperativity + Competitive inhibition (eq. 3.2)	V_m	${K_m}^{h_{app}}\left(1+\frac{I}{K_I}\right)$	h		
Competitive inhibition + activation (eq. 3.3)	$V_m\left(1+\frac{I}{K_A}\right)$	$K_m\frac{\left(1+\frac{I}{K_I}\right)}{\left(1+\frac{I}{K_A}\right)}$	1		
Cooperativity modulated by inhibitor + competitive inhibition + activation (eq. 3.5)	$V_m\left(1+\frac{I}{K_A}\right)$	${K_m}^{h_{app}}\frac{\left(1+\frac{I}{K_I}\right)}{\left(1+\frac{I}{K_A}\right)}$	$h\left(1+\frac{I}{K_A}\right)$		
Substrate non-competitive inhibition (Sub. NCI) (eq. 2.7)	$\frac{V_m}{\left(1+\frac{S}{K_{SI}}\right)}$	K_m	1		
Sub. NCI modulated by inhibitor + Non-competitive inhibition (eq. 4.9)	$\frac{V_m}{\left(1+\frac{S}{K_{SI}}\frac{I}{K_I}\right)\left(1+\frac{I}{K_I}\right)}$	K_m	1		
Sub. NCI modulated by inhibitor + Mixed inhibition (eq. 6.9)	$\frac{V_m}{\left(1+\frac{S}{K_{SI}}\frac{I}{K_I}\right)\left(1+\frac{I}{K_{i2}}\right)}$	$K_m\frac{\left(1+\frac{I}{K_{i1}}\right)}{\left(1+\frac{I}{K_{i2}}\right)}$	1		

Table 15. Influence of cooperativity, activation and inhibition by the substrate on enzymatic constants.

Note that when [*I*] = 0, all the equations become the Michaelis–Menten equation. Except for the substrate inhibition (eq. 2.7), which becomes the Michaelis–Menten equation when [*S*] = 0.

Model evaluation

For each tested model, the parameters were estimated by non-linear regression of all experimental conditions simultaneously. In order to find which model fits the data sets the better, statistical approaches were used to discriminate between them. One of the most widely used methods to compare multiple models is the Akaike's Information Criterion (AIC), which estimates the information lost between model and data. The AIC is defined by the equation below and the model chosen is the one with the lowest AIC score (reviewed in [276, 277])

$$AIC = -2log_{10}(likelihood) + kn_{par}$$

Where, n_{par}: number of parameters in the fitted model; k: penalty per parameter to be used (default is 2); log_{10} *likelihood:* measure of the fit between predicted and observed values; kn_{par}*:* penalty for over-fitting when increasing the number of parameters.

When the number of data points (N) is small or the number of parameters (n_{par}) is large (N / n_{par} < 40), it is recommended to use the corrected AIC as follows,

$$AICc = AIC + 2k(K+1)/(N-K-1)$$, where K is the number of parameters + 1.

According to the Figure 50, when the difference between two models is equal to two, the model A has 70 % of chance to be the corrected one (red line in the Figure 50). When this difference is zero, they have the same probability to be correct.

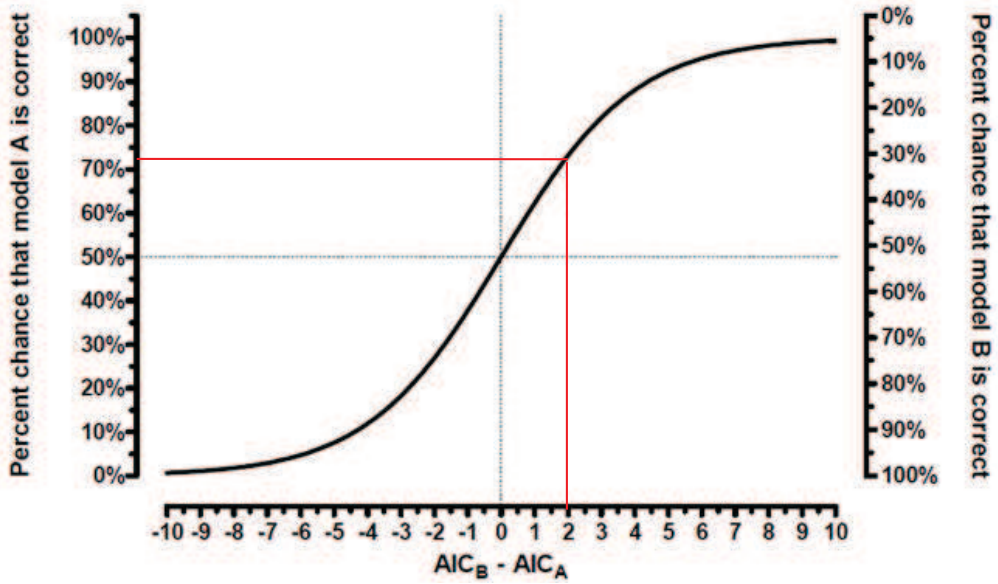


Figure 50. Relationship between the difference in AIC (or AICc) scores and the probability of each model to be true.

To obtain the AIC score the R software was used [278], or Tinn-R, which is one of its interfaces. In practical, we need to (i) have the information into columns (e.g., concentration of substrate, effect observed, and error), (ii) have all the information in a work folder, (iii) save the data as csv format, (iv) import data. For example, for the competitive inhibition model the script is as follows:

```
# Format des données (enregistrer en .csv)
# Informations en colonnes
# 1. S: Concentration Substrat   (Variable indépendante) , µM
# 2. I: Concentration Inhibiteur   (Variable indépendante), µM
# 3. TR   (Variable dépendante) transport rate, pmol/mg/s
# 4. Km, Vm = Michaelis parameters
# 5. KI, Ki1, Ki2, KA, KSI = inhibition constants

>modelCOMPET= as.formula("TR~Vm*S/(Km*(1+I/KI)+S)")

>fitCOMPET=nls(modelCOMPET, data=data, start=list(Vm=30,
Km=0.8, KI=0.5),)

>summary(fitCOMPET)

>AIC(fitCOMPET, k=2))
```

The results are:

Parameters	Estimate	Std. Error	t value	Pr(>|t|)
Vm	43.2473	2.6652	16.227	< 2e-16
Km	1.5700	0.2589	6.065	4.00e-08
KI	0.7890	0.1123	7.023	6.09e-10

AIC = 423.382

The R-scripts of final models used in chapter III are displayed in appendix Figure 4.

CHAPTER I: NON P-GP-SUBSTRATES HIV PROTEASE INHIBITORS SET UP

Context

As mentioned above, HIV-1 protease is a protein crucial to allow the virus to infect new cells, therefore one of the targets in the therapeutic treatment of AIDS. Unfortunately, several HIV protease inhibitors used for the AIDS treatment are transported by P-gp, ABCG2 and MRPs. Some of them also inhibit such transport (IC_{50} > 10 µM) or even induce the expression of these pumps (reviewed in [16, 17]). In order to have a good therapeutic dose without having to take multiple molecules (e.g. HIV inhibitor and P-gp inhibitor), our goal was to find a molecule capable of inhibiting both HIV protease and P-gp without being transported by P-gp. In this way, this molecule will not be evacuated by P-gp and may accumulate in normally inaccessible areas due to the presence of these pumps (e.g., brain). These molecules can also be used in HARRT therapy to improve the bioavailability of other anti-HIV agents which are substrates of P-gp. This dual ligand strategy has been described not only for HIV treatment [279, 280] but also for MDR cancer [281, 282].

Approach I: Could compounds that prevent P-gp translocation also inhibit HIV-1 protease activity?

Dipeptide analogs inhibiting P-gp [257] were recently developed in our laboratory and due to their peptide nature were hypothesized to inhibit the activity of HIV protease activity. With the help of Dr. Raphaël Terreux (Associate professor, IBCP), we carried out docking experiments using this dipeptide library to determine whether these compounds could bind to the protease.

The "Sybyl molecular-modeling" software (SYBYL 1.3. Tripos Inc.) was used with Surflex-Dock as search algorithm and Cscore as the scoring function. The receptor used was the HIV-1 protease (PDB code: 1mui). The ligand structures were designed in Sybyl software, with the exception of protease inhibitors that were available on the PDB. Those ligands can be divided on 3 categories:

- Nine protease inhibitors were used as positive controls: amprenavir (APV, PDB code: 3NU3), saquinavir (SQV, PDB code: 3DIX), daurunavir (DRV, PDB code: 3PWM), tripanavir (TPV, PDB code: 2O4P), atazanavir (ATV, PDB code: 3EKY), nelfinavir (NFV, PDB code: 3EKX), ritonavir (RTV, PDB code: 2B60), indinavir (IDV, PDB code: 2B7Z) and lopinavir (LPV, PDB code: 1MUI);

- Seven inhibitors of P-gp were used as possible negative controls: QZ59-RRR, QZ59-SSS, zosuquidar, GF120918, verapamil, tariquidar and progesterone;

- 28 reversine 121 derivatives (molecules of interest): CT1300 to CT1361 molecules.

Docking with default options was conducted in a way to allow protein movement hydrogen and consider ring flexibility. The maximal conformation per fragment was set to 40. The maximal number of rotatable bonds per molecule was set to 200 and the maximum number of poses per ligand to 100.

As expected, protease inhibitors are in the best or modest affinity group (blue and green in Table 16), and most of the negative controls in the worst affinity group (orange and red in Table 16). According to their score, many inhibitors derived from reversine 121 would present a good affinity for the protease (score> 7.5). However this score function is not very effective to discriminate active from inactive compounds. Therefore, a second analysis (AuPosSOMS [283]) was performed to classify molecules according to the similarity of contact between the HIV-1 protease (considered as the receptor) and our compounds (considered as putative ligand) during docking. This analysis (Figure 51) shows that some CT molecules could inhibit the activity of HIV-1 protease (CT1300, CT1357, CT1327, CT1355, CT1340 and CT1345, circled in blue Figure 51).

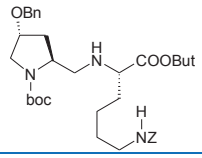
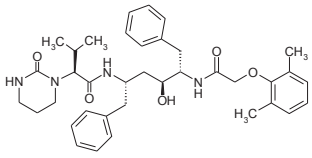
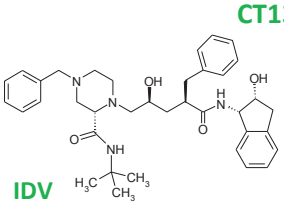
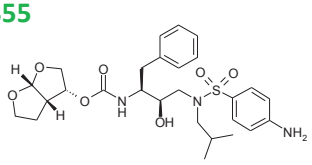
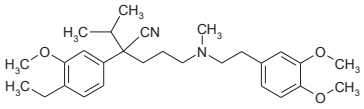
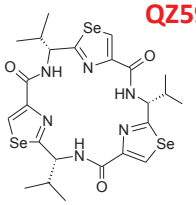
Score >10	Score 7.5-10
LPV, CT1316, CT1361, CT1326, CT1327, CT1347, CT1338, CT1344, CT1336, CT1342, CT1301, CT1337, CT1302, NFV, CT1346, CT1345, SQV, CT1339	**CT1357, TPV, CT1343, CT1329, CT1345, IDV, CT1348, DRV, RTV, CT1340, CT1354, APV, CT1341, CT1300, Zosuquidar, GF120918, CT1356, ATV, CT1355**
CT1347   **LPV**	 **IDV**  **DRV**
Score 5-7.5	**Score 0-5**
CT1333, Verapamil, Tariquidar, CT1351, CT1353, Progesterone	**QZ59(RRR), QZ59(SSS),**
Verapamil 	 **QZ59-RRR**

Table 16. Rankings of tested compounds according to their score (docking results).
Compounds with the best affinity for HIV-1 protease are presented in blue (score > 10) while compounds with the worst affinity are presented in red (score between 0 and 5).

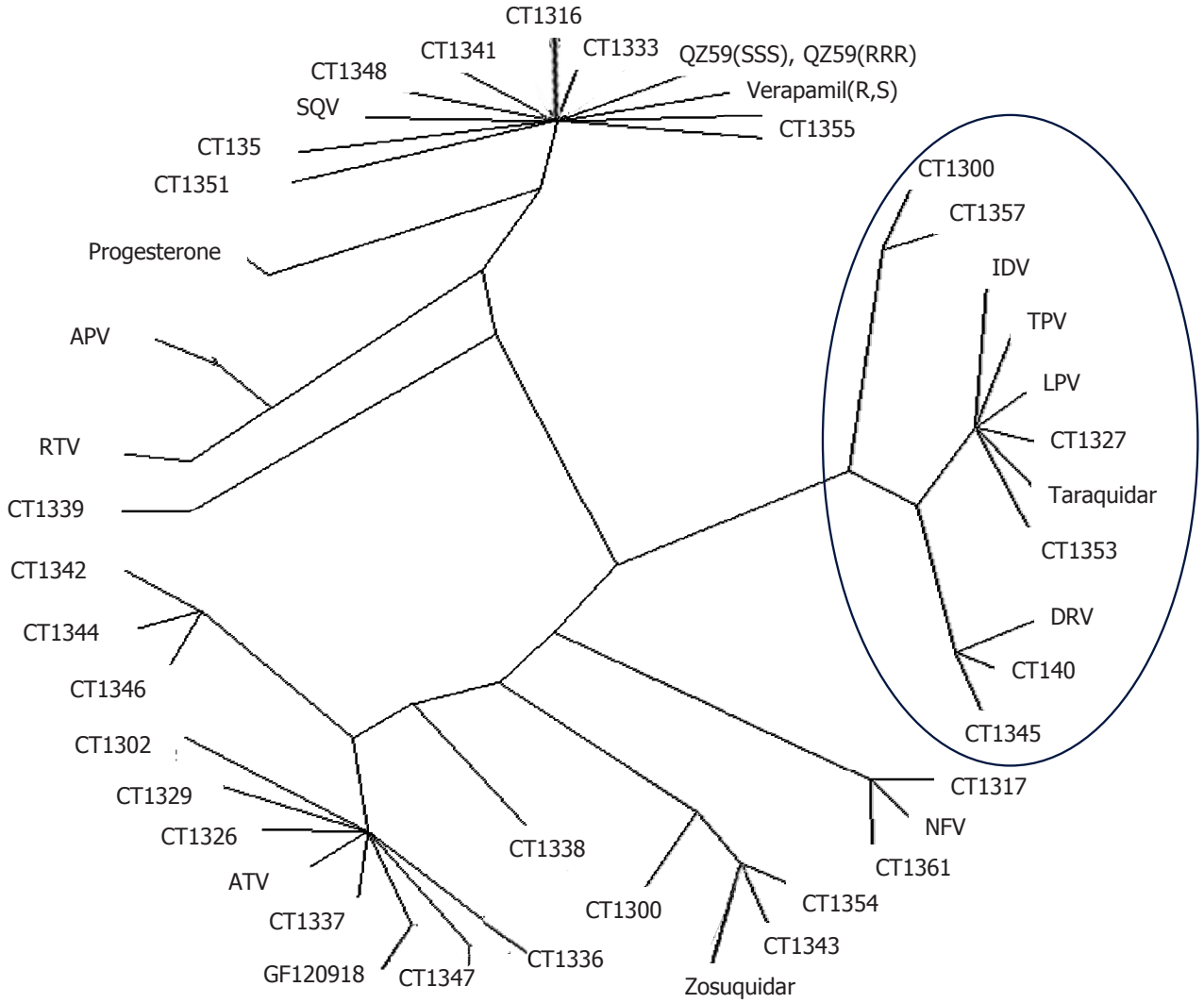


Figure 51. Docking analysis by automatic analysis of poses using SOM (AuPosSOMS). The group of putative inhibitor of HIV-1 protease is encircled in blue.

The efficiency predicted for CT1347 and CT1353 differs according to the analytical method. Thus the inhibition activity of these compounds was tested toward the protease using the system for monitoring the HIV-1 protease activity in T-cells (see Methods). The Gal4/PR cells in presence of doxycycline (Doxy) and a “good” inhibitor allow the expression of GFP. This was observed for the reference inhibitors, indinavir and saquinavir (red and pale green circles respectively in Figure 52). On the contrary, the Gal4/PR cells in presence of Doxy and a “bad” inhibitor does not express the GFP. It was the case of both, CT1347 and CT1353 compounds (orange and green circles respectively in Figure 52), suggesting that they are not able to inhibit the protease in this condition.

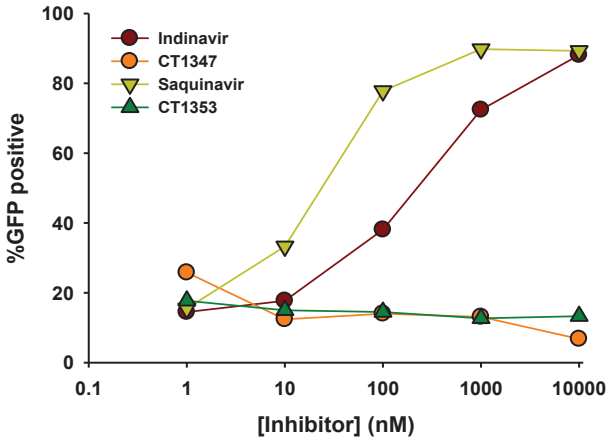
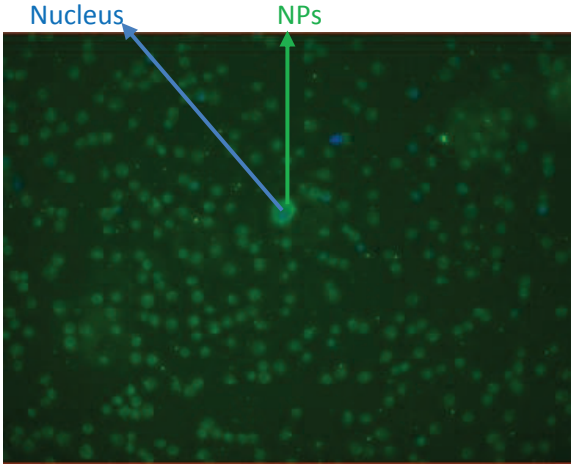


Figure 52. Analysis of HIV-1 protease inhibition activity of CT1347 and CT1353 compounds. Indinavir and saquinavir were used as positive controls.

This lack of inhibition may results either from a lack of efficacy, or from the relatively hydrophobic character of those compounds. Indeed, they could be stuck inside the plasma membrane and could not reach the protease of HIV-1, in the cytoplasm. To circumvent this problem, these compounds has to be more hydrophilic. Thus, the penetration of these compounds into the cell was improved via biodegradable nanoparticles (NPs) of poly lactic acid (PLA) (provided by Dr. Bernard Verrier from IBCP). Penetration of nanoparticles containing green fluorophor (coumarin 6) into the cells was confirmed by fluorescent microscopy (Figure 53A). Inhibitor activity of compounds encapsuled into the NPs was analyzed as previously described. Inhibitor activity of indinavir was not changed by the NPs (GFP expression) (Figure 53B). Although CT1347 and CT1353 where delivered into the cytoplasm by NPs, they did not inhibit the protease (no expression of GFP) (Figure 53B). Thus, these compounds are not able to inhibit the HIV-1 protease activity.

A) Superimposed images of SupT1 Gal4 cells treated with NPs



B) Inhibition of the CT1347 and CT1353 compound mixed with NPs

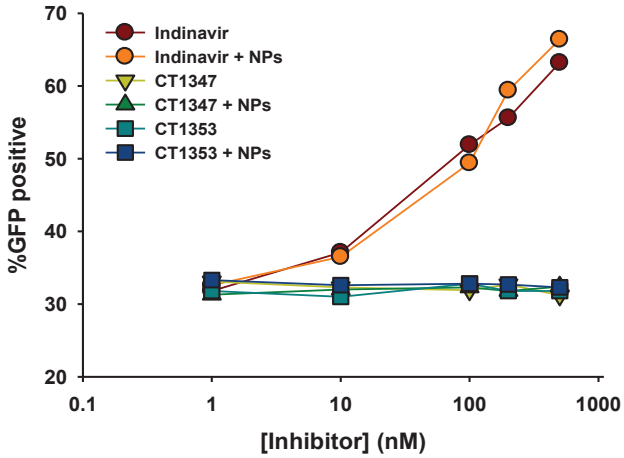


Figure 53. Delivery of CT1347 and CT1353 by NPs of PLA into the cells.

A) Cells were incubated with nanoparticles enclosing the CT1347 or CT1353 and containing green fluorophor (coumarin 6 [284]). The nucleus (in blue) was market with DAPI.

B) As previously described, the inhibition of HIV-1 protease was monitored in T-cells.

In parallel, we tested the inhibitory activity of these compounds in a well-established FRET (fluorescence resonance energy transfer) assay with the HIV-1 Protease Assay Kit from ProteinOne. The substrate is labelled with a fluorescent donor and a non-fluorescent acceptor. When this molecule is entire, the acceptor is close enough to quench the fluorescence emission of the donor. Contrary, when this molecule is cleaved (by the HIV protease) the quenching effect is lost, leading to the appropriate excitation wavelength of the donor (Figure 54A). Thus, the HIV protease in presence of fluorogenic substrate give the maximum of fluorescence. In presence of a "good" inhibitor the fluorescence decreases, indicating the successful inhibition of the protease.

As expected, reference protease inhibitors (pepstatine, indinavir and saquinavir) blocked the HIV-protease substrate cleavage (decrease of fluorescence in Figure 54B). In presence of our compounds, the fluorescence did not change, suggesting that no inhibition of the activity of HIV-1 protease (Figure 54B).

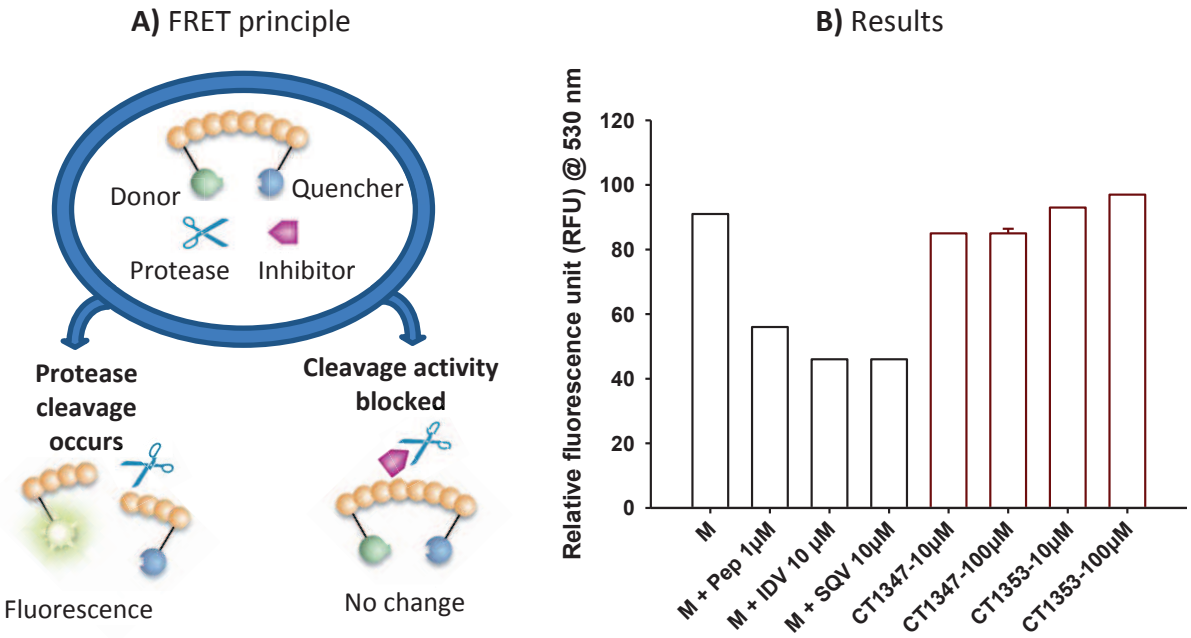


Figure 54. Inhibition of compounds by FRET-based spectrometer assay.
Excitation/emission = 490 nm/530 nm, using a microplate spectrofluorimeter (Tecan Infinite M1000)
M: reaction mix (buffer+ HIVPR+ Substrate); Pep: pepstatine; IDV: indinavir; SQV: saquinavir

These results were presented in the following conferences:

- Scientific Day "ARC 1 - Health" (September 2012, Isle d'Abeau, France). Oral presentation: Structural and functional study of human ABC transporters involved in the efflux of anti-retroviral agents. Martinez L.

- Science Day of Cluster 10 "Infectiologie" (January 2012, Lyon, France). Poster presentation: Novel compounds inhibiting human P-glycoprotein and the VIH-1 protease. Martinez L., Hilton B.,Wolkowicz R., Ettouati L., Paris J., Le Borgne M., Terreux R., Andrieu T., Dussurgey S., Di Pietro A ., and Falson P.

Approach II: Could compounds that block HIV-1 protease activity also prevent P-gp translocation?

The HIV-1 protease is a homodimer, each polypeptide chain containing 99 amino acid residues. The dimer interface forms a tunnel to the active site. This tunnel is covered by two flexible flaps that can twist open and allows proteins to enter. HIV-1 Protease is mostly composed of β sheet arranged in antiparallel way (Figure 55A). The active site contains two catalytic triads, both Asp-Thr-Gly, in which Asp-25 (monomer A) uses a water molecule to cleave the protein at its amide backbone by attacking that bond's carbonyl group. Asp-25' (monomer B) stabilizes the resulting intermediate, allowing the cleavage to proceed to completion (Figure 55B).

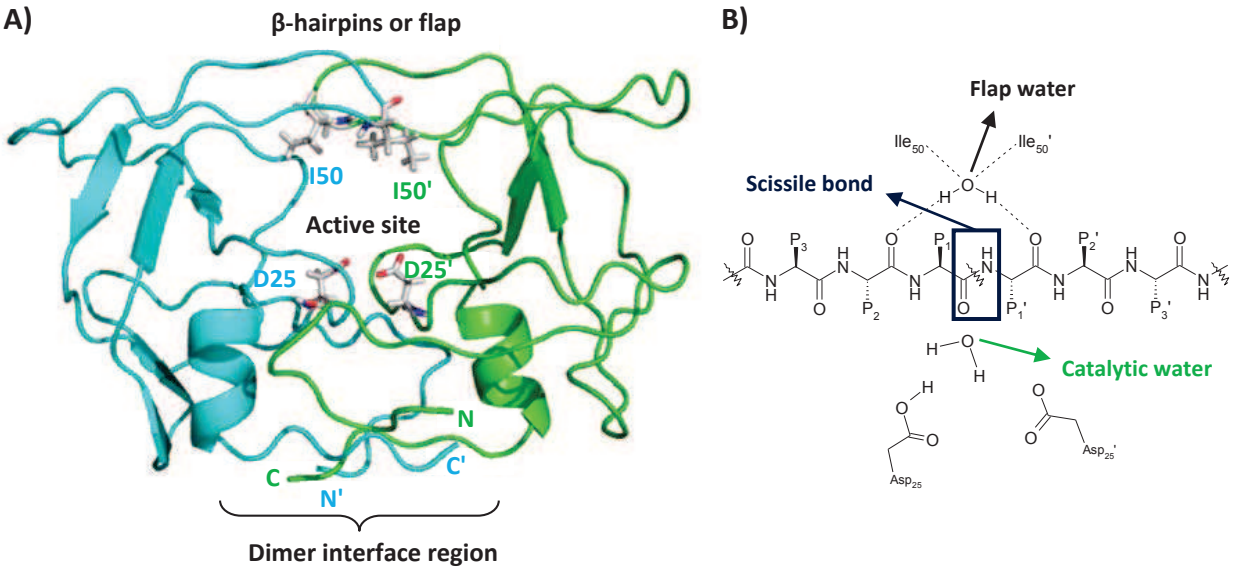


Figure 55. Aspartyl HIV-protease.

A) Representation of the HIV protease (PDB code: 1MUI). Each monomer is shown as ribbons in blue and green. The catalytic aspartic acid and isoleucine residues (enzyme flaps) are shown in CPK-colored stick molecules (red, O atom; grey, C atom). **B)** Representation of peptide substrate binding to the protease of HIV-1 via the "Flap water". According to [285].

The established approach to design inhibitors has been to incorporate a transition-state analog with a peptidomimetic reproducing the linear shape of the peptide substrate, such as: hydroxyethylene, hydroxyethylamine, statine, reduced amide, α-perfluorinated ketones, etc. Up to now, all the protease inhibitors used in AIDS treatment are peptidomimetics using a nonhydrolyzable hydroxyethylene moiety as basic core (Figure 56A), excepting for tipranavir which is a potent non-peptidic inhibitor using a rigid heterocyclic core (Figure 56B). Alternatively, Dr. J. Hasserodt (ENS-Lyon), explored an unusual interaction of two functional groups (a tertiary amine with a ketone or aldehyde function, N→C=O interaction) giving a urea moiety motif (Figure 56C).

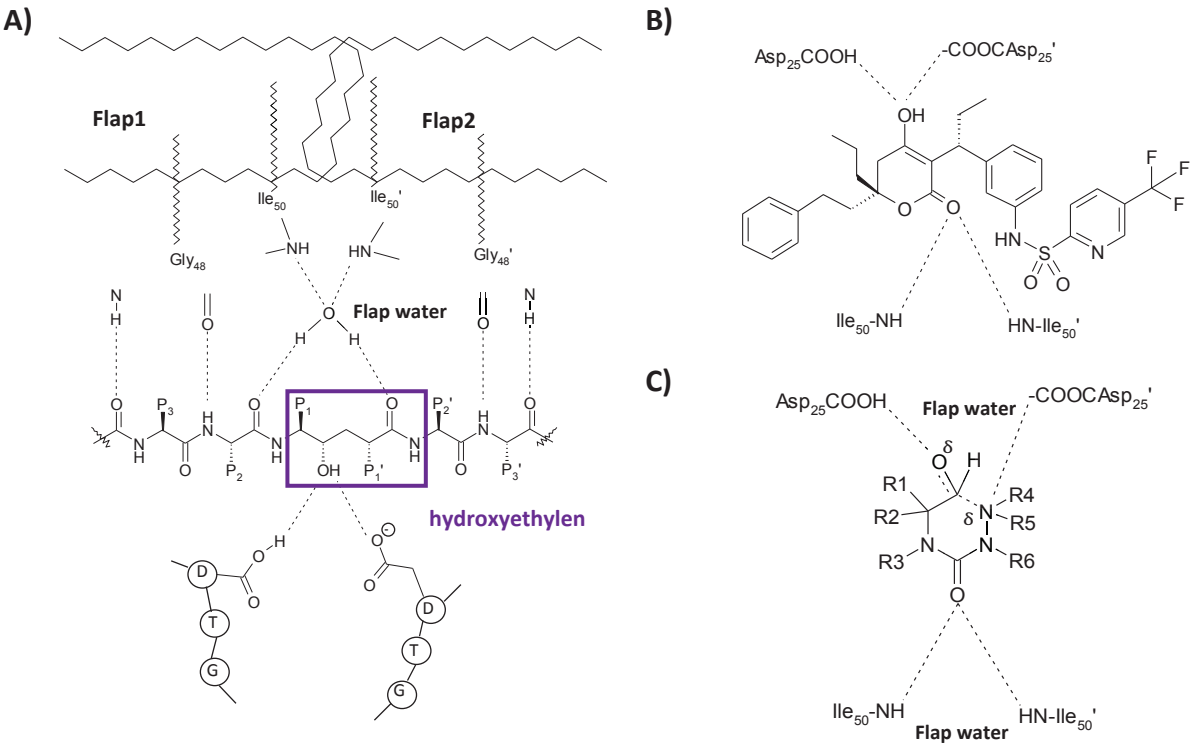


Figure 56. Representation of HIV PR inhibitors and their hydrogen bonds with the water molecule. Hydrogen bonding between the catalytic aspartic acid residues as well as isoleucines with: **A)** a non-hydrolyzable core hydroxyethylene, **B)** a heterocyclic inhibitor Tipranavir, and **C)** a quaternary hemiaminal-based inhibitor candidate.
According to [285] and [286].

This N→C=O interaction is modulated according to the media in which is dissolved: in a deuterated chloroform (CDCl3) the molecule is linear, in a polar protic media (MeOH) the cyclic interaction is pre-organized and at physiological pH adopt the status of a quaternized hemiaminal (Figure 57).

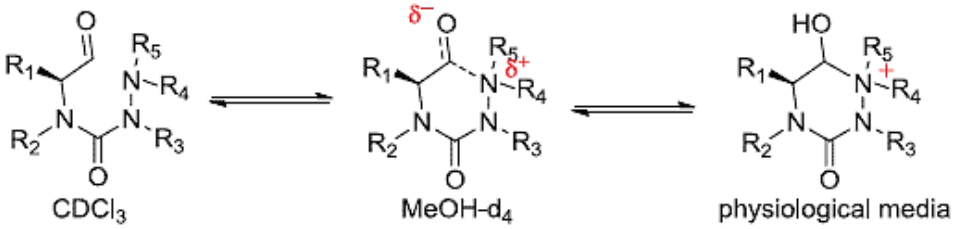


Figure 57. Behavior of the N→C=O interaction.

In the frame of a collaboration with Dr. J. Hasserodt (ENS-Lyon), compounds numbered 36 to 43 were tested as putative HIV-PR inhibitors. *In cellulo* assays showed that half-maximal inhibitory concentration (IC_{50}) were about 20 µM for compounds 37, 41 and 42 while the other candidates did not display significant inhibition activity (Figure 2 in the publication I). Similar results were obtained by FRET-based spectrometer assay; compounds number 37, 41 and 42 were once again the most potent inhibitors (Supporting Information of the publication I).

Efflux of these compounds by P-gp or ABCG2 was determined by analyzing the compounds cytotoxicity and cross-resistance eventually induced by these pumps expression. The half-maximal effective concentrations EC_{50} of the compounds in control and expressing (P-gp or ABCG2) cell lines were similar. This indicated that the expression of these pumps did not change the cytotoxicity pattern (Supporting Information of the publication I) and suggested that they were not transported by these pumps. Nevertheless they do not prevent the transport of mitoxantrone by these efflux pumps (Supporting Information of the publication I), indicated that these compounds are not good inhibitors for P-gp or ABCG2.

These results were published in the Journal Bioorganic & Medicinal Chemistry and were presented in the following conference:

- Journées Jeunes Chercheur de la Société de Chimie Thérapeutique (February 2013, Romainville- Biocitech, France). Poster presentation: Design and synthesis of N-CO-interacted inhibitor candidates and their biological assessment with HIV-1 protease. Gros G., Martinez L., Falson P., Hasserodt J.

Publication I

Bioorganic & Medicinal Chemistry 21 (2013) 5407–5413

Contents lists available at SciVerse ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Modular construction of quaternary hemiaminal-based inhibitor candidates and their in cellulo assessment with HIV-1 protease

Guillaume Gros [a], Lorena Martinez [b], Anna Servat Gimenez [a], Paula Adler [a,†], Philippe Maurin [a], Roland Wolkowicz [c], Pierre Falson [b], Jens Hasserodt [a,*]

[a] Laboratoire de Chimie, Université de Lyon – ENS, 46 allée d'Italie, 69364 Lyon, France
[b] Drug Resistance Mechanism and Modulation Laboratory, BMSSI, UMR 5086 CNRS, Université Lyon 1 – IBCP, 7 passage du Vercors, 69367 Lyon, France
[c] Department of Biology, San Diego State University, 5500 Campanile Drive NLS304, 92182 San Diego, CA, USA

ARTICLE INFO

Article history:
Received 28 February 2013
Revised 30 May 2013
Accepted 6 June 2013
Available online 14 June 2013

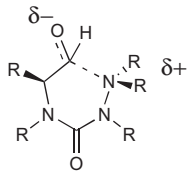
Keywords:
Cyclic urea
HIV-1 protease inhibitors
Quaternary hemiaminal
Transition-state isostere
Antiviral agents
Chemical synthesis
Cellular assay

ABSTRACT

Non-peptidomimetic drug-like protease inhibitors have potential for circumventing drug resistance. We developed a much-improved synthetic route to our previously reported inhibitor candidate displaying an unusual quaternized hemi-aminal. This functional group forms from a linear precursor upon passage into physiological media. Seven variants were prepared and tested in cellulo with our HIV-1 fusion-protein technology that result in an eGFP-based fluorescent readout. Three candidates showed inhibition potency above 20 μM and toxicity at higher concentrations, making them attractive targets for further refinement. Importantly, our class of original inhibitor candidates is not recognized by two major multidrug resistance pumps, quite in contrast to most clinically applied HIV-1 protease inhibitors.

1. Introduction

Over the past 40 years, a limited number of functional groups have been considered and explored for the simulation of the transition states of peptide bond hydrolysis.[1] Only very few have been considered in earnest for drug design targeting the class of aspartic proteases, including hydroxyethylene, α,α-difluoroketones, phosphinates, and α-keto-carboxamides.[2] More recently, the focus has veered away from linear-shaped peptidomimetics incorporating these groups, aiming more at drug-like test candidates.[2,3] Expected benefits of this measure include smaller size, reduced bio-degradability and preventing development of pathogen resistance.[4]



We have originated the exploration of an unusual functional group, or rather, the unusual interaction of two functional groups, for its capacity to simulate electrostatic properties found in the transition states stabilized by aspartic proteases.[5–7] The weak interaction between a tertiary amine and a ketone gives rise to a highly polar and stable molecular motif, provided the two functional groups belong to the same molecule, that their cyclic interaction is pre-organized, and that the molecule is dissolved in polar protic media (MeOH).[8,9] Simple synthetic molecules showing this interaction have sporadically surfaced in the literature.[10] Only a few examples of natural products (alkaloids) showing this interaction are known.[11–17] In water, the pK_a range of such species is estimated to range from 7 to 9,[18–21] so that one can reasonably assume this functional group to become protonated at physiological pH and adopt the status of a quaternized hemiaminal, a highly unusual molecular species. Any preliminary conclusion on the hypothesis that this group mimics the electrostatics of the transition states of peptide hydrolysis may be drawn from our ability to observe elevated affinities with a model peptidase whose choice is motivated by a great supply of SAR studies with inhibitor candidates, namely HIV-1 protease (HIV-1 PR).[2] In 2008, we reported on the diversity-orientated synthesis of a hydrazino urea that displays this N–C═O interaction.[22] This molecular constitution was chosen for two reasons: (a) because the carbonyl contained in a urea motif found in the HIV-1 PR inhibitor DMP450[23,24] has been proven to interact directly via

* Corresponding author. Tel./fax: +33 472728860.
E-mail address: jens.hasserodt@ens-lyon.fr (J. Hasserodt).
† On leave from the University of Ottawa, Canada.

0968-0896/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.bmc.2013.06.018

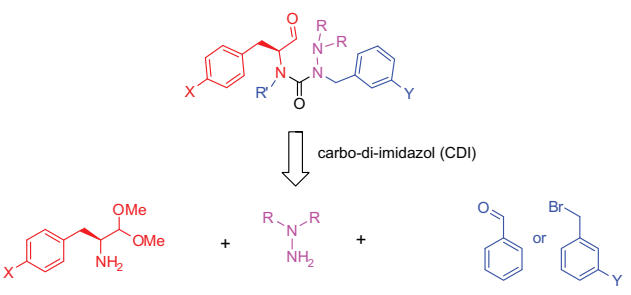
hydrogen-bonding with two protein flaps of HIV-1 protease covering the active site; this results in the extrusion of a key structural water molecule usually found in X-ray crystal structures of complexes of HIV-1 PR with linear, peptidomimetic inhibitors[25] and thus contributes to high binding affinity,[1] and (b) in order to maximize chances of observing total passage to the interacted form in methanol, and indeed this was proven by NMR studies. The measures that achieved this include the choice of (1) a urea that ensure coplanarity of all substituents on the two nitrogens, (2) only two possible orientations (linear or concave) for both displayed functional groups, an aldehyde and a tertiary hydrazine nitrogen, and (3) a distance between both functional groups that allows for closure of a (favored) six-membered ring. These entropy-related precautions allowed us to see almost quantitative ring formation when transferring the compound to methanol, thus depriving the equilibrium of the presence of any free aldehyde, a prerequisite for consideration as an inhibitor/drug candidate. In 2009, we reported the first in vitro inhibition data on this system with HIV-1 protease and obtained values down to 25 μM.[6] Here, we describe the preparation of seven new variants on this system, including the first example of an exhaustively substituted one, a feat we had previously been unable to accomplish.

2. Results and discussion

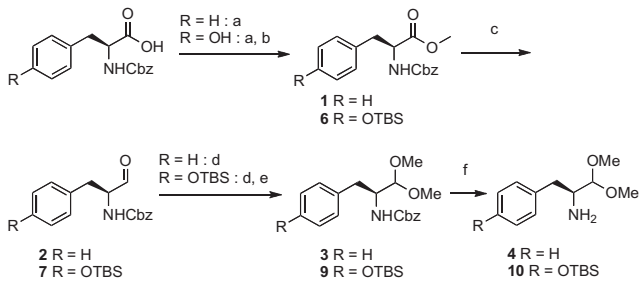
2.1. Chemistry

Our hydrazino urea can be built up in a highly modular fashion from three different building blocks (Scheme 1). To this end, reduced derivatives of any α-amino acid may serve if protected properly, as well as any N,N-disubstituted hydrazine, and further alkylating reagents that will introduce substituents into the two urea nitrogens found in the target structures.

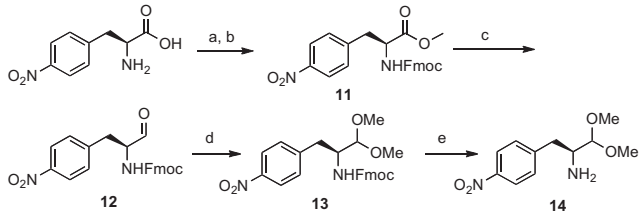
While we were previously unable to produce significant amounts of the first generation of a hydrazino urea,[22] we could develop a highly efficient route to the protected α-amino aldehyde building blocks allowing us to run the synthesis at the multi-gram scale.[26] This became possible due to the choice of simple Cbz-protected α-amino acids (Schemes 2 and 3) that are by far the most affordable among all commercialized protected α-amino acids. Initial esterification of Cbz-protected phenylalanine and tyrosine, followed by protection of the tyrosine hydroxyl with a TBS group gave intermediates **1** and **6**. Their partial reduction to aldehydes **2** and **7** with DIBAL-H is high-yielding while retaining the enantiopurity.[27] These are easily transformed into their dimethyl acetals but the conditions cause the loss of the TBS protective group in the tyrosine series. After its reprotection, both systems could be prepared for introduction into the key urea coupling step by liberation of the amino groups under mild catalytic hydrogenation conditions, and produce aminoacetals **4** and **10** in respectively 77% and 38% yield over four and six steps. Aminoacetal **4** in particular could thus be produced at the level of 10 grams.



Scheme 1. Modular built-up of target compounds capable of showing the N–CO interaction.



Scheme 2. Preparation of α-amino acid precursors before urea coupling; **Cbz** protection; reagents and conditions: (a) MeOH, H_2SO_4 cat, reflux, 12 h (**1**: 92%); (b) TBSCl, imidazole, DMF, rt, 12 h (**6**: 73% two steps); (c) DIBAL-H, CH_2Cl_2, −78 °C, 2 h; (d) MeOH, PTSA, rt, 12 h (**3**: 84% two steps); (e) TBSCl, imidazole, DMF, rt, 12 h (**9**: 52% three steps); and (f) H_2, 5% Pd/C, MeOH, 3 h (quant).

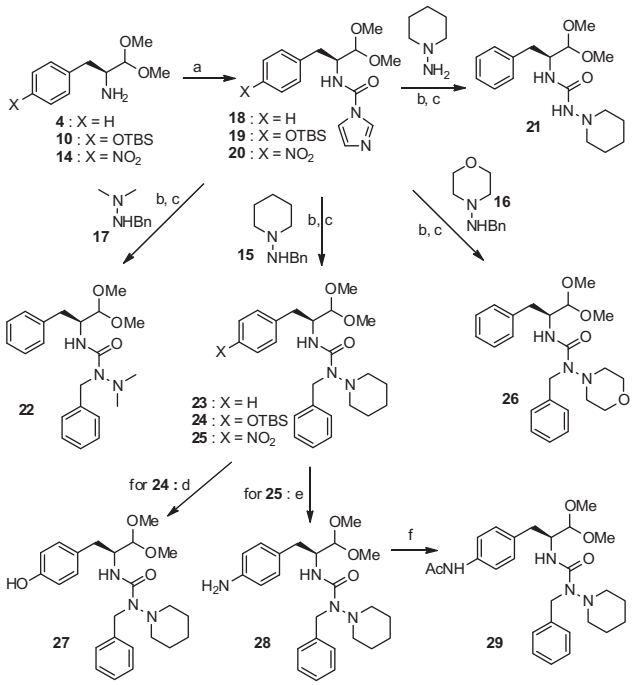


Scheme 3. Preparation of α-amino acid precursor before urea coupling; *Fmoc* protection; reagents and conditions: (a) MeOH, $SOCl_2$, reflux, 12 h (quant); (b) FmocCl, K_2CO_3, H_2O–dioxane, rt, 12 h (79%); (c) DIBAL-H, CH_2Cl_2, −78 °C, 2 h; (d) MeOH, PTSA, rt, 12 h (53% two steps); (e) piperidine, DMF, 40 min (80%).

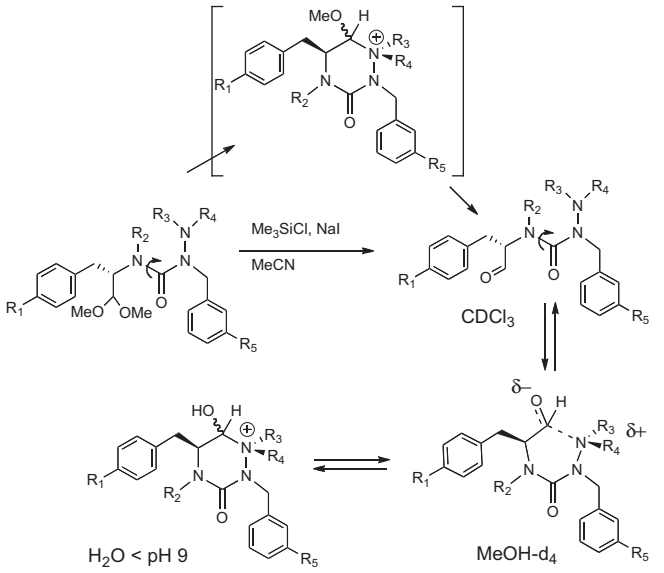
We were very much interested in extending the substituent on the amino aldehyde building block in order to possibly satisfy more remote binding pockets of the protease for higher affinities. However, as the introduction of a carboxamide junction in the para position of a phenylalanine side chain cannot be done prior to DIBAL-H reduction, this junction has to be established post-urea coupling by working initially with a nitro group. However, an aromatic nitro group is highly susceptible to catalytic hydrogenation, and we therefore chose to work in this particular case with a Fmoc protective group that is as resistant to DIBAL-H as is Cbz and that can be removed with piperidine instead of hydrogenation. The corresponding five-step synthesis of the aminoacetal precursor (**14**, Scheme 3) gives an overall yield of 33% from commercial unprotected *p*-nitrophenylalanine.

The second building block (Scheme 1) was varied so as to allow for the presentation of respectively, a cyclic *N*-piperidinyl substituent, a cyclic *N*-morpholinyl substituent, and an open-chain dimethylamine unit in the final target compounds. Hydrazines **15**,[22] **16**, and **17** were prepared by a two-step procedure consisting of imine formation from benzaldehyde followed by reduction with sodium cyanoborohydride (80% average yield over two steps for the cyclic substituents and 40% for the open-chain). These two sets of building blocks then served in the key coupling step towards a urea unit. The three-step protocol (Scheme 4) using carbonyl-di-imidazol (CDI) followed by activation/quaternization of the mono-reacted intermediate with iodomethane and final reaction with the second building block (tri-substituted hydrazine) was previously reported by us.[22] We thus prepared five out of the eight variants contained in this work (ureic acetals **21**, **22**, **26**, **27** and **29**) in yields ranging from 43% to 71%.

Initially, we designed our cyclic ureas in accordance with insights gained from the inhibitor DMP450 (see Section 1). However, up to now, our three-step coupling protocol did not allow for the preparation of a derivative of our N–CO interacted hydrazino ureas that displays benzyl groups on *both* nitrogens of the urea motif as

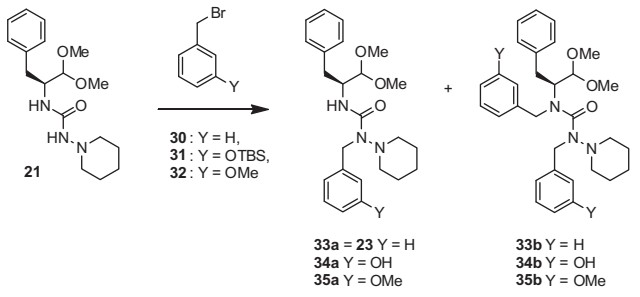


Scheme 4. Protected precursors of target compounds capable of showing the N–CO interaction; reagents and conditions: (a) CDI, THF, rt, 15 min; (b) MeI, MeCN, rt, 12 h; (c) hydrazine, NEt_3, CH_2Cl_2, rt, 12 h; (d) TBAF, THF, rt, 10 min; (e) H_2, 5% Pd/C, MeOH, 2 h; Ac_2O, Et_3N, DCM, rt, 30 min.



Scheme 6. Deprotection of all acetal precursors to hydrazino-aldehydes, constitutional equilibrium in different solvents, and hypothetical form in neutral aqueous media.

seen in DMP450 (Scheme 7); we believe steric hindrance to be responsible for the lack of any reactivity of a secondary amine contained in the aminoacetal (rather than a primary one as seen in Scheme 4) towards the quaternized imidazolyl urea intermediate. Even our measure of switching our coupling reagent from CDI to carbonyl-di-triazol did not trigger the reaction. In the present report, we therefore opted for the simultaneous introduction of two identical pendent arms starting from a urea with only two substituents (**21**, Scheme 5). Use of sodium hydride and an excess of substituted or unsubstituted benzyl bromide did indeed yield small amounts of the tetra-substituted ureas (**33b**, **34b**, **35b**) alongside tri-substituted material. **33b** and **35b** were isolated and characterized while the formation of **34b** was only confirmed by LCMS monitoring. The exclusive presence of tri-substituted ureas having received a benzyl substituent in the hydrazine portion proves that alkylation of this nitrogen is faster than the one on the opposite side of the urea unit. For the synthesis of **34a** and **34b**, a protected version of meta-hydroxy benzyl bromide proved necessary in order to resist the harsh condition of the substitution.



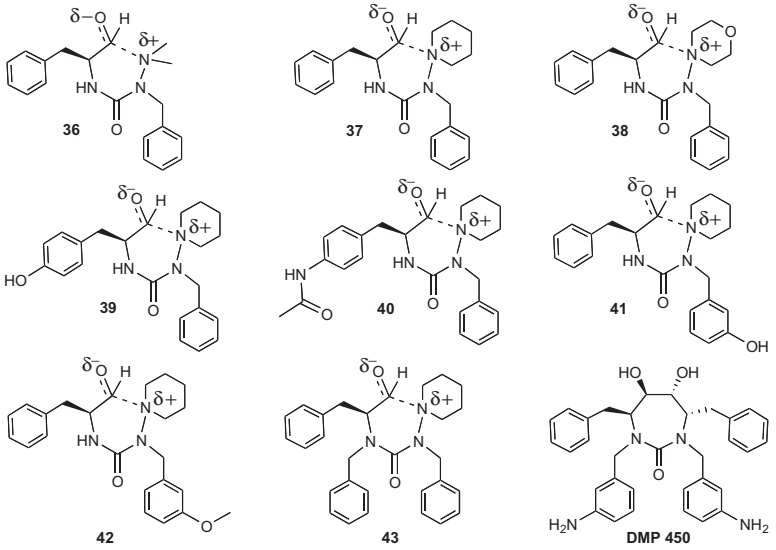
Scheme 5. Outcome of a double alkylation reaction to reach an exhaustively derivatized urea; reagents and conditions: 6 equiv bromide compound, 20 equiv NaH, DMF, rt, 48 h.

A simple TBS protected version was chosen and synthesized from commercial 3-hydroxybenzoate.[28]

Acetals obtained by this process can be stored indefinitely, while their corresponding free aldehydes run the risk of racemization over longer storage times. In order to prepare our target aldehydes for study we ran the deprotection (**35b** was not deprotected in view of the small amounts available) in anhydrous media using Me_3SiI generated in situ from Me_3SiCl and sodium iodide (Scheme 6). While six out of eight targets were obtained as oils, two turned out to be solids (**39** and **41** that have as a common feature a phenol substituent). LCMS monitoring of the deprotection reactions revealed the formation of an intermediate that could in two cases be isolated: a quaternized N/O acetal (Scheme 6). Its salt-like character may be responsible for the total retention of the two products at the base of a silicagel column if chromatographic purification is attempted. The existence of this unusual species may be explained with the high pre-organization that we deliberately chose for this system. Aqueous workup however hydrolyzes it to the desired aldehyde. This tendency to halt cleavage at the quaternized N/O acetal is most pronounced for the exhaustively substituted targets **33b** further proof for our strategy of maximum pre-organization in order to achieve 100% cyclization and the absence of free aldehyde in physiological media.

2.2. Folding of deprotected 39

As previously reported, the NCO interaction is observable only in polar protic solvents. In the presence of a urea moiety, it has been shown that the adoption of the required configuration for NCO interaction requires time. In deuterated chloroform the proton NMR spectrum for compound **39** corresponds solely to the non-interacted form, which is confirmed by the presence of a characteristic aldehyde signal at 9.6 ppm (Fig. 1). On the other hand, this aldehydic signal disappears over time when **39** is dissolved in deuterated methanol; a new signal around 4.1 ppm takes the place of that at 9.6 ppm. The kinetics are slow enough to be monitored by NMR. This demonstrates that about 40% is cyclized after 25 min at room temperature, 60% conversion is reached after 35 min, and cyclization is complete in 4 h. In view of the similarity of the structure to that of the benchmark structure published by Waibel et al. (**37**) it was not surprising that **39** also folds to a



Scheme 7. Formula of all eight target compounds as present in methanol, including **37** (reported previously) and DMP 450.

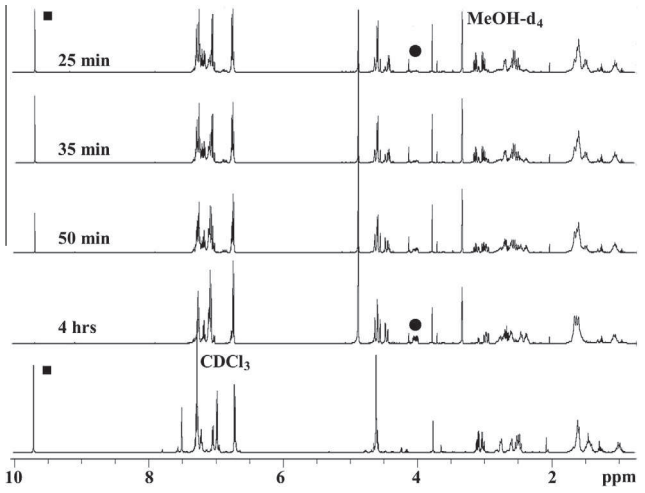


Fig. 1. 1H spectra (500 MHz) of 39 in $CDCl_3$, and at certain time intervals beginning with initial dissolution in MeOH-d_4. (■) aldehydic proton, (●) proton of the NCO carbonyl moiety.

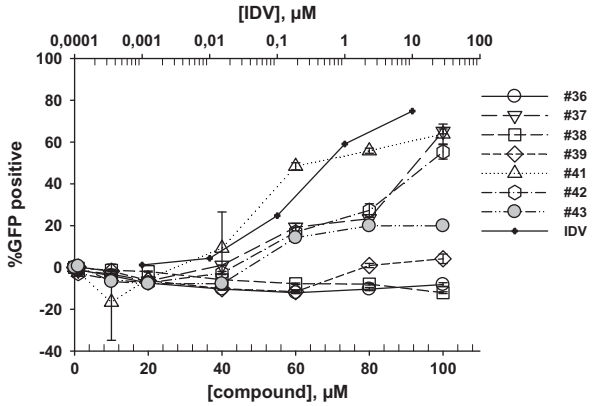


Fig. 2. Assay response to inhibitors. SupT1 HIV-1 protease-Gal4 cells were grown for 24 h in the presence of 1–100 µM of compounds **36–43** or 0.0001–10 µM Indinavir (IDV) as reference inhibitor, and then activated with 1 mg/mL doxocycline. eGFP fluorescence was quantified 48 h later by flow cytometry.

100%. This cyclization process is as reversible as has been reported for **37**: MeOH evaporation followed by a further dissolution in $CDCl_3$ gives rise to a proton NMR spectrum in full congruence with open chain form.

2.3. Inhibition assays of HIV-1 Protease activity in T-Cells

To investigate the inhibitory activity of our compounds, we utilized a new system for monitoring the HIV-1 protease activity in T-cells, natural targets of HIV.[29] In this system, the HIV-1 protease is fused to Gal4 and placed under the control of a tetracycline-inducible promoter. When expressed in T cells, the fusion protein binds to a Gal4 Responsive Element which controls the expression of a gene coding for Enhanced Green Fluorescent Protein (eGFP). When the HIV-1 protease is active, it prevents the binding of GAL4. The more the HIV-1 protease activity is inhibited by our molecule candidates, the more eGFP is expressed. Cells are grown for 24 h in the presence of inhibitor, and then eGFP fluorescence is quantified by flow cytometry. Indinavir is used as positive control; as shown in Fig. 2 (black circles), this inhibitor triggers eGFP expression via HIV-1 protease inhibition at concentrations as low as 0.1 µM.

The same experiments carried out with compounds **36**–**43** reveal an HIV-1 protease inhibition pattern with **37**, **41** and **42** above 20 µM. The other candidates did not display significant inhibition activity. For comparison, in vitro inhibition data obtained by a classic, FRET-based spectrometer assay was included in the Supporting information. In cellulo assays contribute an extra level of reliability to the inhibition data in view of the variable quality of purified protease samples.

2.4. Cytotoxicity and ABC pumps efflux

One obstacle to efficient chemotherapy and antiretroviral therapeutic is the multidrug resistance (MDR) phenotype developed by cells to prevent cytotoxicity. This mechanism involves ATP-binding cassette (ABC) transporters[30] such as P-gp, MRP1[31] and BCRP.[32–34] Most of antiretroviral drugs used to treat HIV-infected patients are unfortunately substrates of such efflux pumps.[35,36] We thus evaluated the efflux of the present compounds by these pumps by looking for the compounds cytotoxicity and a cross-resistance induced by P-gp or BCRP expression (Table 1).

Table 1
Compounds cytotoxicity. Cell survival was estimated by MTT assays as described in the experimental section. Half-maximal cytotoxic concentrations, IC_{50} μM, were the mean ± sd of at least three independent experiments

#	P-gp		BCRP	
	Expressing cells (%)	Control cells (%)	Expressing cells (%)	Control cells (%)
Saquinavir	49[37]	37[37]	—	—
36	100 ± 1.8	68.8 ± 2.1	57.2 ± 0.8	56.3 ± 0.9
37	48.6 ± 0.8	42.6 ± 0.9	34.8 ± 2.1	37.6 ± 1.1
38	—	79.8 ± 1.2	—	—
39	71.2 ± 1.7	58.6 ± 1.6	47.9 ± 1.6	52.4 ± 0.4
41	43.5 ± 0.7	42.1 ± 1.5	23.1 ± 1.1	35.3 ± 1.5
42	46.8 ± 0.9	39.4 ± 0.7	37.2 ± 0.9	42.6 ± 0.7
43	12.8 ± 0.4	11.2 ± 0.2	12.8 ± 0.2	17.6 ± 0.4

As shown, these compounds displayed either no (**38**), limited (>30–50 μM, **36**, **37**, **39**, **41** and **42**) or rather high (∼10 μM, **43**) cytotoxicity. Cytotoxicity of the more active compounds **37**, **41** and **42** (Fig. 2) is in the same range of concentrations showing that the inhibition power needs to be improved and the cytotoxicity reduced.

The specific expression of P-gp or BCRP did not change the cytotoxicity pattern of all products, indicating that multidrug resistance pumps do not transport this type of compound, contrary to what is being observed for established, clinically applied HIV-1 protease inhibitors such as Saquinavir (Table 1) or Indinavir for which Pgp confers a twofold resistance ratio.[38] Finally, we checked if such compounds could inhibit the transport activity of the same ABC pumps by looking at their effect on the efflux of mitoxantrone mediated by P-gp or BCRP (Supporting information Table 1). We could not observe any effect of them on the mitoxantrone efflux suggesting that they probably are not inhibitors of such pumps.

3. Conclusion

In this report, we significantly improved synthetic access to the previously reported **37** that allows the investigator to have access to this compound at the gram scale. We demonstrated the quick access to a number of substitution variants and thus the modularity of our synthetic route. Our in cellulo assay of HIV-1 PR activity allowed us to position the inhibitory potency of our molecule candidates with respect to two established and clinically applied inhibitors; the values for our candidates have to be interpreted in light of the considerably reduced size compared to the benchmark inhibitors (60%) and the associated decrease in interacting surface. Among our new molecule variants, **41** and **42** are promising hits, as they display a significant inhibition potential and toxicity profile that encourage further refinement. Significantly, they are not transported out of cells by multidrug efflux pumps, quite in contrast to clinically applied inhibitors. Future attempts to obtain an X-ray diffraction structure of an enzyme-inhibitor complex may furnish the molecular insights to shift the inhibitory potency into the nanomolar range.

4. Experimental

4.1. Synthesis

Further synthetic protocols and analytical data of the target compounds are contained in the Supplementary information.

4.1.1. (S)-(1-(4-tert-Butyldimethylsilyloxy)benzyl-2-oxo-ethyl)-carbamic acid benzyl ester (7)

To a cold (−78 °C) solution of **6** (10.3 g, 23.2 mmol) in dry CH_2Cl_2 (100 mL) was added dropwise a 1.0 M solution of DIBAL-H in CH_2Cl_2 (50 mL, 50 mmol) over 40 min. The reaction mixture, monitored by TLC, was quenched at −78 °C with MeOH (30 mL) after 2 h and then warmed to room temperature. The mixture was poured into 300 mL of an ice cold 1.2 M HCl aqueous solution, extracted twice with 300 mL of CH_2Cl_2 and the combined organic phases were then washed with 300 mL brine, dried over $MgSO_4$ and concentrated under vacuum. Because of the presence of a chiral center in alpha of the aldehyde, the next step was performed immediately with the crude product.

4.1.2. (S)-(1-(4-Hydroxy)benzyl-2,2-dimethoxy-ethyl)-carbamic acid benzyl ester (8)

The crude **7** (23.2 mmol) was dissolved in 180 mL of anhydrous MeOH, and *p*-toluenesulfonic acid monohydrate (3.0 g, 15.8 mmol) was added. The reaction was stirred overnight, and then most of the solvent was removed under vacuum. 100 mL of a solution of aqueous saturated $NaHCO_3$ was added, and the mixture was extracted twice with 100 mL of CH_2Cl_2. The combined organic layers were dried over $MgSO_4$, and the solvent was evaporated under vacuum. Purification by flash chromatography (gradient 1:9 to 3:7, EtOAc:Cyclohexane) gave **8** as a white solide (5.67 g, 61%). ^1H NMR (500 MHz, $CDCl_3$) δ 7.30 (m, 5H), 7.04 (d, J = 8.0 Hz, 2H), 6.73 (d, J = 8.0 Hz, 2H), 5.07 (m, 2H), 4.20 (d, J = 2.8 Hz, 1H), 4.09 (m, 1H), 3.45 (s, 3H), 3.40 (s, 3H), 2.87 (dd, J = 5.8 Hz, J_{AB} = 14.1 Hz, 1H), 2.71 (dd, J = 5.8 Hz, J_{AB} = 14.1 Hz, 1H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 156.5, 154.8, 136.4, 130.3, 128.5, 128.1, 127.9, 115.4, 104.9, 68.9, 55.8, 55.7, 52.4, 35.3. HRMS (ESI) calcd for $C_{19}H_{23}NNaO_5$ $[M+Na]^+$ 368.1468, found 368.1463.

4.1.3. General coupling procedure

A solution of amine in anhydrous THF (1 mL per mmol) was added dropwise to a stirred suspension of CDI (1.1 equiv) in anhydrous THF (1 mL per mmol) at room temperature. After 30 min the solvent was evaporated, and the residue redissolved in CH_2Cl_2 (5 mL per mmol) and washed twice with water (2 mL per mmol). The organic phase was dried over $MgSO_4$, and the solvent evaporated under vacuum to give pure imidazole carboxamide.

The imidazole carboxamide compound was dissolved in anhydrous acetonitrile (1.5 mL per mmol) and iodomethane (4 equiv) was added. The reaction was stirred overnight at room temperature, then the solvent was evaporated, and the yellow orange oil was dried under vacuum. The residue was redissolved in dry CH_2Cl_2 (5 mL per mmol), and the hydrazine (1 equiv) and triethylamine (1 equiv) were added. After stirring for 24 h at room temperature, the reaction was quenched by adding saturated aqueous $NaHCO_3$ (5 mL per mmol), and the mixture was extracted CH_2Cl_2 (twice with 5 mL per mmol). The combined organic extracts were dried over $MgSO_4$, and the solvent was removed under vacuum. After purification by flash chromatography (1:4 to 1:1, EtOAc:Cyclohexane), the urea compound was obtained.

4.1.4. (S)-Imidazole-1-carboxylic Acid (1-(p-tert-butyldimethylsily-loxy)benzyl-2,2-dimethoxy-ethyl)-amide (19)

Yellow oil, yield: 86%. ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.06 (s, 1H), 7.29 (m, 1H), 7.09 (m, 3H), 6.79 (d, J = 8.2 Hz, 2H), 4.35 (m, 1H), 4.26 (d, J = 2.7 Hz, 1H), 3.49 (s, 3H), 3.40 (s, 3H), 2.93 (dd, J = 7.0 Hz, J_{AB} = 14.4 Hz, 1H), 2.90 (dd, J = 7.6 Hz, J_{AB} = 14.4 Hz, 1H), 0.98 (s, 9H), 0.19 (s, 6H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 154.5, 148.6, 135.9, 130.5, 130.1, 129.6, 120.3, 115.8, 104.1, 56.0, 55.7, 53.7, 35.3, 25.7, 18.2, −4.4.

4.1.5. (S)-1-Benzyl-3-(1-(p-tert-butyldimethylsilyloxy)benzyl-2,2-dimethoxy-ethyl)-1-piperidin-1-yl-urea (24)

From **19** + **15**. Colorless oil, yield: 82%. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.27 (m, 2H), 7.19 (m, 3H), 7.12 (d, J = 8.5 Hz, 2H), 6.78 (d, J = 8.5 Hz, 2H), 6.64 (d, J = 9.0 Hz, 1H; NH), 4.62 (d, J_{AB} = 16.4 Hz, 1H), 4.50 (d, J_{AB} = 16.4 Hz, 1H), 4.31 (d, J = 3.7 Hz, 1H), 4.28 (m, 1H), 3.48 (s, 3H), 3.47 (s, 3H), 3.00 (dd, J = 5.0 Hz, J_{AB} = 13.8 Hz, 1H), 2.74 (m, 2H), 2.50 (m, 2H), 2.39 (m, 1H), 1.63 (m, 3H), 1.51 (m, 2H), 1.01 (m, 10H), 0.19 (s, 6H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 157.9, 154.0, 140.6, 131.2, 130.4, 128.3, 127.3, 126.5, 119.8, 105.8, 56.2, 54.9, 53.5, 53.2, 52.2, 42.1, 34.7, 26.5, 25.7, 18.2, −4.4. HRMS (ESI) calcd for $C_{30}H_{47}N_3NaO_4Si$ $[M+Na]^+$ 564.3228, found 564.3221.

4.1.6. General procedure for the deprotection of all acetal precursors to hydrazino-aldehydes

To a solution of the acetal precursor (63 mg, 0.14 mmol) in anhydrous acetonitrile (20 mL per mmol) were added NaI (3 equiv) and TMSCl (2 equiv). The reaction was stirred at room temperature for 1.5 h and was then quenched by adding $NaHCO_3$ (saturated aqueous solution, 20 mL per mmol). The mixture was extracted with CH_2Cl_2 (three times 60 mL per mmol), and the combined organic extracts were washed with $Na_2S_2O_3$ (saturated solution, 100 mL per mmol) and then with water (60 mL per mmol). The organic phase was dried over Na_2SO_4 and the solvent was evaporated under vacuum to give the crude aldehyde.

4.1.7. (S)-1-Benzyl-3-(1-(4-hydroxy)benzyl-2,2-dimethoxy-ethyl)-1-piperidin-1-yl-urea (27)

Compound **26** (415 mg, 0.77 mmol) was dissolved in 25 mL THF and 1.5 mL of a 1 M solution of tetra-*n*-butylammonium fluoride in THF was added dropwise. The reaction was stirred for 10 min and the solvent was evaporated. The residue was redissolved in CH_2Cl_2 (20 mL) and washed with water (20 mL). The aqueous phase was extracted with CH_2Cl_2 (10 mL), and the combined organic phase was dried over $MgSO_4$ and concentrated under vacuum. After purification by flash chromatography (gradient 1:4 to 1:1, EtOAc:Cyclohexane), **27** (332 mg, quant) was obtained as a colorless gel-oil. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.25 (m, 2H), 7.17 (m, 3H), 7.02 (d, J = 8.2 Hz, 2H), 6.77 (d, J = 9.1 Hz, 1H; NH), 6.69 (d, J = 8.5 Hz, 2H), 4.61 (d, J_{AB} = 16.0 Hz, 1H), 4.54 (d, J_{AB} = 16.0 Hz, 1H), 4.29 (d, J = 3.3 Hz, 1H), 4.26 (m, 1H), 3.48 (s, 3H), 3.46 (s, 3H), 2.95 (dd, J = 5.0 Hz, J_{AB} = 14.0 Hz, 1H), 2.70 (m, 2H), 2.51 (m, 2H), 2.39 (m, 1H), 1.63 (m, 3H), 1.53 (m, 2H), 1.01 (m, 1H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 158.3, 155.3, 140.2, 130.2, 129.3, 128.3, 127.2, 126.6, 115.4, 105.9, 56.2, 55.3, 53.4, 53.3, 52.8, 42.3, 35.0, 26.4. HRMS (ESI) calcd for $C_{24}H_{33}N_3NaO_4$ $[M+Na]^+$ 450.2363, found 450.2355.

4.1.8. (S)-1-Benzyl-3-(1-(4-hydroxy)benzyl-2-oxo-ethyl)-1-piperidin-1-yl-urea (39)

From acetal **27** (32 mg, 0.07 mmol). Purification by flash chromatography (gradient 1:4 to 1:1, EtOAc:Cyclohexane), gave 39 as a colorless oil (17 mg, 59%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.71 (s, 1H; CHO), 7.28 (m, 5H; Ar_{Phenyl}), 7.04 (d, J = 7.8 Hz, 1H; NH), 6.98 (d, J = 8.6 Hz, 2H; Ar_{Phenol}), 6.71 (d, J = 8.6 Hz, 2H; Ar_{Phenol}), 4.61 (m, 3H; NCH_2Ph + NHC*H*), 3.10 (dd, J = 6.2 Hz, J_{AB} = 14.0 Hz, 1H; CH_AH_BPhenol), 3.02 (dd, J = 7.0 Hz, J_{AB} = 14.0 Hz, 1H; CH_AH_BPhenol), 2.75 (m, 1H; $NNCH_{2piperidine}$), 2.60 (m, 1H; $NNCH_{2pip}$), 2.50 (m, 2H; $NNCH_{2pip}$), 1.61 (m, 3H; CH_{2pip}), 1.45 (m, 2H; CH_{2pip}), 1.00 (m, 1H; CH_{2pip}). ^{13}C NMR (127 MHz, $CDCl_3$) δ 200.9 (CHO), 158.3 (NCON), 155.7 ($Ar_{Phenol}COH$), 139.8 ($Ar_{Phenyl}CCH_2N$), 130.3 (Ar_{Phenol}), 128.4 ($2Ar_{Phenyl}$), 127.4 ($2Ar_{Phenyl}$), 126.8 (Ar_{Phenyl}), 115.7 (Ar_{Phenol}), 60.5 (NHCH), 53.6 ($NNCH_{2pip}$), 53.4 ($NNCH_{2pip}$), 42.5 (NCH_2Ph), 34.7 (CH_2Phenol), 26.4 (CH_{2pip}), 26.3 (CH_{2pip}), 23.1 (CH_{2pip}). HRMS (ESI) calcd for $C_{22}H_{27}N_3NaO_3$ $[M+Na]^+$ 404.1945, found 404.1941.

4.2. Cell culture

Human T-cell line SupT1, PR/Gal4 and Gal4 was from Dr. R. Wolkowicz, San Diego, California, USA. Cells were maintained in complete RPMI 1640 media with Stable Glutamine (PAA laboratories) supplemented with 10% fetal bovine serum (PAA laboratories), 1% penicillin/streptomycin (PAA laboratories). Cells were incubated at 37 °C in humidified 5% CO_2.

The NIH3T3, drug-sensitive, parental cell line and that transfected with human MDR1-G185, NIH3T3 Pgp,[39] derived from NIH Swiss mouse embryo cultures, were from ATCC and used as previously described.[40] Human embryonic kidney (HEK293) cell lines transfected with either BCRP (HEK293-BCRP) or the empty vector (HEK293-pcDNA3) were obtained as previously described.[38] Cells were maintained in Dulbecco's modified Eagle's medium (DMEM high glucose, PAA laboratories), supplemented with 10% fetal bovine serum (FBS, PAA laboratories), 1% penicillin/streptomycin (PAA laboratories). The NIH3T3 Pgp growth medium was supplemented with 60 ng/mL colchicine. The HEK293-pcDNA3 and HEK293-BCRP was supplemented with 0.75 mg/mL G418 (PAA laboratories). Cells were incubated at 37 °C in humidified 5% CO_2.

4.3. Flow cytometry

4.3.1. HIV-1 protease activity in T-cells

Doxycycline (Sigma Aldrich) was dissolved in water at 10 mg/mL stock concentration and stored at −20 °C. Indinavir was a gift from Merck and Saquinavir from Roche. They were dissolved in 100% DMSO at 10 mM stock concentration and stored at −20 °C.

In a 96 well plate, 50,000 cells were incubated for 10 min in growth medium and then inhibitors were added at varying concentrations and incubated for at least 10 min. Cells were expressing the HIV-1–eGFP fusion protein with 1 μg/ml doxycycline. After 48 h, cells were washed with PBS and the expression of eGFP was quantified by flow cytometry carried out with a FACS Calibur cytometer (Becton Dickinson). Excitation and emission were set up at 488 and 530 nm, respectively. Data were collected on CellQuest Pro (version 4.0) software and then exported to FlowJo for analysis.

4.4. Cytotoxicity assays

Parental cell line (NIH3T3 and HEK293-pcDNA3) and transfected with each pump (NIH2T3 Pgp and HEK293-BCRP) were incubated for 72 h in the presence of increasing concentrations of the novel HIV-1 protease inhibitors, and then cells survival was estimated by a 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) (Sigma–Aldrich) colorimetric assay as detailed previously.[38]

Acknowledgments

We thank T. ANDRIEU and S. DUSSURGEY (Tour Inserm CERVI, Lyon France) for their help with the flow cytometer analysis. We would like to acknowledge the kind supply with indinavir by Merck Sharp & Dohme and GF120918 by GlaxoSmithKline.

This work was supported by the Centre National de la Recherche Scientifique (CNRS) and University of Lyon 1 (UMR5086), and the French Ministry of Research (EA3741, Lyon 1). It was funded by the National Research Agency (ANR-06-BLAN_0420, ANR-06-PCVI-0019-01, ANR piribio09_444706), the Association pour la Recherche sur le Cancer (ARC), and the Ligue Nationale Contre le Cancer (Labellisation 2009–14).

G.G. and L.M. were recipients of doctoral fellowships from the Ministère de la Recherche (France) and the Région Rhône-Alpes (France), respectively.

Supplementary data

Supplementary data associated (further synthetic protocols, NMR data, protocol for ABC pump inhibition) with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bmc.2013.06.018.

References and notes

[1]. Tyndall, J. D. A.; Nall, T.; Fairlie, D. P. *Chem. Rev.* **2005**, *105*, 973.
[2]. Brik, A.; Wong, C. H. *Org. Biomol. Chem.* **2003**, *1*, 5.
[3]. Bursavich, M. G.; Rich, D. H. *J. Med. Chem.* **2002**, *45*, 541.
[4]. De Clercq, E. *Nat. Rev. Drug Disc.* **2007**, *6*, 1001.
[5]. Gautier, A.; Pitrat, D.; Hasserodt, J. *Bioorg. Med. Chem.* **2006**, *14*, 3835.
[6]. Waibel, M.; Pitrat, D.; Hasserodt, J. *Bioorg. Med. Chem.* **2009**, *17*, 3671.
[7]. Hasserodt, J.; Gautier, A.; Barbe, R.; Waibel, M. Attanazi, O. A.; Spinelli, D., Eds.; Targets in Heterocyclic Systems, Societa Chimica Italiana: Rome, 2009; Vol. 13, pp 1–26.
[8]. Leonard, N. *Rec. Chem. Prog.* **1956**, *17*, 243.
[9]. Leonard, N. J. *Acc. Chem. Res.* **1979**, *12*, 423.
[10]. Zhang, X. A.; Song, D. T.; Lippard, S. J. *J. Org. Chem.* **2008**, *73*, 734.
[11]. Gadamer, J. *Arch. Pharm.* **1920**, *258*, 148.
[12]. Kermack, W.; Robinson, R. *J. Chem. Soc.* **1922**, *121*, 427.
[13]. Huisgen, R.; Wieland, H.; Eder, H. *Liebigs Ann. Chem.* **1949**, *561*, 193.
[14]. Boit, H. G.; Paul, L. *Chem. Ber. Recl.* **1954**, *87*, 1859.
[15]. Birnbaum, G. I. *J. Am. Chem. Soc.* **1974**, *96*, 6165.
[16]. Cimino, G.; Destefano, S.; Scognamiglio, G.; Sodano, G.; Trivellone, E. *Bull. Soc. Chim. Belg.* **1986**, *95*, 783.
[17]. Becker, M. H.; Chua, P.; Downham, R.; Douglas, C. J.; Garg, N. K.; Hiebert, S.; Jaroch, S.; Matsuoka, R. T.; Middleton, J. A.; Ng, F. W.; Overman, L. E. *J. Am. Chem. Soc.* **2007**, *129*, 11987.
[18]. Leonard, N. J.; Fox, R. C.; Oki, M. *J. Am. Chem. Soc.* **1954**, *76*, 5708.
[19]. Hine, J.; Kokesh, F. C. *J. Am. Chem. Soc.* **1970**, *92*, 4383.
[20]. McCrindle, R.; McAlees, A. J. *J. Chem. Soc., Chem. Commun.* **1983**, 61.
[21]. Kirby, A. J.; Komarov, I. V.; Bilenko, V. A.; Davies, J. E.; Rawson, J. M. *Chem. Commun.* **2002**, 2106.
[22]. Waibel, M.; Hasserodt, J. *J. Org. Chem.* **2008**, *73*, 6119.
[23]. DeLucca, G. V.; Ericksonviitanen, S.; Lam, P. Y. S. *Drug Discovery Today* **1997**, *2*, 6.
[24]. Nair, A. C.; Jayatilleke, P.; Wang, X.; Miertus, S.; Welsh, W. J. *J. Med. Chem.* **2002**, *45*, 973.
[25]. Dunitz, J. D. *Science* **1994**, *264*, 670.
[26]. Jurczak, J.; Golebiowski, A. *Chem. Rev.* **1989**, *89*, 149.
[27]. Dinh, T. Q.; Armstrong, R. W. *J. Org. Chem.* **1995**, *60*, 8118.
[28]. Kwon, S.; Myers, A. G. *J. Am. Chem. Soc.* **2005**, *127*, 16796.
[29]. Hilton, B. J.; Wolkowicz, R. *PLoS ONE* **2010**, *5*, e10940.
[30]. Dean, M.; Hamon, Y.; Chimini, G. *J. Lipid Res.* **2001**, *42*, 1007.
[31]. Cole, S. P.; Bhardwaj, G.; Gerlach, J. H.; Mackie, J. E.; Grant, C. E.; Almquist, K. C.; Stewart, A. J.; Kurz, E. U.; Duncan, A. M.; Deeley, R. G. *Science* **1992**, *258*, 1650.
[32]. Allikmets, R.; Schriml, L. M.; Hutchinson, A.; Romano-Spica, V.; Dean, M. *Cancer Res.* **1998**, *58*, 5337.
[33]. Doyle, L. A.; Yang, W.; Abruzzo, L. V.; Krogmann, T.; Gao, Y.; Rishi, A. K.; Ross, D. D. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 15665.
[34]. Miyake, K.; Mickley, L.; Litman, T.; Zhan, Z.; Robey, R.; Cristensen, B.; Brangi, M.; Greenberger, L.; Dean, M.; Fojo, T.; Bates, S. E. *Cancer Res.* **1999**, *59*, 8.
[35]. Weiss, J.; Haefeli, W. E. In *International Review of Cell and Molecular Biology*; Jeon, K. W., Ed.; Academic Press, 2010; Vol. 280, pp 219–279. T2.
[36]. Kis, O.; Robillard, K.; Chan, G. N. Y.; Bendayan, R. *Trends Pharmacol. Sci.* **2010**, *31*, 22.
[37]. Kim, A. E.; Dintaman, J. M.; Waddell, D. S.; Silverman, J. A. *J. Pharmacol. Exp. Ther.* **1998**, *286*, 1439.
[38]. Kim, R. B.; Fromm, M. F.; Wandel, C.; Leake, B.; Wood, A. J.; Roden, D. M.; Wilkinson, G. R. *J. Clin. Invest.* **1998**, *101*, 289.
[39]. Cardarelli, C. O.; Aksentijevich, I.; Pastan, I.; Gottesman, M. M. *Cancer Res.* **1995**, *55*, 1086.
[40]. Arnaud, O.; Koubeissi, A.; Ettouati, L.; Terreux, R.; Alamé, G.; Grenot, C.; Dumontet, C.; Di Pietro, A.; Paris, J.; Falson, P. *J. Med. Chem.* **2010**, *53*, 6720.

Supporting Information

Modular construction of N-CO-interacted inhibitor candidates and their *in cellulo* assessment with HIV-1 protease

Guillaume Gros[a] , Lorena Martinez[b] , Anna Servat Gimenez[a] , Paula Adler[a,c] , Philippe Maurin[a] , Pierre Falson[b] and Jens Hasserodt[a,] *

[a]*Laboratoire de Chimie, Universite de Lyon – ENS, 46 allée d'Italie, Lyon 69364, France*
[b]*IBCP, rue Vercors, Lyon 67000, France*
[c]*on leave by the University of Ottawa, Canada*

* Corresponding author. Tel.: +33-472-728394; fax: +33-472-728860; e-mail: jens.hasserodt@ens-lyon.fr

Table des matières

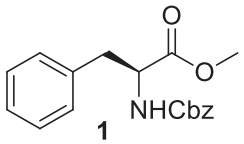
1. Chemistry

1.1. General procedures

All reactions were performed in anhydrous solvents and under argon atmosphere unless stated otherwise. Reaction progress was monitored by thin-layer chromatography (TLC) using Merck silica gel 60 aluminium sheets. The plates were visualized by either UV light (254 nm), or by a solution of phosphomolybdic acid. Column chromatography was performed using Merck silica gel Si 60 (40-63 µm). ^{1}H and ^{13}C NMR spectra were obtained at room temperature with a Bruker 500 MHz spectrometer. Chemical shifts (δ) are reported in ppm (s = singlet, d = doublet, t = triplet, m = multiplet, br = broad) and referenced from tetramethyl silane or from solvent references. NMR coupling constants (*J*) are reported in Hertz.

1.2. Synthesis of amino-acetal 4

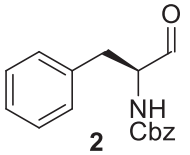
N-Cbz-L-phenylalanine Methyl Ester (1)



1

To a solution of N-Cbz-L-phenylalanine (8.3 g, 27.8 mmol) in 100 mL of anhydrous MeOH was added 1 mL of concentrated sulfuric acid and the resulting mixture was heated at reflux overnight. The solution was then cooled to room temperature and concentrated under reduced pressure. The residue was diluted in 50 mL Et_2O was added and washed successively with 5% aqueous $NaHCO_3$, brine, and finally dried over $MgSO_4$. After drying under vacuum **1** was obtained as a pale yellow resin (8.0 g, 92 %). ^{1}H NMR (200 MHz, $CDCl_3$) δ 7.1-7.3 (m, 10H), 5.30 (d, *J* = 8.2 Hz, 1H; NH), 5.13 (s, 2H), 4.69 (m, 1H), 3.74 (s, 3H), 3.14 (m, 2H). ^{13}C NMR (50.3 MHz, $CDCl_3$) δ 172.0, 155.7, 136.3, 135.7, 129.3, 128.7, 128.6, 128.2, 128.1, 127.2, 67.0, 54.9, 52.4, 38.2.

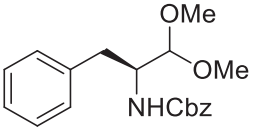
(S)-(1-benzyl-2-oxo-ethyl)-carbamic Acid Benzyl Ester (2)



2

To a cold (-78 °C) solution of 7 (6.8 g, 21.7 mmol) in dry CH_2Cl_2 (100 mL) was added dropwise a 1.0 M solution of DIBAL-H in CH_2Cl_2 (45 mL, 45 mmol) over 1 h. The reaction mixture, monitored by TLC, was quenched at -78°C with MeOH (30 mL) after 2 hrs and then warmed to room temperature. The mixture was poured into 250 mL of an ice cold 1.2 M HCl aqueous solution, extracted twice with 200 mL of CH_2Cl_2 and the combined organic phases were then washed with 300 mL brine, dried over $MgSO_4$ and concentrated under vacuum. Because of the presence of a chiral center in alpha of the aldehyde, the next step was performed immediately with the crude product.

(S)-(1-benzyl-2,2-dimethoxy-ethyl)-carbamic Acid Benzyl Ester (3)

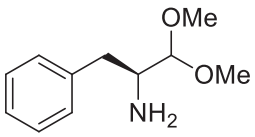


3

The crude **2** (23.2 mmol) was dissolved in 160 mL of anhydrous MeOH, and *p*- toluenesulfonic acid monohydrate (3.0 g, 15.8 mmol) was added. The reaction was stirred overnight, and then most of the solvent was removed under vacuum. 100 mL of a solution of aqueous saturated $NaHCO_3$ was added, and the mixture was extracted twice with 100 mL of CH_2Cl_2. The combined organic layers were dried over $MgSO_4$, and the solvent was evaporated under vacuum. Purification by flash chromatography (gradient 1:4 to 1:1, EtOAc: Cyclohexane) gave **3** as a white solide (5.67 g, 84 %). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.35-7.19 (m, 10H), 5.07 (d, J_{AB} = 12.4 Hz, 1H), 5.01 (d, J_{AB} = 12.4 Hz, 1H), 4.94 (d, *J* = 9.2 Hz, 1H), 4.17 (d, *J* = 3.3 Hz, 1H), 4.11-4.09 (m, 1H), 3.43 (s, 3H), 3.39 (s, 3H), 2.93 (dd, J_{AB} = 13.8, *J* = 6.1 Hz, 1H), 2.79 (dd,

J_{AB} = 13.8, J = 8.0 Hz, 1H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 155.9, 137.7, 136.5, 129.2, 128.3, 127.8, 127.7, 126.2, 104.6, 66.4, 55.6, 55.4, 53.4, 35.8.

(S)-1-benzyl-2,2-dimethoxyethan-1-amine (4)

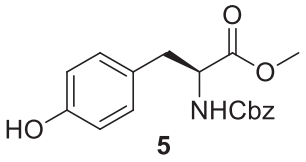


4

To a suspension of Pd on C (0.9 g, 5% wet Degussa type) in MeOH (50 mL) under hydrogen atmosphere, 2.18 g of 3 (6.6 mmol) were added and the reaction was stirred at room temperature. After 3 h, the reaction mixture was filtered on a Celite pad and the solvent was evaporated under reduced pressure to yield **4** (1.2 g, quant) as a yellow oil. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.32-7.21 (m, 5H), 4.11 (d, J = 5.7 Hz, 1H), 3.45 (s, 3H), 3.44 (s, 3H), 3.15-3.11 (m, 1H), 2.97 (dd, J_{AB} = 13.5, J = 4.0 Hz, 1H), 2.51 (dd, J_{AB} = 13.5, J = 9.5 Hz, 1H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 138.7, 129.2, 128.3, 126.1, 107.5, 55.0, 54.8, 53.9, 38.4.

1.3. Synthesis of amino-acetal 10

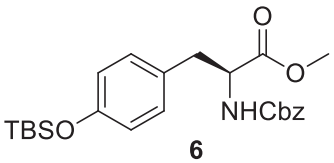
N-Cbz-L-tyrosine Methyl Ester (5)



5

To a solution of N-Cbz-L-tyrosine (10.0 g, 31.7 mmol) in 100 mL of anhydrous MeOH was added 1 mL of concentrated sulfuric acid and the resulting mixture was heated at reflux overnight. The solution was then cooled to room temperature and concentrated under reduced pressure. The residue was diluted in 30 mL Et_2O was added and washed successively with 5% aqueous $NaHCO_3$, brine, and finally dried over $MgSO_4$. After drying under vacuum **5** was obtained as a pale yellow resin (9.0 g, 86 %). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.35 (m, 5H), 6.96 (d, J = 8.2 Hz, 2H), 6.72 (d, J = 8.2 Hz, 2H), 5.69 (bs, 1H; OH), 5.29 (d, J = 8.2 Hz, 1H; NH), 5.14 (d, J_{AB} = 12.2, 1H), 5.10 (d, J_{AB} = 12.2, 1H), 4.65 (m, 1H), 3.75 (s, 3H), 3.08 (dd, J = 5.5 Hz, J_{AB} = 14.0 Hz, 1H), 3.02 (dd, J = 5.5 Hz, J_{AB} = 14.0 Hz, 1H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 172.3, 155.8, 155.0, 136.1, 130.4, 128.6, 128.3, 128.1, 127.4, 115.6, 67.1, 55.0, 52.4, 37.5.

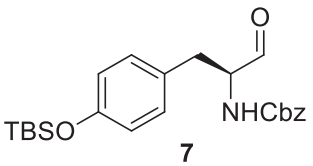
O-TBS-N-Cbz-L-tyrosine methyl ester (6)



6

To a solution of N-Cbz-L-tyrosine methyl ester **5** (9.0 g, 27.3 mmol) in 85 mL of anhydrous DMF under argon atmosphere was added *tert*-butyldimethylsilyl chloride (6.5 g, 43.1 mmol, 1.6 eq) and imidazole (3.5 g, 51.4 mmol, 1.9 eq). The reaction mixture was stirred at room temperature for 12 h. The solution was concentrated under reduced pressure and the residue was diluted in 70 mL of EtOAc which was then washed twice with 100 mL of aqueous 10 % LiCl. The organic layer was then dried over $MgSO_4$ and evaporated to give **6** (10.3 g, 85%) as a yellowish oil. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.33 (m, 5H), 6.97 (d, J = 8.3 Hz, 2H), 6.76 (d, J = 8.3 Hz, 2H), 5.29 (d, J = 8.0 Hz, 1H; NH), 5.13 (d, J_{AB} = 12.4, 1H), 5.10 (d, J_{AB} = 12.4, 1H), 4.63 (m, 1H), 3.72 (s, 3H), 3.08 (dd, J = 5.8 Hz, J_{AB} = 14.4 Hz, 1H), 3.03 (dd, J = 5.8 Hz, J_{AB} = 14.4 Hz, 1H), 1.00 (s, 9H), 0.21 (s, 6H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 172.1, 155.7, 154.8, 136.3, 130.3, 128.5, 128.3, 128.2, 128.1, 120.2, 66.9, 55.0, 52.3, 37.5, 25.7, 18.2, -4.4.

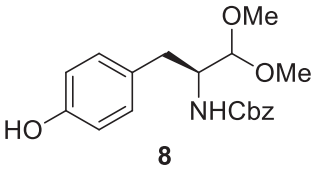
(S)-(1-(4-*tert*-Butyldimethylsilyloxy)benzyl-2-oxo-ethyl)-carbamic Acid Benzyl Ester (7)



7

To a cold (-78 °C) solution of **6** (10.3 g, 23.2 mmol) in dry CH_2Cl_2 (100 mL) was added dropwise a 1.0 M solution of DIBAL-H in CH_2Cl_2 (50 mL, 50 mmol) over 40 min. The reaction mixture, monitored by TLC, was quenched at -78°C with MeOH (30 mL) after 2 hrs and then warmed to room temperature. The mixture was poured into 300 mL of an ice cold 1.2 M HCl aqueous solution, extracted twice with 300 mL of CH_2Cl_2 and the combined organic phases were then washed with 300 mL brine, dried over $MgSO_4$ and concentrated under vacuum. Because of the presence of a chiral center in alpha of the aldehyde, the next step was performed immediately with the crude product.

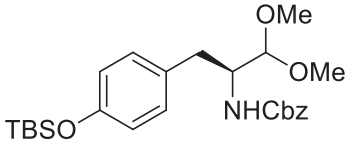
(S)-(1-(4-Hydroxy)benzyl-2,2-dimethoxy-ethyl)-carbamic Acid Benzyl Ester (8)



8

The crude **7** (23.2 mmol) was dissolved in 180 mL of anhydrous MeOH, and *p*- toluenesulfonic acid monohydrate (3.0 g, 15.8 mmol) was added. The reaction was stirred overnight, and then most of the solvent was removed under vacuum. 100 mL of a solution of aqueous saturated $NaHCO_3$ was added, and the mixture was extracted twice with 100 mL of CH_2Cl_2. The combined organic layers were dried over $MgSO_4$, and the solvent was evaporated under vacuum. Purification by flash chromatography (gradient 1:9 to 3:7, EtOAc: Cyclohexane) gave **8** as a white solide (5.67 g , 61 %). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.30 (m, 5H), 7.04 (d, *J* = 8.0 Hz, 2H), 6.73 (d, *J* = 8.0 Hz, 2H), 5.07 (m, 2H), 4.20 (d, *J* = 2.8 Hz, 1H), 4.09 (m, 1H), 3.45 (s, 3H), 3.40 (s, 3H), 2.87 (dd, *J* = 5.8 Hz, J_{AB} = 14.1 Hz, 1H), 2.71 (dd, *J* = 5.8 Hz, J_{AB} = 14.1 Hz, 1H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 156.5, 154.8, 136.4, 130.3, 128.5, 128.1, 127.9, 115.4, 104.9, 68.9, 55.8, 55.7, 52.4, 35.3.

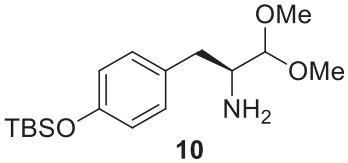
(S)-(1-(*tert*-Butyldimethylsilyloxy)benzyl)-2,2-dimethoxy-ethyl)-carbamic Acid Benzyl Ester (9)



9

To a solution of **8** (1.67 g, 4.8 mmol) in 12 mL of anhydrous DMF was added under argon atmosphere *tert*-butyldimethylsilyl chloride (1.0 g,7.2 mmol, 1.5 eq) and imidazole (0.5 g, 8.2 mmol, 1.7 eq) at room temperature for 12 h. The solution was concentrated under reduced pressure and the residue was diluted in 20 mL of EtOAc which was then washed twice with 20 mL of aqueous 10 % LiCl. The organic layer was then dried over $MgSO_4$ and evaporated to give **9** (1.89 g, 86%) as a yellowish oil. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.32 (m, 5H), 7.07 (d, *J* = 7.8 Hz, 2H), 6.77 (d, *J* = 7.8 Hz, 2H), 5.09 (d, J_{AB} = 12.3, 1H), 5.04 (d, J_{AB} = 12.3, 1H), 4.95 (d, *J* = 9.0 Hz, 1H; NH), 4.17 (d, *J* = 2.5 Hz, 1H), 4.07 (m, 1H), 3.44 (s, 3H), 3.39 (s, 3H), 2.87 (dd, *J* = 6.4 Hz, J_{AB} = 14.2 Hz, 1H), 2.75 (dd, *J* = 6.4 Hz, J_{AB} = 14.2 Hz, 1H), 1.00 (s, 9H), 0.21 (s, 6H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 156.1, 154.2, 136.6, 130.2, 128.5, 128.0, 127.9, 120.0, 104.7, 66.6, 55.7, 55.6, 53.5, 35.3, 25.7, 18.2, -4.4.

(S)-1-(*tert*-Butyldimethylsilyloxy)benzyl-2,2-dimethoxyethan-1-amine (10)



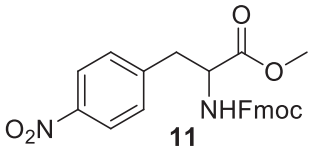
10

To a suspension of Pd on C (0.7 g, 5% wet Degussa type) in MeOH (45 mL) under hydrogen atmosphere, 1.88 g of **9** (4.1 mmol) were added and the reaction was stirred at room temperature. After 3 h, the reaction mixture was filtered on a Celite pad and the solvent was evaporated under reduced pressure to yield **10** (1.36 g, quant) as a yellow oil. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.09 (d, *J* =

8.2 Hz, 2H), 6.79 (d, *J* = 8.2 Hz, 2H), 4.10 (d, *J* = 5.8 Hz, 1H), 3.46 (s, 3H), 3.45 (s, 3H), 3.10 (m, 1H), 2.91 (dd, *J* = 4.1 Hz, J_{AB} = 13.6 Hz, 1H), 2.47 (dd, *J* = 9.3 Hz, J_{AB} = 13.6 Hz, 1H), 1.00 (s, 9H), 0.21 (s, 6H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 154.1, 131.4, 130.2, 120.1, 107.8, 55.1, 55.0, 54.2, 37.8, 25.7, 18.2, -4.4.

1.4. Synthesis of amino-acetal 14

N-Fmoc-L-*p*-Nitrophenylalanine Methyl Ester (11)

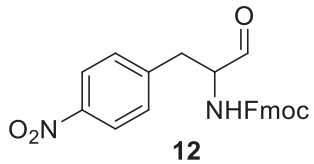


11

Thionyl chloride (1.03 ml, 14.2 mmol) was added dropwise to a suspension of L-*p*- Nitrophenylalanine (0.995 g, 4.73 mmol) in 10 ml of MeOH at 0°C. After the addition, the solution formed was allowed to warm to room temperature. The reaction mixture was heated under reflux overnight and then concentrated. The crude material was recrystallized from methanol/ethyl acetate to give 1.012 g (95%) of L-p-Nitrophenylalanine Methyl Ester Hydrochloride.

The obtained ester (1.178 g, 5.23 mmol) was dissolved in a solution of K_2CO_3 (1.59 g, 11.5 mmol) in dioxane/water (1:1, 20 mL). The mixture was cooled to 0°C and Fmoc-chloride (1.42 g, 5.5 mmol) was added in one batch. The reaction mixture was stirred at 0 °C for 30 min, and the overnight at room temperature. The reaction mixture was subsequently concentrated, the residue was diluted in water (10 mL) and extracted with CH_2Cl_2 (3 x 20 mL). The combined organic extracts were washed with brine, dried over Na_2SO_4, and the solvent was evaporated under vacuum. Purification by flash chromatography (gradient 1:5 to 1:4, EtOAc: Cyclohexane) gave **11** as a white solide (1.807 g , 78 %).^{1}H NMR (500 MHz, $CDCl_3$) δ 8.18 (d, *J* = 8.1 Hz, 2H), 7.83 (d, *J* = 7.4 Hz, 2H), 7.61(d, *J* = 7.4 Hz, 2H), 7.47 (t, *J* = 7.4 Hz, 2H), 7.37 (t, *J* = 7.4 Hz, 2H), 7.26 (d, *J* = 8.1 Hz, 2H), 5.30 (d, *J* = 7.4 Hz, 2H), 4.74 (m, 1H), 4.57 (dd, *J* = 6.2 Hz, J_{AB} = 10.5 Hz, 1H), 4.47 (dd, *J* = 6.2 Hz, J_{AB} = 10.5 Hz, 1H), 4.25 (t, *J* = 6.2 Hz, 1H), 3.80 (s, 3H), 3.31 (dd, *J* = 5.5 Hz, J_{AB} = 13.6 Hz, 1H), 3.20 (dd, *J* = 5.5 Hz, J_{AB} = 13.6 Hz, 1H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 171.2, 155.4, 147.2, 143.7, 143.6, 141.5, 130.3, 127.9, 127.1, 124.9, 123.8, 120.1, 66.8, 54.5, 52.7, 47.2, 38.1.

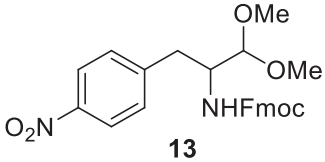
(S)-(1-(*p*-Nitro)benzyl-2-oxo-ethyl)-carbamic Acid 9H-fluoren-9-ylmethyl Ester (12)



12

To a cold (-78 °C) solution of **11** (1.807 g, 4.047 mmol) in dry CH_2Cl_2 (30 mL) was added dropwise a 1.0 M solution of DIBAL-H in CH_2Cl_2 (9.5 mL, 9.5 mmol) over 40 min. The reaction mixture, monitored by TLC, was quenched at -78°C with MeOH (8 mL) after 4 hrs and then warmed to room temperature. The mixture was poured into 30 mL of an ice cold 2 M HCl aqueous solution, extracted twice with 30 mL of CH_2Cl_2 and the combined organic phases were then washed with 50 mL brine, dried over Na_2SO_4 and concentrated under vacuum. Because of the presence of a chiral center in alpha of the aldehyde, the next step was performed immediately with the crude product.

(S)-(1-(*p*-Nitro)benzyl-2,2-dimethoxy-ethyl)-carbamic Acid 9H-fluoren-9-ylmethyl Ester (13)

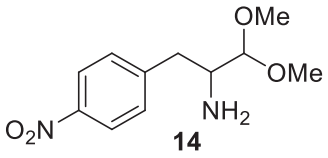


13

The crude **12** (4.047 mmol) was dissolved in 30 mL of anhydrous MeOH, and *p*- toluenesulfonic acid monohydrate (0.6

g, 3.2 mmol) was added. The reaction was stirred overnight, and then most of the solvent was removed under vacuum. 25 mL of a solution of aqueous saturated $NaHCO_3$ was added, and the mixture was extracted twice with 25 mL of CH_2Cl_2. The combined organic layers were dried over Na_2SO_4, and the solvent was evaporated under vacuum. Purification by flash chromatography (1:4, EtOAc: Cyclohexane) gave **13** as a pale yellow solide (1.162 g , 53 %, two steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.17 (d, *J* = 8.1 Hz, 2H), 7.81 (d, *J* = 7.4 Hz, 2H), 7.57(m, 2H), 7.45 (t, *J* = 7.4 Hz, 2H), 7.39 (m, 2H), 7.35 (m, 2H), 4.94 (d, *J* = 7.4 Hz, 2H), 4.42 (m, 2H), 4.20 (m, 2H), 3.49 (s, 3H), 3.47 (s, 3H), 3.12 (dd, *J* = 4.3 Hz, J_{AB} = 13.2 Hz, 1H), 2.89 (dd, *J* = 8.7 Hz, J_{AB} = 13.2 Hz, 1H).^{13}C NMR (127 MHz, $CDCl_3$) δ 156.0, 146.7, 146.1, 143.8, 141.4, 130.2, 127.8, 127.1, 124.9, 123.6, 120.1, 105.1, 66.6, 56.3, 55.9, 53.4, 47.3, 35.5.

(S)-1-(p-Nitro)benzyl-2,2- dimethoxyethan-1-amine (14)

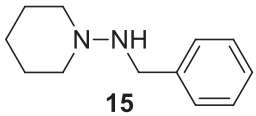


14

The N-Fmoc protected a-aminoacetal **13** (0.156 g, 0.34 mmol) was treated with 20% piperidine in DMF (1.0 mL, 6 eq) and stirred for 20 min at room temperature. After concentration in vacuo, the residue was purified by column chromatography (7:3, EtOAc: hexane; 1:9, MeOH/CH_2Cl_2) to yield **14** as pale yellow oil (64 mg, 79 %).^{1}H NMR (500 MHz, $CDCl_3$) δ 8.10 (d, *J* = 8.6 Hz, 2H), 7.35 (d, *J* = 8.6 Hz, 2H), 4.05 (d, *J* = 5.5 Hz, 1H), 3.40 (s, 3H), 3.39 (s, 3H), 3.13-2.95 (m, 2H), 2.59 (dd, *J* = 9.0 Hz, J_{AB} = 13.4 Hz, 1H), 1.82 (br s, 2H, NH_2).

1.5. Synthesis of hydrazines 15, 16 and 17

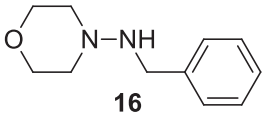
Benzyl-piperidin-1-yl-amine (15)



15

Benzaldehyde (2.44 g, 23.02 mmol) and 1-Aminopiperidine (2.00 g, 20.00 mmol) were stirred at room temperature in anhydrous MeOH (150 mL) overnight. Acetic acid (30 mL, 525 mmol) and $NaCNBH_3$ (6.30 g, 100 mmol) were added, and stirring was continued for 2 h. Most of the solvent was removed under vacuum. Then a pH value around 8 was adjusted by adding $NaHCO_3$ (saturated aqueous solution, 40 mL) and NaOH (aqueous solution, 10 mL), and the mixture was extracted with CH_2Cl_2 (3 × 40 mL). The combined organic extracts were dried over Na_2SO_4, and the solvent was evaporated under vacuum. After purification by flash chromatography (50% EtOAc/cyclohexane), **15** was obtained as a colorless oil (3.44 g, 91%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.36-7.23 (m, 5H), 3.97 (s, 2H), 2.67 (br s, 4H), 1.66-1.62 (m, 4H), 1.57-1.52 (m, 2H). ^{13}C NMR (50 MHz, $CDCl_3$) δ 138.9, 128.3, 128.0, 126.7, 57.3, 52.6, 25.8, 23.7.

Benzyl-morpholin-4-yl-amine (16)



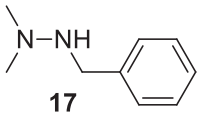
16

Benzaldehyde (1.74 g, 16.5 mmol) and 4-Aminomorpholine (1.59 mL, 16.5 mmol) were stirred at room temperature in anhydrous toluene (10 mL) overnight. The solvent was evaporated and a white solid was obtained.

The solid was dissolved in anhydrous MeOH, acetic acid (21 mL, 367 mmol) and $NaCNBH_3$ (5.0 g, 80 mmol) were added, and stirring was continued for 2 h. Most of the solvent was removed under vacuum. Then a pH value around 8 was adjusted by adding $NaHCO_3$ (saturated aqueous solution, 30 mL) and NaOH (aqueous solution, 8 mL), and the mixture was extracted with CH_2Cl_2 (3 × 30 mL). The combined organic extracts were dried

over Na_2SO_4, and the solvent was evaporated under vacuum. After purification by flash chromatography (50% EtOAc/cyclohexane), **16** was obtained as a yellow oil (2.67 g, 80%). ^{1}H NMR (200 MHz, $CDCl_3$) δ 7.35-7.22 (m, 5H), 3.97 (s, 2H), 3.74 (m, 4H), 2.72 (m, 4H), 1.97 (br s, 1H).

2-Benzyl-1,1-dimethylhydrazine (17)



17

Benzaldehyde (1.58 g, 15 mmol) and *N,N*-dimethylhydrazine (2.69 g, 45 mmol) were stirred at room temperature in anhydrous ethanol (30 mL) for 1 h. The reaction was quenched by pouring the mixture into water (30 mL) and was then extracted with CH_2Cl_2 (2 x 30 mL). After drying the combined organic phase over $MgSO_4$ and removing the solvent under reduced pressure, the target *N,N*-dimethylhydrazone was obtained as a colorless oil (2.21 g, 33 %).

The oil was dissolved in anhydrous MeOH, acetic acid (20 mL, 345 mmol) and $NaCNBH_3$ (4.71 g, 75 mmol) were added, and stirring was continued for 2 h. Most of the solvent was removed under vacuum. Then a pH value around 8 was adjusted by adding $NaHCO_3$ (saturated aqueous solution, 30 mL) and NaOH (aqueous solution, 8 mL), and the mixture was extracted with CH_2Cl_2 (3 × 30 mL). The combined organic extracts were dried over Na_2SO_4, and the solvent was evaporated under vacuum. After purification by flash chromatography (50% EtOAc/cyclohexane), **17** was obtained as a colorless oil (1.97 g, 88 %). ^{1}H NMR (200 MHz, $CDCl_3$) δ 7.33-7.19 (m, 5H), 3.89 (s, 2H), 2.45 (s, 6H), 2.29 (br s, 1H).

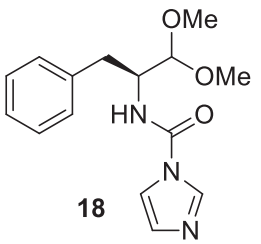
1.6. Procedures for the synthesis of ureic acetals 21 to 29 and their imidazole carboxamide intermediates 18 to 20

General coupling procedure

A solution of amine in anhydrous THF (1 mL per mmol) was added dropwise to a stirred suspension of CDI (1.1 eq) in anhydrous THF (1 mL per mmol) at room temperature. After 30 min the solvent was evaporated, and the residue redissolved in CH_2Cl_2 (5 mL per mmol) and washed twice with water (2 mL per mmol). The organic phase was dried over $MgSO_4$, and the solvent evaporated under vacuum to give pure imidazole carboxamide.

The imidazole carboxamide compound was dissolved in anhydrous acetonitrile (1.5 mL per mmol) and iodomethane (4 eq) was added. The reaction was stirred overnight at room temperature, then the solvent was evaporated, and the yellow orange oil was dried under vacuum. The residue was redissolved in dry CH_2Cl_2 (5 mL per mmol), and the hydrazine (1 eq) and triethylamine (1 eq) were added. After stirring for 24h at room temperature, the reaction was quenched by adding saturated aqueous $NaHCO_3$ (5 mL per mmol), and the mixture was extracted CH_2Cl_2 (twice with 5 mL per mmol). The combined organic extracts were dried over $MgSO_4$, and the solvent was removed under vacuum. After purification by flash chromatography (1:4 to 1:1, EtOAc: Cyclohexane), the urea compound was obtained.

(S)-Imidazole-1-carboxylic Acid (1-Benzyl-2,2-dimethoxyethyl)-amide (18)

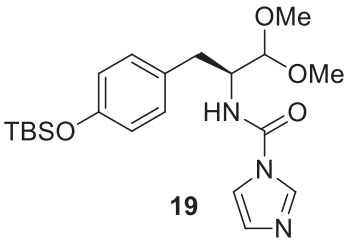


18

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.04 (s, 1H), 7.33-7.22 (m, 6H), 7.07 (s, 1H), 5.90 (d, *J* = 8.5 Hz, 1H), 4.42-4.37 (m, 1H), 4.26 (d, *J* = 2.7 Hz, 1H), 3.49 (s, 3H), 3.39 (s, 3H), 3.02-2.94 (m, 2H); ^{13}C NMR (50 MHz,

CDCl3) δ 148.6, 137.0, 135.9, 130.5, 129.2, 128.7, 126.8, 115.8, 104.1, 56.0, 55.7, 53.5, 36.0.

(S)-Imidazole-1-carboxylic Acid (1-(*p*-t*ert*-Butyldimethylsilyloxy)benzyl-2,2-dimethoxy-ethyl)-amide (19)

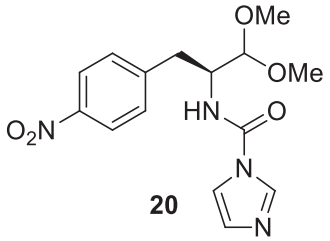


19

Yellow oil, yield : 86%. ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.06 (s, 1H), 7.29 (m, 1H), 7.09 (m, 3H), 6.79 (d, *J* = 8.2 Hz, 2H), 4.35 (m, 1H), 4.26 (d, *J* = 2.7 Hz, 1H), 3.49 (s, 3H), 3.40 (s, 3H), 2.93 (dd, *J* = 7.0 Hz, J_{AB} = 14.4 Hz, 1H), 2.90 (dd, *J* = 7.6 Hz, J_{AB} = 14.4 Hz, 1H), 0.98 (s, 9H), 0.19 (s, 6H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 154.5, 148.6, 135.9, 130.5, 130.1, 129.6, 120.3, 115.8, 104.1, 56.0, 55.7, 53.7, 35.3, 25.7, 18.2, -4.4.

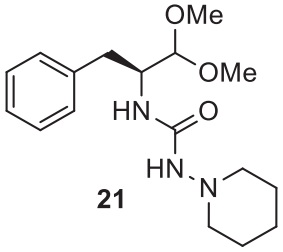
(S)-Imidazole-1-carboxylic Acid (1-(*p*-Nitro)benzyl-2,2-dimethoxy-ethyl)-amide (20)

This product was not isolated. Next step was conducted on the crude product.



20

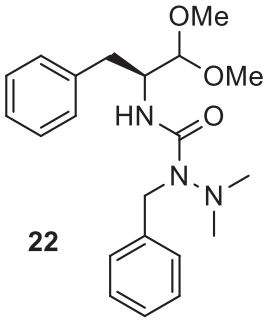
(S)-3-(1-benzyl-2,2-dimethoxy-ethyl)-1-piperidin-1-yl-urea (21)



21

From **18** + commercial Piperidin-1-ylamine. Colorless oil, yield : 49%.^{1}H NMR (500 MHz, $CDCl_3$) δ 7.2-7.3 (m, 5H), 6.13 (d, *J* = 8.5 Hz, 1H; NH), 4.98 (s, 1H; NH), 4.27 (d, *J* = 4.0 Hz, 1H), 4.22 (m, 1H), 3.45 (s, 6H), 3.02 (dd + bs, *J* = 4.9 Hz, J_{AB} = 13.9 Hz, 2H), 2.77 (dd, *J* = 9.0 Hz, J_{AB} = 13.9 Hz, 1H), 2.69 (b, 1H), 2.20 (b, 1H), 2.08 (b, 1H), 1.62 (b, 5H), 1.09 (b, 1H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 157.8, 138.4, 129.4, 128.3, 126.2, 105.5, 57.4, 56.2, 55.1, 51.7, 35.2, 25.8, 23.1.

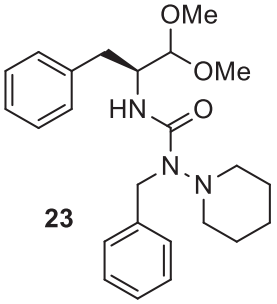
(S)-1-Benzyl-3-(1-benzyl-2,2-dimethoxyethyl)-1-dimethylamino-urea (22)



22

From **18** + **17**. Colorless oil, yield 52 %.^{1}H NMR (500 MHz, $CDCl_3$) δ 7.30-7.15 (m, 5H), 6.59 (d, *J* = 9.0 Hz, 1H; NH), 4.60 (d, *J* = 16.5 Hz, 1H), 4.43 (d, *J* = 16.5, 2H), 4.34 (m, 2H), 3.50 (s, 6H), 3.09 (dd, *J* = 4.3 Hz, J_{AB} = 14.0 Hz, 1H), 2.80 (dd, *J* = 8.3 Hz, J_{AB} = 14.0 Hz, 1H), 2.43 (s, 3H), 2.25 (s, 3H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 157.9, 140.5, 138.7, 129.5, 128.3, 127.2, 126.6, 126.1, 106.0, 56.3, 55.3, 52.4, 43.9, 43.6, 41.1, 35.5.

(S)-1-Benzyl-3-(1-benzyl-2,2-dimethoxy -ethyl)-1-piperidin-1-yl-urea (23)

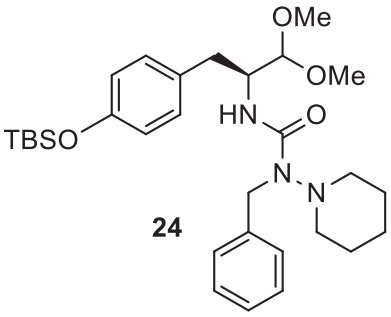


23

From **18** + **15**. ^{1}H NMR (500 MHz, CDCl3) δ 7.29-7.13 (m, 10H), 6.61 (d, *J* = 9.0 Hz, 1H), 4.60 (d, *J* = 16.3 Hz, 1H), 4.45 (d, *J* = 16.3 Hz, 1H), 4.31-4.27 (m, 2H), 3.473 (s, 3H), 3.468 (s, 3H), 3.06 (dd, *J* = 13.9, 4.5 Hz, 1H), 2.78 (dd, *J* = 13.9, 8.9 Hz, 1H), 2.71-2.69 (m, 1H), 2.52-2.48 (m, 1H), 2.42-2.40 (m, 1H), 2.36-2.32 (m, 1H), 1.65-1.42

(m, 5H), 1.03-0.95 (m, 1H). ^{13}C NMR (50 MHz, CDCl3) δ 157.8, 140.5, 138.6, 129.4, 128.2, 128.1, 127.2, 126.4, 126.0, 105.9, 56.3, 55.0, 53.4, 53.1, 52.1, 42.0, 35.3, 26.4, 23.2.

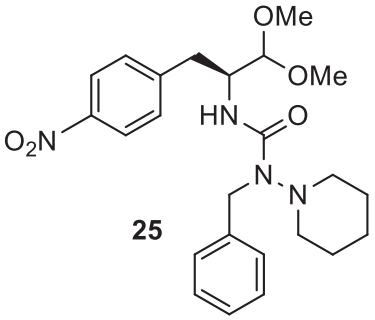
(S)-1-Benzyl-3-(1-(*p-tert*-Butyldimethylsilyloxy)benzyl-2,2-dimethoxy-ethyl)-1-piperidin-1-yl-urea (24)



24

From **19** + **15**. Colorless oil, yield : 82%. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.27 (m, 2H), 7.19 (m, 3H), 7.12 (d, *J* = 8.5 Hz, 2H), 6.78 (d, *J* = 8.5 Hz, 2H), 6.64 (d, *J* = 9.0 Hz, 1H; NH), 4.62 (d, J_{AB} = 16.4 Hz, 1H), 4.50 (d, J_{AB} = 16.4 Hz, 1H), 4.31 (d, *J* = 3.7 Hz, 1H), 4.28 (m, 1H), 3.48 (s, 3H), 3.47 (s, 3H), 3.00 (dd, *J* = 5.0 Hz, J_{AB} = 13.8 Hz, 1H), 2.74 (m, 2H), 2.50 (m, 2H), 2.39 (m, 1H), 1.63 (m, 3H), 1.51 (m, 2H), 1.01 (m, 10H), 0.19 (s, 6H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 157.9, 154.0, 140.6, 131.2, 130.4, 128.3, 127.3, 126.5, 119.8, 105.8, 56.2, 54.9, 53.5, 53.2, 52.2, 42.1, 34.7, 26.5, 25.7, 18.2, -4.4.

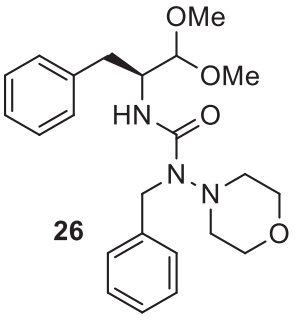
(S)-1-Benzyl-3-(1-(4-nitro)benzyl-2,2-dimethoxy-ethyl)-1-piperidin-1-yl-urea (25)



25

From **20** + **15**. Colorless oil, yield : 46%. . ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.18 (d, *J* = 8.3 Hz, 2H), 7.45 (d, *J* = 8.3 Hz, 2H), 7.28-7.16 (m, 5H), 6.72 (d, *J* = 9.2 Hz, 1H), 4.64 (d, J_{AB} = 16.6 Hz, 1H), 4.46 (d, J_{AB} = 16.6 Hz, 1H), 4.38 (m, 1H), 4.35 (m, 1H), 3.53 (s, 3H), 3.52 (s, 3H), 3.21 (dd, *J* = 4.6 Hz, J_{AB} = 12.9 Hz, 1H), 2.92 (dd, *J* = 8.3 Hz, J_{AB} = 12.9 Hz, 1H), 2.74 (m, 2H), 2.59 (m, 2H), 2.47 (m, 1H), 1.67 (m, 3H), 1.55 (m, 2H), 1.05 (m, 1H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 157.7, 147.3, 146.6, 140.3, 130.4, 128.3, 127.3, 126.8, 123.4, 106.2, 56.8, 55.6, 53.8, 53.2, 52.1, 42.3, 35.5, 26.5.

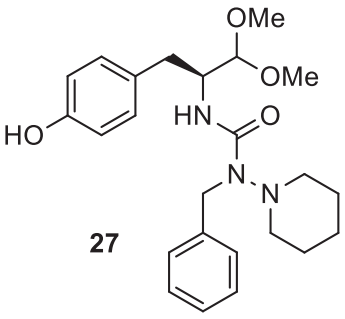
(S)-1-benzyl-1-morpholino-3-(1-benzyl-2,2-dimethoxy-ethyl)-urea (26)



26

From **19** + **15**. Colorless oil, yield : 82%. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.27 (m, 2H), 7.19 (m, 3H), 7.12 (d, *J* = 8.5 Hz, 2H), 6.78 (d, *J* = 8.5 Hz, 2H), 6.64 (d, *J* = 9.0 Hz, 1H; NH), 4.62 (d, J_{AB} = 16.4 Hz, 1H), 4.50 (d, J_{AB} = 16.4 Hz, 1H), 4.31 (d, *J* = 3.7 Hz, 1H), 4.28 (m, 1H), 3.48 (s, 3H), 3.47 (s, 3H), 3.00 (dd, *J* = 5.0 Hz, J_{AB} = 13.8 Hz, 1H), 2.74 (m, 2H), 2.50 (m, 2H), 2.39 (m, 1H), 1.63 (m, 3H), 1.51 (m, 2H), 1.01 (m, 10H), 0.19 (s, 6H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 157.9, 154.0, 140.6, 131.2, 130.4, 128.3, 127.3, 126.5, 119.8, 105.8, 56.2, 54.9, 53.5, 53.2, 52.2, 42.1, 34.7, 26.5, 25.7, 18.2, -4.4.

(S)-1-Benzyl-3-(1-(4-hydroxy)benzyl-2,2-dimethoxy-ethyl)-1-piperidin-1-yl-urea (27)

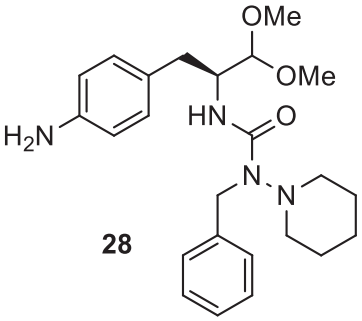


27

Compound **26** (415 mg, 0.77 mmol) was dissolved in 25 mL THF and 1.5 mL of a 1M solution of tetra-n-butylammonium fluoride in THF was added dropwise. The reaction was stirred for 10 min and the solvent was evaporated. The residue was redissolved in CH_2Cl_2 (20 mL) and washed with water (20 mL). The aqueous phase was extracted with CH_2Cl_2 (10 mL), and the combined organic phase was dried over $MgSO_4$ and

concentrated under vacuum. After purification by flash chromatography (gradient 1:4 to 1:1, EtOAc: Cyclohexane), **27** (332 mg, quant) was obtained as a colorless gel-oil. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.25 (m, 2H), 7.17 (m, 3H), 7.02 (d, *J* = 8.2 Hz, 2H), 6.77 (d, *J* = 9.1 Hz, 1H; NH), 6.69 (d, *J* = 8.5 Hz, 2H), 4.61 (d, J_{AB} = 16.0 Hz, 1H), 4.54 (d, J_{AB} = 16.0 Hz, 1H), 4.29 (d, *J* = 3.3 Hz, 1H), 4.26 (m, 1H), 3.48 (s, 3H), 3.46 (s, 3H), 2.95 (dd, *J* = 5.0 Hz, J_{AB} = 14.0 Hz, 1H), 2.70 (m, 2H), 2.51 (m, 2H), 2.39 (m, 1H), 1.63 (m, 3H), 1.53 (m, 2H), 1.01 (m, 1H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 158.3, 155.3, 140.2, 130.2, 129.3, 128.3, 127.2, 126.6, 115.4, 105.9, 56.2, 55.3, 53.4, 53.3, 52.8, 42.3, 35.0, 26.4.

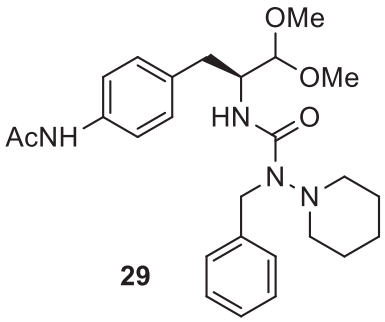
(S)-1-Benzyl-3-(1-(4-amino)benzyl-2,2-dimethoxy-ethyl)-1-piperidin-1-yl-urea (28)



28

To a suspension of Pd on C (0.1 g, 10% wet Degussa type) in MeOH (10 mL) under hydrogen atmosphere, 231 mg of **25** (0.52 mmol) were added and the reaction was stirred at room temperature. After 2 h, the reaction mixture was filtered on a Celite pad and the solvent was evaporated under reduced pressure to yield **10** (216 mg, quant) as a yellow oil. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.32-7.21 (m, 5H), 7.07 (d, *J* = 8.3 Hz, 2H), 6.66 (d, *J* = 8.3 Hz, 2H), 4.64 (d, J_{AB} = 16.0 Hz, 1H), 4.54 (d, J_{AB} = 16.0 Hz, 1H), 4.32 (d, *J* = 4.2 Hz, 1H), 4.26 (m, 1H), 3.50 (s, 6H), 2.96 (dd, *J* = 4.8 Hz, J_{AB} = 13.9 Hz, 1H), 2.75 (m, 2H), 2.55 (m, 2H), 2.43 (m, 1H), 1.67 (m, 3H), 1.55 (m, 2H), 1.05 (m, 1H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 158.1, 144.6, 140.7, 130.3, 128.5, 128.3, 127.3, 126.5, 115.2, 105.8, 56.2, 55.1, 53.5, 53.3, 52.3, 42.2, 34.8, 26.5, 23.3.

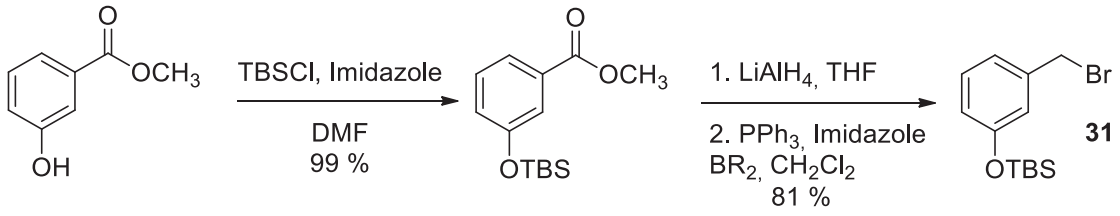
(S)-1-Benzyl-3-(1-(4-acetylamido)benzyl-2,2-dimethoxy-ethyl)-1-piperidin-1-yl-urea (29)



29

To a solution of **28** (102 mg, 0.25 mmol) in CH_2Cl_2 (6 mL) were added acetic anhydride (0.15 mL, 1.6 mmol) and Et_3N (50 µL, 0.36 mmol). The reaction was stirred at room temperature for 30 min and was then quenched by adding 1N HCl aqueous solution (10 mL). The mixture was extracted extracted with CH_2Cl_2 (3 times 10 mL), and the combined organic extracts were washed with a saturated aqueous solution of $NaHCO_3$ (20 mL) , dried over Na_2SO_4 and the solvent was evaporated under vacuum. After purification by flash chromatography (gradient 1:1 to 3:1, EtOAc: Cyclohexane), **29** (105 mg, 94%) was obtained as a colorless oil. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.38 (d, *J* = 7.8 Hz, 2H), 7.24-7.15 (m, 5H), 6.63 (d, *J* = 9.0 Hz, 2H), 4.54 (d, J_{AB} = 16.0 Hz, 1H), 4. 74 (d, J_{AB} = 16.0 Hz, 1H), 4.26 (d, *J* = 4.2 Hz, 1H), 4.22 (m, 1H), 3.44 (s, 6H), 2.97 (dd, *J* = 5.3 Hz, J_{AB} = 13.9 Hz, 1H), 2.71 (m, 2H), 2.47 (m, 2H), 2.37 (m, 1H), 2.13 (s, 3H), 1.59 (m, 3H), 1.48 (m, 2H), 0.98 (m, 1H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 168.2, 158.0, 140.6, 136.3, 134.4, 130.0, 128.3, 127.3, 126.6, 119.7, 105.9, 56.3, 55.2, 53.5, 53.4, 52.4, 42.2, 35.0, 26.5, 24.7, 23.3.

1.7. Preparation of substituted benzyl bromide 31



Ref : S. Kwon, A. G. Myers, *J. Am. Chem. Soc.* **2005**, *127*, 16796–16797.

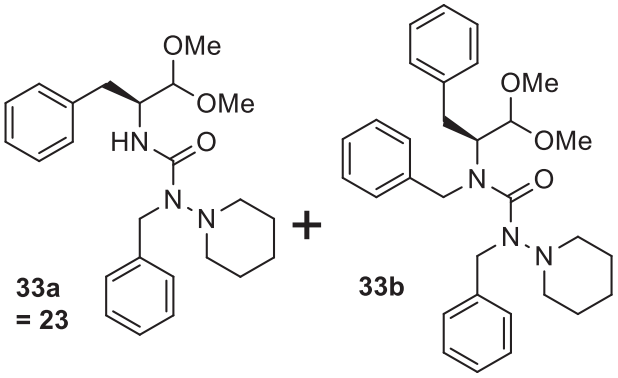
1.8. Synthesis of 33a to 35b

General procedure for the double alkylation of 21 to obtain 33 to 35 (a and b)

To a solution of NaH (400 mg, 16.6 mmol) in anhydrous DMF (6 mL) was added dropwise a solution of compound **1** (400 mg, 1.25 mmol) in anhydrous DMF (4 mL). Benzyl bromide (1.0 g, 5.85 mmol) was then added dropwise. The reaction was stirred at room temperature for 48 h and was then quenched by adding iso-propanol (1.5 mL) at 0 °C and then water (1.5 mL) still at 0°C. Water (20 mL) was added and the mixture was extracted with EtOAc (2 x 20 mL). The combined organic extracts were washed with LiCl (10% aqueous solution, 2 x 20 mL), then dried over Na_2SO_4 and the solvent was evaporated under vacuum.

(S)-1-benzyl-3-(1-benzyl-2,2-dimethoxy-ethyl)-1-piperidin-1-yl-urea (33a = 23) and (S)-1,3-dibenzyl-3-(1-benzyl-2,2-dimethoxy -ethyl)-1-piperidin-1-yl-urea (33b)

After purification by flash chromatography (50 % EtOAc/cyclohexane), **2** and **3** were obtained as colorless oils (**33a**: 182 mg, 35 %, **33b**: 77 mg, 12%).



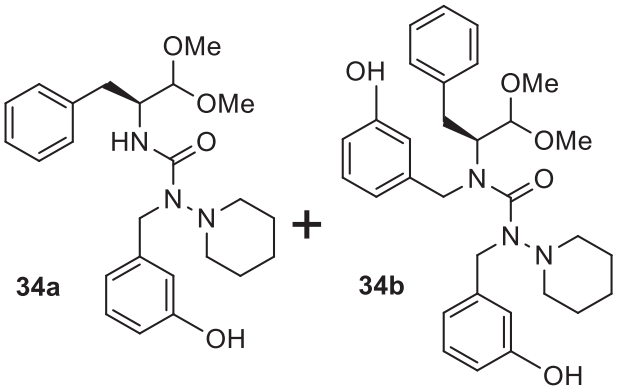
33a = 23 **33b**

Analysis **33a** : cf 23

Analysis **33b** : ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.0-7.35 (m, 15H), 4.81 (m, 1H), 4.55 (m, 2H), 4.38 (m, 3H), 3.34 (s, 3H), 3.26 (s + m, 3H), 3.02 (dd, J = 4.6 Hz, J_{AB} = 14.0 Hz, 1H), 2.55-2.65 (b, 4H), 1.10-1.34 (b, 6H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 162.7, 139.8, 139.7, 130.1, 129.6, 128.3, 128.2, 128.0, 127.8, 127.0, 126.4, 126.0, 125.9, 107.5, 52.8, 43.3, 27.0, 25.5, 23.4. MS (ESI) m/z 502.2 ([M + H]$^+$, 100)

(S)-1-(3-hydroxy-benzyl)-3-(1-benzyl-2,2-dimethoxy-ethyl)-1-piperidin-1-yl-urea (34a) and (S)-1,3-di(3-hydroxy-benzyl)-3-(1-benzyl-2,2-dimethoxy-ethyl)-1-piperidin-1-yl-urea (34b)

After purification by flash chromatography (50 % EtOAc/cyclohexane), **2** and **3** were obtained as colorless oils (**34a**: 244 mg, 46 %, **34b**: 55 mg, 8%).



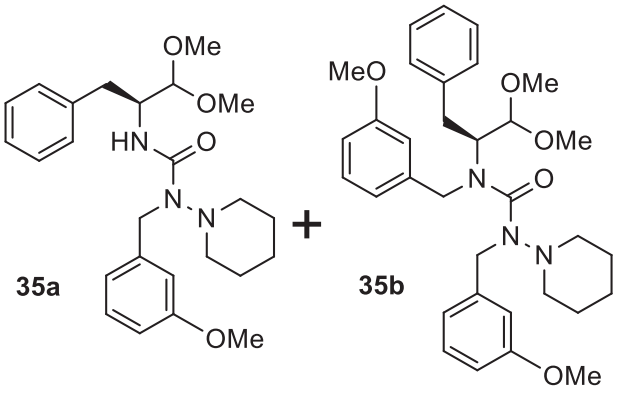
34a **34b**

Analysis **34a** : ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.20-7.30 (m, 5H), 7.04 (t, 1H), 6.77 (d, 1H), 6.75 (s, 1H), 6.67 (d, 1H), 6.64 (d, 1H),

4.54 (d_{AB}, 1H), 4.41 (d_{AB}, 1H), 4.17 (m, 1H), 3.47 (s, 6H), 3.06 (dd_{AB}, 1H), 2.83 (dd_{AB}, 1H), 2.71 (d, 1H), 2.49 (t, 1H), 2.43 (d, 1H), 2.37 (t, 1H), 1.61 (m, 3H), 1.48 (m, 2H), 1.00 (m, 1H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 158.3, 156.8, 141.7, 138.4, 129.5, 129.2, 128.3, 126.2, 118.6, 114.5, 114.0, 105.8, 56.2, 55.3, 53.3, 53.1, 52.3, 42.2, 35.6, 26.4, 23.2. MS (ESI) m/z 428.2 ($[M + H]^+$, 100)

Analysis **34b**: too many impurities on NMR. MS (ESI) m/z 534.2 ($[M + H]^+$, 100)

(S)-1-(3-methoxy-benzyl)-3-(1-benzyl-2,2-dimethoxy-ethyl)-1-piperidin-1-yl-urea (35a) and (S)-1,3-di(3-methoxy-benzyl)-3-(1-benzyl-2,2-dimethoxy-ethyl)-1-piperidin-1-yl-urea (35b)



After purification by flash chromatography (gradient 1:9 to 1:4, EtOAc: cyclohexane), **35a** and **35b** were obtained as colorless oils (**35a**: 146 mg, 48 %, **35b**: 90 mg, 23%).

Analysis **35a**: ^1H NMR (500 MHz, $CDCl_3$) δ 7.30-7.17 (m, 6H), 6.82 (s, 1H), 6.76 (t, J = 8.6 Hz, 1H), 6.65 (d, J_{AB} = 16.5 Hz, 1H), 4.58 (d, J = 8.6 Hz, 1H), 4.49 (d, J_{AB} = 16.5 Hz, 1H), 4.33 (m, 1H), 4.30 (m, 1H), 3.78 (s, 3H), 3.49 (s, 6H), 3.07 (dd, J = 4.6 Hz, J_{AB} = 13.8 Hz, 1H), 2.82 (dd, J = 9.2 Hz, J_{AB} = 13.8 Hz, 1H), 2.73 (m, 1H), 2.52 (m, 1H), 2.46 (m, 1H), 2.41 (m, 1H), 1.64 (m, 3H), 1.50 (m, 2H), 1.02 (m, 1H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 159.6, 157.9, 142.4, 138.7, 129.5, 129.2, 128.3, 126.1, 119.7, 112.9, 112.0, 105.8, 56.3, 55.1, 55.0, 53.3, 53.2, 52.2, 42.1, 35.4, 26.5, 23.3. MS (ESI) m/z 442.2 ($[M + H]^+$, 100)

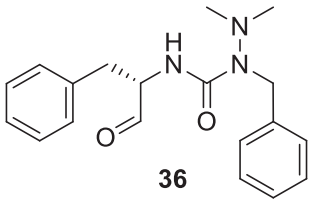
Analysis **35b**: ^1H NMR (500 MHz, $CDCl_3$) δ 7.18-7.23 (m, 4H), 7.06-7.13 (m, 3H), 6.86-6.87 (m, 2H), 6.82 (s, 1H), 6.78 (d, J = 7.2 Hz, 1H), 6.71 (d, J = 7.2 Hz, 1H), 6.59 (d, J = 7.2 Hz, 1H), 4.83 (m, 1H), 3.81 (s, 3H), 3.78 (m, 1H), 3.73 (s, 1H), 3.40 (s, 1H), 3.27 (s + m, 4H), 3.02 (dd, J = 4.6 Hz, J_{AB} = 13.8 Hz, 1H), 2.55-2.65 (b, 4H), 1.10-1.34 (b, 6H). ^{13}C NMR (127 MHz, $CDCl_3$) δ 162.4, 159.5, 159.4, 141.6, 139.6, 129.5, 129.0, 128.8, 128.2, 126.0, 120.4, 119.4, 113.7, 112.2, 111.9, 111.7, 107.2, 61.9, 56.6, 55.15, 55.05, 52.7, 43.3, 34.3, 25.5, 23.4. MS (ESI) m/z 562.2 ($[M + H]^+$, 100).

1.9. Synthesis of 36 to 43

General procedure for the deprotection of all acetal precursors to hydrazino-aldehydes

To a solution of the acetal precursor (63 mg, 0.14 mmol) in anhydrous acetonitrile (20 mL per mmol) were added NaI (3 eq) and TMSCl (2 eq). The reaction was stirred at room temperature for 1.5 h and was then quenched by adding $NaHCO_3$ (saturated aqueous solution, 20 mL per mmol). The mixture was extracted with CH_2Cl_2 (3 times 60 mL per mmol), and the combined organic extracts were washed with $Na_2S_2O_3$ (saturated solution, 100 mL per mmol) and then with water (60 mL per mmol). The organic phase was dried over Na_2SO_4 and the solvent was evaporated under vacuum to give the crude aldehyde.

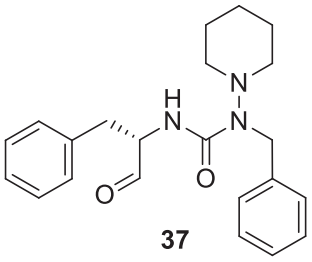
(S)-1-Benzyl-3-(1-benzyl-2-oxo-ethyl)-1-dimethylamino-urea (36)



36

From acetal **22** (41 mg, 0.11 mmol). Purification by flash chromatography (gradient 1:2 to 1:1, EtOAc: Cyclohexane), gave **36** as a colorless oil (14 mg, 39 %). ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.78 (s, 1H; CHO), 7.35-7.25 (m, 10H; Ar), 6.95 (d, *J* = 7.3 Hz, 1H; NH), 4.64 (m, 1H; NHCH), 4.60 (d, *J* = 4.3 Hz, 2H; NCH_2Ph), 3.28 (dd, *J* = 6.8 Hz, J_{AB} = 14.0 Hz, 1H; CH_2Ph), 3.14 (dd, *J* = 7.4 Hz, J_{AB} = 14.0 Hz, 1H; CH_2Ph), 2.46 (s, 3H; $NNCH_3$), 2.36 (s, 3H; $NNCH_3$). ^{13}C NMR (127 MHz, $CDCl_3$) δ 200.5 (CHO), 157.9 (NCON), 140.0 (Ar), 136.7 (Ar), 129.5 (Ar), 128.6 (Ar), 128.5 (Ar), 127.4 (Ar), 126.9 (Ar), 126.9 (Ar), 60.7 (NHCH), 43.9 ($NNCH_3$), 41.6 (NCH_2Ph), 35.7 (CH_2Ph). HRMS (ESI) calcd for $C_{19}H_{24}N_3O_2$ $[M + H]^+$ 326.1863 found 326.1860.

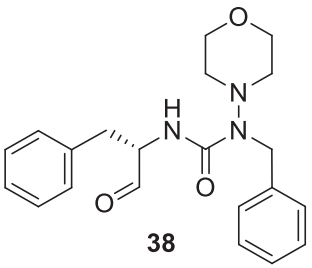
(S)-1-Benzyl-3-(1-benzyl-2-oxo-ethyl)-1-piperidin-1-yl-urea (37)



37

From acetal **23** (220 mg, 0.53 mmol). Purification by flash chromatography (gradient 1:2 to 1:1, EtOAc: Cyclohexane), gave **37** as a colorless oil (164 mg, 85 %). ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.66 (s, 1H; CHO), 7.25-7.12 (m, 10H; Ar), 6.85 (d, *J* = 6.9 Hz, 1H; NH), 4.51 (s, 2H; NCH_2Ph), 4.50-4.46 (m, 1H; NHCH), 3.12 (dd, *J* = 6.0 Hz, J_{AB} = 14.0 Hz, 1H; CH_2Ph), 3.05 (dd, *J* = 7.4 Hz, J_{AB} = 14.0 Hz, 1H; CH_2Ph), 2.64-2.62 (m, 1H; $NNCH_{2piperidine}$), 2.48-2.36 (m, 3H; $NNCH_{2pip}$), 1.54-1.50 (m, 3H; CH_{2pip}), 1.36-1.22 (m, 2H; CH_{2pip}), 0.95-0.87 (m, 1H; CH_{2pip}). ^{13}C NMR (127 MHz, $CDCl_3$) δ 200.7 (CHO), 158.0 (NCON), 140.0 (Ar), 136.4 (Ar), 129.3 (Ar), 128.6 (Ar), 128.3 (Ar), 127.4 (Ar), 126.9 (Ar), 126.7 (Ar), 60.2 (NHCH), 53.5 ($NNCH_{2pip}$), 53.4 ($NNCH_{2pip}$), 42.3 (NCH_2Ph), 35.3 (CH_2Ph), 26.3 (CH_{2pip}), 26.2 (CH_{2pip}), 23.1 (CH_{2pip}). ^{1}H NMR (500 MHz, MeOH-d_4), (1:1 mixture of two diasteromers) δ 7.33-7.10 (m, 18H; Ar), 6.94-6.88 (m, 2H; Ar), 4.61-4.57 (m, 4H; NCH_2Ph, $^{\delta+}N\rightarrow CH=O^{\delta-}$), 4.49-4.45 (m, 2H; NCH_2Ph), 4.14-4.05 (m, 2H; NHCH), 3.12-3.05 (m, 2H; CH_2Ph), 2.81-2.72 (m, 4H; CH_2Ph, $NNCH_{2pip}$), 2.64-2.58 (m, 2H; $NNCH_{2pip}$), 2.41-2.37 (m, 4H; $NNCH_{2pip}$), 1.67-1.60 (m, 10H; CH_{2pip}), 1.11-1.04 (m, 2H; CH_{2pip}). ^{13}C NMR (127 MHz, MeOH-d_4) δ 160.9 (NCON), 160.8 (NCON), 142.4 (Ar), 140.9 (Ar), 131.39 (Ar), 131.37 (Ar), 130.2 (Ar), 130.0 (Ar), 129.03 (Ar), 129.02 (Ar), 128.4 (Ar), 128.1 (Ar), 100.5 ($^{\delta+}N\rightarrow CH=O^{\delta-}$), 100.1($^{\delta+}N\rightarrow CH=O^{\delta-}$), 57.5 (NHCH), 57.4 (NHCH), 55.41 ($NNCH_{2pip}$), 55.38 ($NNCH_{2pip}$), 54.92 ($NNCH_{2pip}$), 54.86 ($NNCH_{2pip}$), 43.9 (NCH_2Ph), 37.44 (CH_2Ph), 37.36 (CH_2Ph), 28.36 (CH_{2pip}), 28.34 (CH_{2pip}), 28.30 (CH_{2pip}), 28.24 (CH_{2pip}), 25.1 (CH_{2pip}). MS (ESI) m/z 366.2 ($[M + H]^+$, 100). HRMS (ESI) calcd for $C_{22}H_{28}N_3O_2$ $[M + H]^+$ 366.2182, found 366.2187. Anal. Calcd for $C_{22}H_{27}N_3O_2$: C, 72.30; H, 7.45; N, 11.50. Found: C, 72.19; H, 7.48; N 11.46.

(S)-1-benzyl-1-morpholino-3-(1-oxo-3-phenylpropan-2-yl)-urea (38)

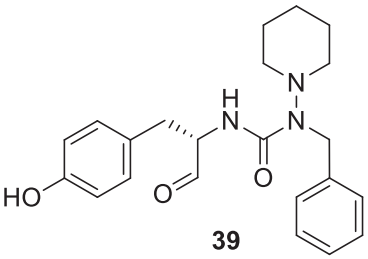


38

From acetal **26** (70 mg, 0.17 mmol). Purification by flash chromatography (gradient 1:4 to 2:1, EtOAc: Cyclohexane), gave **38** as a colorless oil (59 mg, 95 %). ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.78 (s, 1H; CHO), 7.35-7.23 (m, 10H; Ar), 6.88 (d, *J* = 7.0 Hz, 1H; NH), 4.66 (s, 2H; NCH_2Ph), 4.64 (m, 1H; NHC*H*), 3.77 (t, *J* = 12.0 Hz, 2H; $NNCH_{2morpholine}$), 3.45-3.35 (m, 2H; $NNCH_{2morp}$), 3.25 (dd, *J* = 6.2 Hz, J_{AB} = 13.8 Hz, 1H; CH_AH_BPh), 3.19 (dd, *J* = 7.0 Hz, J_{AB} = 14.0 Hz, 1H; CH_AH_BPh), 2.87-2.79 (m, 2H; CH_2O_{morp}), 2.57 (d, *J* = 11.0 Hz, 1H; CH_2O_{morp}), 2.37 (d, *J* = 10.7 Hz, 1H; CH_2O_{morp}). ^{13}C NMR (127 MHz, $CDCl_3$) δ 200.4 (CHO), 157.7

(NCON), 139.7 (PhC_q), 136.4 (PhC_q), 129.4 (2Ph), 128.8 (2Ph), 128.5 (2Ph), 127.5 (2Ph), 127.1 (Ph),127.0 (Ph), 67.1 + 66.9 (CH_2O_{morp}), 60.2 (NHCH), 52.5 + 52.4 ($NNCH_{2morp}$), 42.5 (NCH_2Ph), 35.2 ($CHCH_2Ph$). HRMS (ESI) calcd for $C_{21}H_{26}N_3O_3$ $[M + H]^+$ 368.1969, found 368.1964.

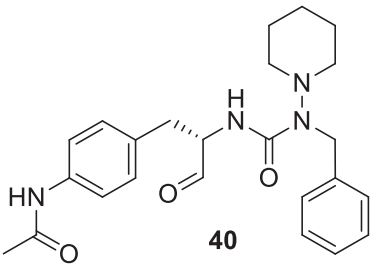
(S)-1-Benzyl-3-(1-(4-hydroxy)benzyl-2-oxo-ethyl)-1-piperidin-1-yl-urea (39)



39

From acetal **27** (32 mg, 0.07 mmol). Purification by flash chromatography (gradient 1:4 to 1:1, EtOAc: Cyclohexane), gave **39** as a colorless oil (17 mg, 59 %). ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.71 (s, 1H; CHO), 7.28 (m, 5H; Ar_{Phenyl}), 7.04 (d, J = 7.8 Hz, 1H; NH), 6.98 (d, J = 8.6 Hz, 2H; Ar_{Phenol}), 6.71 (d, J = 8.6 Hz, 2H; Ar_{Phenol}), 4.61 (m, 3H; NCH_2Ph + NHCH), 3.10 (dd, J = 6.2 Hz, J_{AB} = 14.0 Hz, 1H; C$H_A$$H_B$Phenol), 3.02 (dd, J = 7.0 Hz, J_{AB} = 14.0 Hz, 1H; $CH_A$$H_B$Phenol), 2.75 (m, 1H; $NNCH_{2piperidine}$), 2.60 (m, 1H; $NNCH_{2pip}$), 2.50 (m, 2H; $NNCH_{2pip}$), 1.61 (m, 3H; CH_{2pip}), 1.45 (m, 2H; CH_{2pip}), 1.00 (m, 1H; CH_{2pip}). ^{13}C NMR (127 MHz, $CDCl_3$) δ 200.9 (CHO), 158.3 (NCON), 155.7 ($Ar_{Phenol}COH$), 139.8 ($Ar_{Phenyl}CCH_2N$), 130.3 (Ar_{Phenol}), 128.4 ($2Ar_{Phenyl}$), 127.4 ($2Ar_{Phenyl}$), 126.8 (Ar_{Phenyl}), 115.7 (Ar_{Phenol}), 60.5 (NHCH), 53.6 ($NNCH_{2pip}$), 53.4 ($NNCH_{2pip}$), 42.5 (NCH_2Ph), 34.7 (CH_2Phenol), 26.4 (CH_{2pip}), 26.3 (CH_{2pip}), 23.1 (CH_{2pip}).

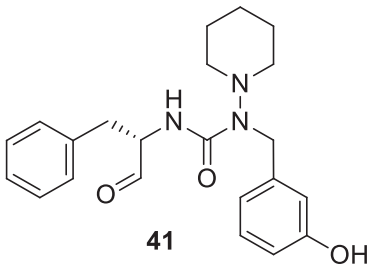
(S)-1-Benzyl-3-(1-(4-acetamido)benzyl-2-oxo-ethyl)-1-piperidin-1-yl-urea (40)



40

From acetal **29** (72 mg, 0.15 mmol). Purification by flash chromatography (gradient 1:4 to 1:1, EtOAc: Cyclohexane), gave **40** as a colorless oil (59 mg, 91 %). ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.73 (s, 1H; CHO), 7.88 (br s, NHAc), 7.48 (d, J = 8.1 Hz, 2H; Ar_{NHAc}), 7.30-7.24 (m, 5H; Ar_{Phenyl}), 7.15 (d, J = 8.1 Hz, 2H; Ar_{NHAc}), 7.02 (d, J = 6.7 Hz, 1H; NH_{urea}), 4.62 (s, 2H; NCH_2Ph), 4.54 (m, 1H; NHCH), 3.17 (dd, J = 6.7 Hz, J_{AB} = 13.5 Hz, 1H; C$H_A$$H_B$ArNHAc), 3.09 (dd, J = 7.4 Hz, J_{AB} = 13.5 Hz, 1H; $CH_A$$H_B$ArNHAc), 2.76 (m, 1H; $NNCH_{2pip}$), 2.63 (m, 1H; $NNCH_{2pip}$), 2.53 (m, 2H; $NNCH_{2pip}$), 2.15 (s, 3H; Ac), 1.64 (m, 3H; CH_{2pip}), 1.46 (m, 2H; CH_{2pip}), 1.03 (m, 1H; CH_{2pip}). ^{13}C NMR (127 MHz, $CDCl_3$) δ 200.8 (CHO), 168.7 (CO_{Ac}), 158.1 (NCON), 140.0 ($Ar_{Phenyl}CCH_2N$), 137.2 ($Ar_{NHAc}CCH_2$), 132.0 (Ar*C*NHAc), 129.8 ($2Ar_{NHAc}$), 128.4 ($2Ar_{Phenyl}$), 127.4 ($2Ar_{Phenyl}$), 126.9 (Ar_{Phenyl}), 120.1 ($2Ar_{NHAc}$), 60.5 (NHCH), 53.6 (d, 2 x $NNCH_{2pip}$), 42.4 (NCH_2Ph), 34.9 (CH_2Ar_{NHAC}), 26.4 (d, 2 x CH_{2pip}), 24.6 (CH_{3Ac}), 23.2 (CH_{2pip}).

(S)-1-(3-hydroxy-benzyl)-3-(1-benzyl-2-oxo-ethyl)-1-piperidin-1-yl-urea (41)

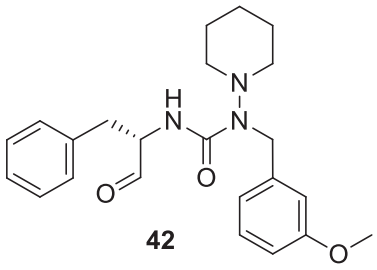


41

From acetal **34a** (78 mg, 0.18 mmol). Purification by flash chromatography (50 % EtOAc/cyclohexane), gave **42** as a colorless oil (41 mg, 59 %). ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.71 (s, 1H; CHO), 7.28 (m, 5H; Ar_{Phenyl}), 7.22 (d, J = 7.8 Hz, 1H; NH), 7.14 (t, 1H; Ar_{Phenol}), 7.00 (d, 1H; Ar_{Phenol}), 6.89 (s, 1H; Ar_{Phenol}), 6.80 (d, 1H; Ar_{Phenol}), 6.71 (d, 1H; Ar_{Phenol}), 4.59 (m, 1H; NHCH), 4.56 (s, 2H; NCH_2Phenol) 3.20 (dd, J = 6.5 Hz, J_{AB} = 14.0 Hz, 1H; C$H_A$$H_B$Ph), 3.16 (dd, J = 7.2 Hz, J_{AB} = 14.0 Hz, 1H; $CH_A$$H_B$Ph), 2.74 (m, 1H; $NNCH_{2piperidine}$), 2.54 (m, 1H; $NNCH_{2pip}$), 2.50 (m, 2H; $NNCH_{2pip}$), 1.67 (m, 3H; CH_{2pip}), 1.42 (m, 2H; CH_{2pip}), 1.00 (m, 1H; CH_{2pip}). ^{13}C NMR (127 MHz, $CDCl_3$) δ 200.6 (CHO), 158.2 (NCON), 156.4 ($Ar_{Phenol}COH$), 141.7 ($Ar_{Phenol}CCH_2N$),

136.3 (Ar_{Phenyl}), 129.4 ($2Ar_{Phenyl}$+Ar_{Phenol}), 128.8 ($2Ar_{Phenyl}$), 127.0 (Ar_{Phenyl}), 119.3 (Ar_{Phenol}), 114.6 (Ar_{Phenol}), 114.2 (Ar_{Phenol}), 60.3 (NHCH), 53.5 ($NNCH_{2pip}$), 53.4 ($NNCH_{2pip}$), 42.3 (CH_2Phenol), 35.4 (CH_2Ph), 26.4 (CH_{2pip}), 26.3 (CH_{2pip}), 23.1 (CH_{2pip}).

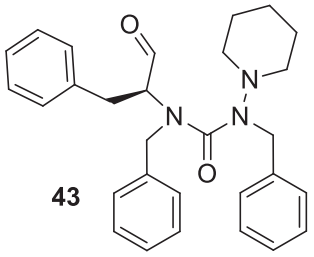
(S)-1-(3-methoxy-benzyl)-3-(1-benzyl-2-oxo-ethyl)-1-piperidin-1-yl-urea (42)



42

From acetal **35a** (63 mg, 0.14 mmol). Purification by flash chromatography (50 % EtOAc/cyclohexane), gave **42** as a colorless oil (38 mg, 67 %). ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.77 (s, 1H; CHO), 7.24-7.32 (m, 5H; $5Ar_{Phenyl}$+$1Ar_{Phenol}$), 6.96 (d, 1H; Ar_{Phenol}), 6.92 (s, 1H; Ar_{Phenol}), 6.80 (d, 1H; Ar_{Phenol}), 4.61 (s, 2H; NCH_2Phenol), 4.58 (m, 1H; NHC*H*), 3.83 (s, 3H; OCH_3), 3.23 (dd, *J* = 6.5 Hz, J_{AB} = 14.0 Hz, 1H; CH_AH$_B$Ph), 3.17 (dd, *J* = 7.2 Hz, J_{AB} = 14.0 Hz, 1H; CH$_A$$H_B$Ph), 2.76 (m, 1H; $NNCH_{2piperidine}$), 2.61 (m, 1H; $NNCH_{2pip}$), 2.55 (m, 2H; $NNCH_{2pip}$), 1.64 (m, 3H; CH_{2pip}), 1.41 (m, 2H; CH_{2pip}), 1.05 (m, 1H; CH_{2pip}). ^{13}C NMR (127 MHz, $CDCl_3$) δ 200.6 (CHO), 159.7 (Ar_{Phenol}COMe), 158.0 (NCON), 141.9 (Ar_{Phenol}*C*CH_2N), 136.6 (Ar_{Phenyl}), 129.4 ($2Ar_{Phenyl}$+Ar_{Phenol}), 128.7 ($2Ar_{Phenyl}$), 127.0 (Ar_{Phenyl}), 119.8 (Ar_{Phenol}), 113.2 (Ar_{Phenol}), 112.2 (Ar_{Phenol}), 60.3 (NHCH), 55.2 (OCH_3), 53.5 ($2NNCH_{2pip}$), 42.4 (CH_2Phenol), 35.3 (CH_2Ph), 26.4 (CH_{2pip}), 26.3 (CH_{2pip}), 23.2 (CH_{2pip}).

(S)-1,3-dibenzyl-1-(1-oxo-3-phenylpropan-2-yl)-3-(piperidin-1-yl)urea (43)



43

From acetal **33b** (46 mg, 0.09 mmol). Purification by flash chromatography (gradient 1:2 to 1:1, EtOAc: Cyclohexane), gave **43** as a colorless oil (20 mg, 48 %). ^{1}H NMR (500 MHz, CD_2Cl_2) δ 9.47 (s, 1H; CHO), 7.39-7.07 (m, 15H; Ar), 4.74 (d, J_{AB} = 15.2 Hz, 1H; CHNCH_2Ph), 4.47 (m, 2H; $NNCH_2$Ph), 3.83 (d, J_{AB} = 15.2 Hz, 1H; CHNCH_2Ph), 3.67 (s, 1H; NHCH), 3.33 (dd, *J* = 4.5 Hz, J_{AB} = 14.2 Hz, 1H; CH_2Ph), 3.06 (dd, *J* = 9.8 Hz, J_{AB} = 14.2 Hz, 1H; CH_2Ph), 2.82 (br, 4H; $NNCH_{2pip}$), 1.43 (br, 6H; CH_{2pip}). ^{13}C NMR (127 MHz, CD_2Cl_2) δ 199.3 (CHO), 161.6 (NCON), 139.6 (Ar), 138.7 (Ar), 138.3 (Ar), 129.2 (Ar), 128.49 (Ar), 128.45 (Ar), 128.2 (Ar), 128.0 (Ar), 127.6 (Ar), 127.1 (Ar), 126.7 (Ar), 126.4 (Ar), 68.0 (NHCH), 53 (under solvent signal, NCH_2Ph), 52.95 ($NNCH_{2pip}$), 43.6 ($NNCH_2$Ph), 33.7 (CH*C*H_2Ph), 25.6 (CH_{2pip}), 23.4 (CH_{2pip}).

2. Biological tests

2.1 In vitro inhibition assay

The capacity of our compounds to inhibit HIV PR was tested in a well-established continuous FRET (fluorescence resonance energy transfer) assay. (ref : Bell, M. R.; Archer, S. J. Am. Chem. Soc. **1960**, *82*, 151. and Kirby, A. J.; Komarov, I. V.; Bilenko, V. A.; Davies, J. E.; Rawson, J. M. Chem. Commun. **2002**, 2106.) Pepstatine was tested simultaneously as standard. Initially, we determined the kinetic parameters of our recombinant source of HIV PR with a commercial FRET substrate in order to evaluate our assay performance in comparison to the literature data. The potency of the inhibitors was first estimated by determining their IC_{50} values by fitting the experimental

data to $v/v_0 = 1/(1 + [I]/IC_{50})$. The exact K_i constants were derived from the same set of experiments (Table 1).

Table 1 show the inhibitory effect (Ki values obtained by use of three different methods) of hydrazine ureas towards HIV-1 peptidase.

Compound	K_i (μM)		
	Lineweaver and Burk	Hanes	Dixon
36	528	406	492
37	172	257	292
38	317	366	996
39	395	342	306
41	272	242	144
42	320	410	334
43	463	1122	993
pepstatine	0.051	0.052	0.036

Compound	K_i		
	Lineweaver and Burk	Hanes	Dixon
36	- -	-	-
37	+++	++	++
38	+	+	- -
39	-	+	+
41	++	++	+++
42	+	-	+
43	- -	- - -	- -

HIV PR was purchased from BioVendor, and HIV PR substrate 1 (DABCYL-GABA-Ser-Asn-Tyr-Pro-Ile-Val-Gln-EDANS) was obtained from TebuBio. Enzymatic activity and inhibition were determined by a fluorogenic assay using a microplate spectrofluorimeter (Spectramax Gemini XS, Molecular Devices) and black 384-well plates (NUNCLONE, Nunc Inc.). Hydrolysis of the fluorogenic substrate was investigated at pH 4.7, 37 °C in 0.1 M acetate, 1 M NaCl, 1 mM EDTA, 1 mM dithiothreitol, and 1 mg/mL bovine serum albumin. The entire time course was recorded by monitoring the fluorescence at 490 nm with an excitation wavelength set to 340 nm. Each initial rate was the average value of five independently performed reactions. KM determination was performed using a range of fluorogenic substrate concentrations from 20 to 80 μM and an HIV PR concentration of 12.3 nM. Final solutions contained 10% DMSO. For determination of IC50 and Ki values, inhibitors were preincubated with HIV PR for 30 min. IC_{50} determinations were performed using 10% DMSO solutions containing a final of 12.3 nM HIV PR and 4 μM fluorogenic substrate. Initial rates were determined for at least five inhibitor concentrations. Ki Determinations were performed using 10% DMSO solutions containing a final of 12.3 nM HIV PR. Initial rates were determined for at least five inhibitor concentrations and four fluorogenic substrate concentrations (2, 3, 4, and 5 μM).

2.2 ABC pumps transport inhibition

It is known that the acquired resistance to mitoxatrone is due to the enhanced drug efflux via P-gp or BCRP over-expression. Mitoxantrone efflux mediated by P-gp and BCRP is recorded by flow cytometry in the presence or absence of componds and 2 know inhibitors of both pumps, GF120918 (elacridar) and Ko123.

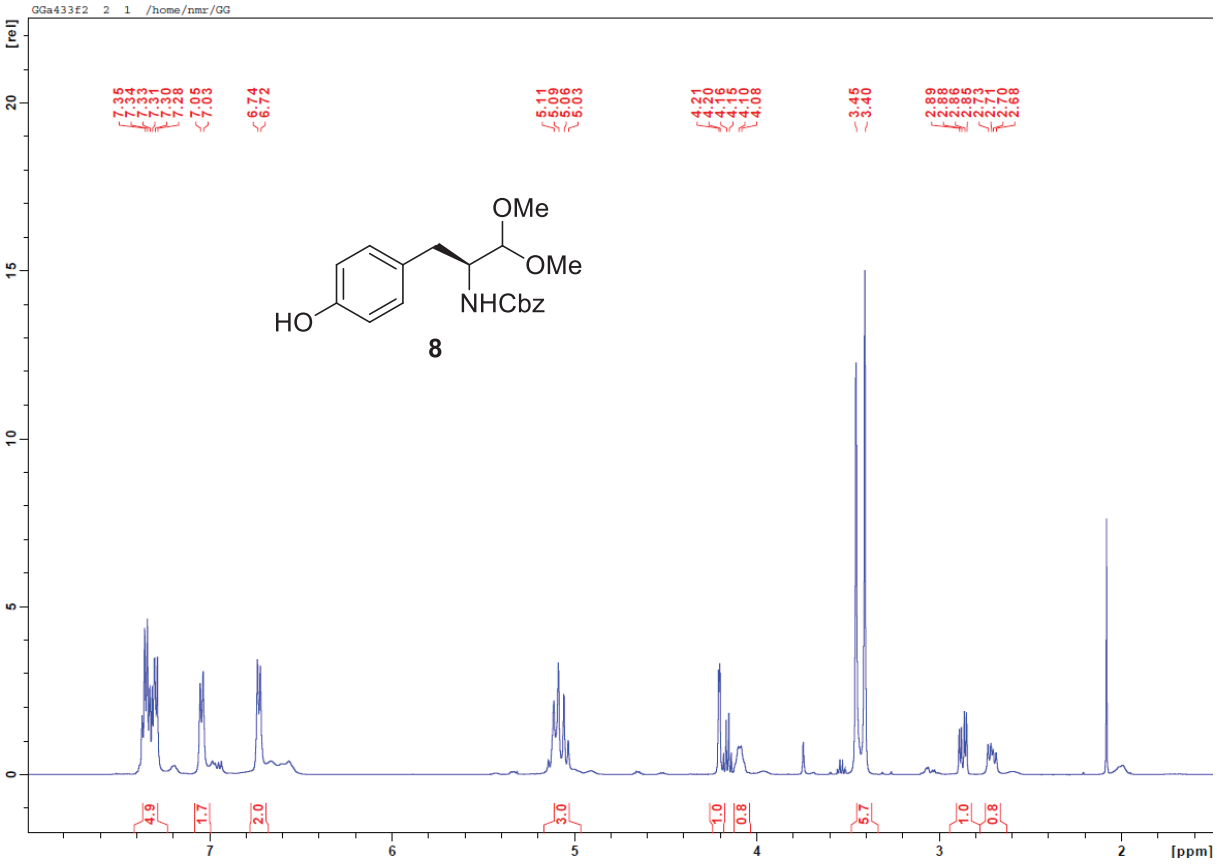
Fifty thousand cells, either control (NIH3T3) or expressing Pgp (NIH3T3 MDR G185), and one hundred thousand cells, either control (pcDNA3) or expressing BCRP (BCRP) were incubated for 24 hours in growth medium and then one hour at 37°C in DMEM with 5 µM mitoxantrone in the presence of 2 µM or 10 µM of each compound of the present study. Cells were washed with PBS, trypsinised, and the mitoxantrone intracellular accumulation was quantified by fluorescence coupled to flow cytometry carried out on a FACS Calibur cytometer (Becton Dickinson), exciting at 488 nm and emitting at 650 nm. Data was collected on CellQuest Pro (version 4.0) software and then exported to FlowJo for analysis. Mitoxantrone (Sigma Aldrich) was dissolved in 100 % DMSO at 20 mM stock concentration and stored at – 20 °C.

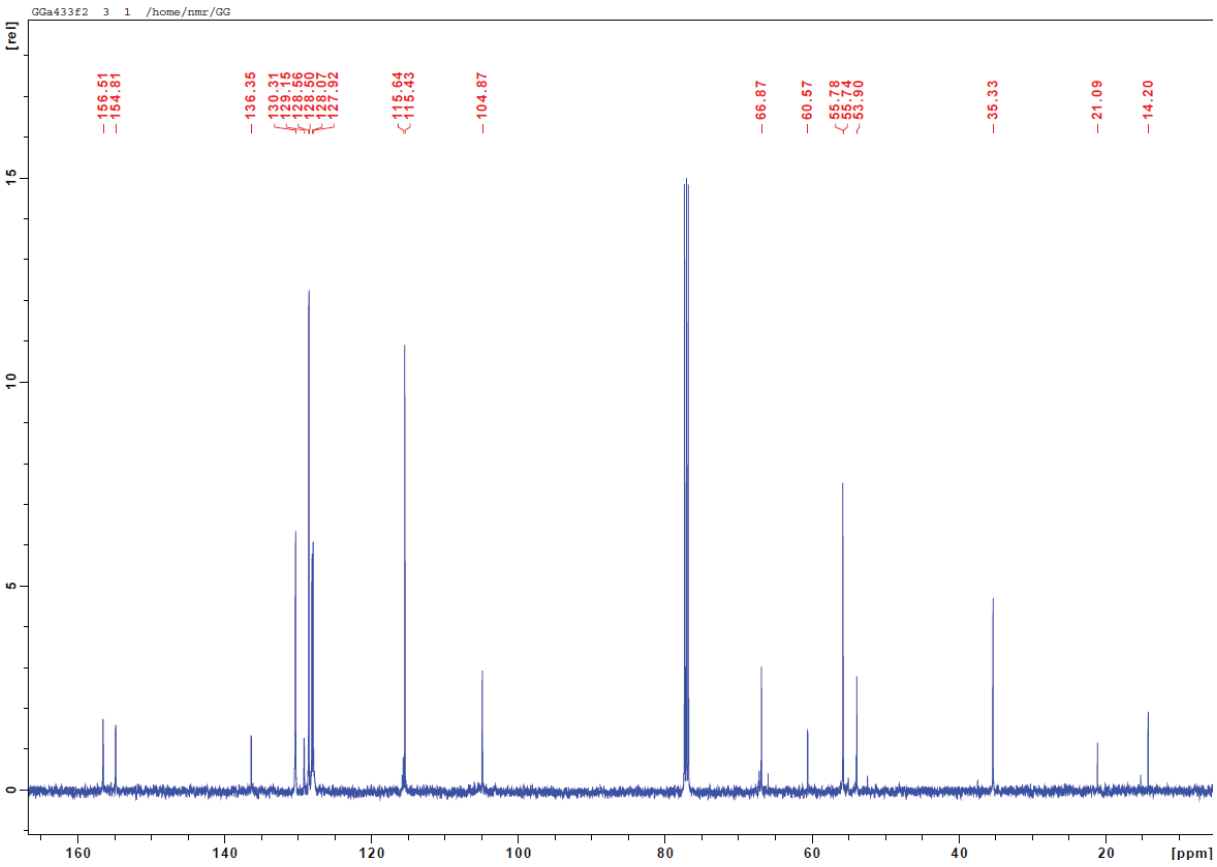
Table 2 show that compounds do not inhibit the efflux of mitoxantrone compared whit GF120918 or Ko123.

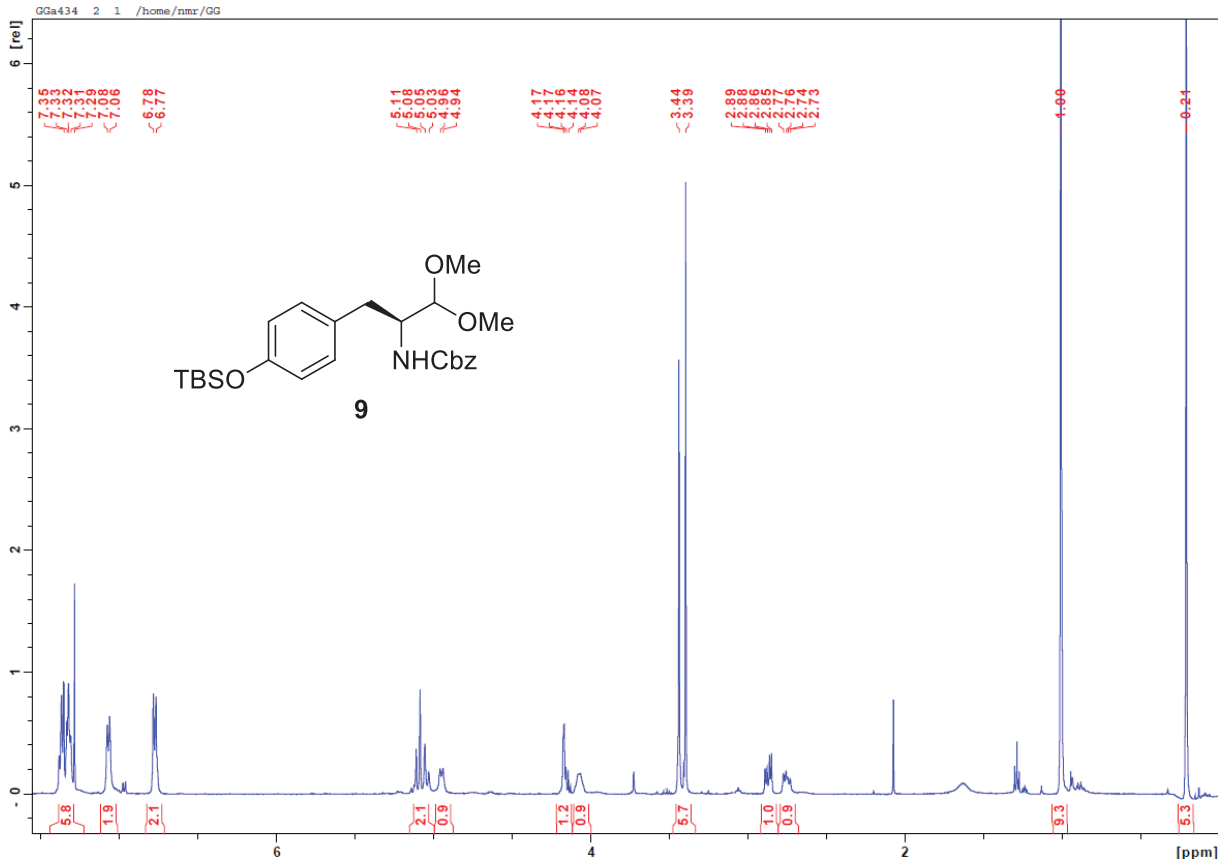
Table 2. Inhibition of mitoxantrone efflux by P-glycoprotein or BCRP. Data are the mean ± SD of at least three independent experiments.

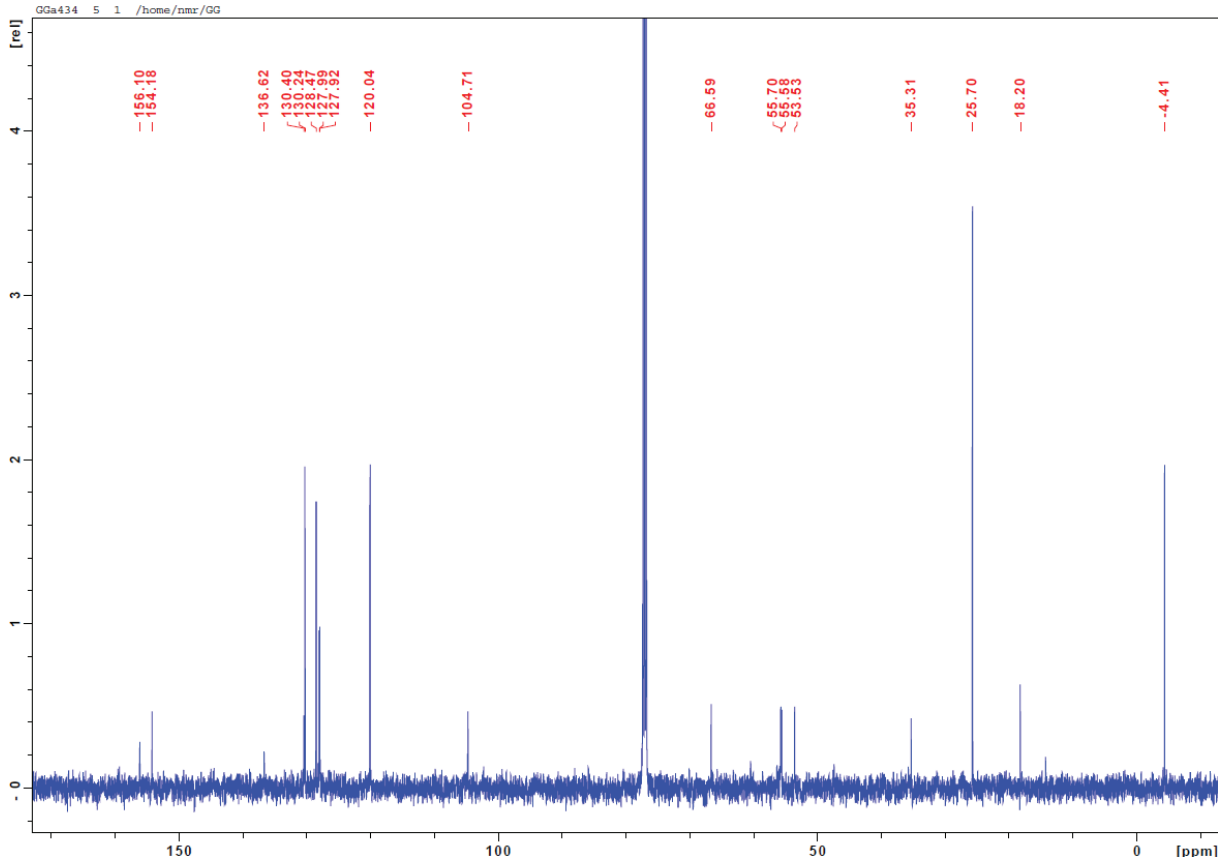
Compound	Inhibition of P-gp (%)		Inhibition of BCRP (%)	
	At 2 µM	At 10 µM	At 2 µM	At 10 µM
36	2.3±0.5	8.1±0.3	0.2±0.03	8.1±0.3
37	5.5±0.5	6.4±0.7	0.2±0.05	0.3±0.01
38	7.1±0.3	11.5±1.2	0.2±0.02	3.3±1.7
39	1.1±0.1	3.0±0.7	0.9±0.6	2.5±0.9
41	4.6±0.2	4.8±0.1	2.5±0.6	8.4±1.1
42	7.7±0.1	22.7±0.5	0.9±0.5	8.1±0.3
43	1.8±0.2	15±2.0	1.8±0.2	8.8±1.1
GF120918	110±2.5			
Ko143			100±1.5	

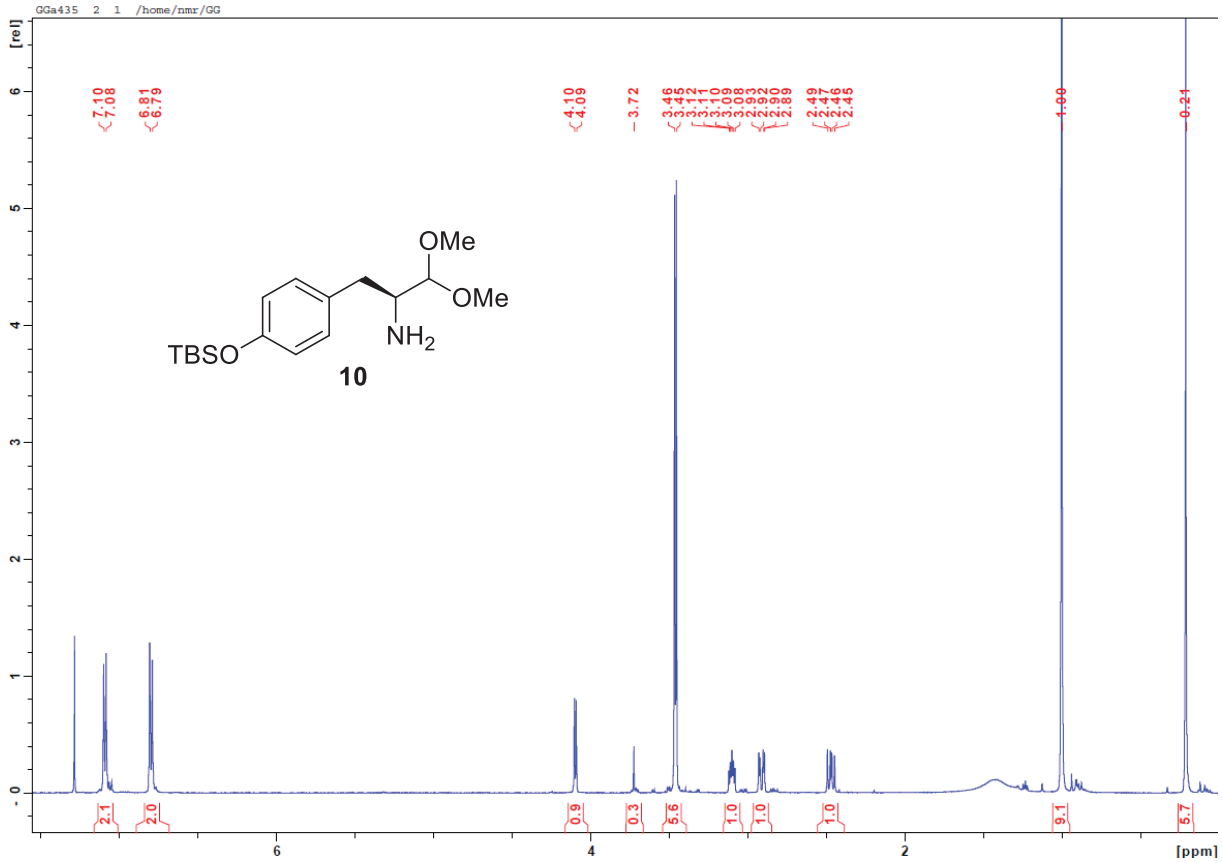
3. NMR spectra of new compounds

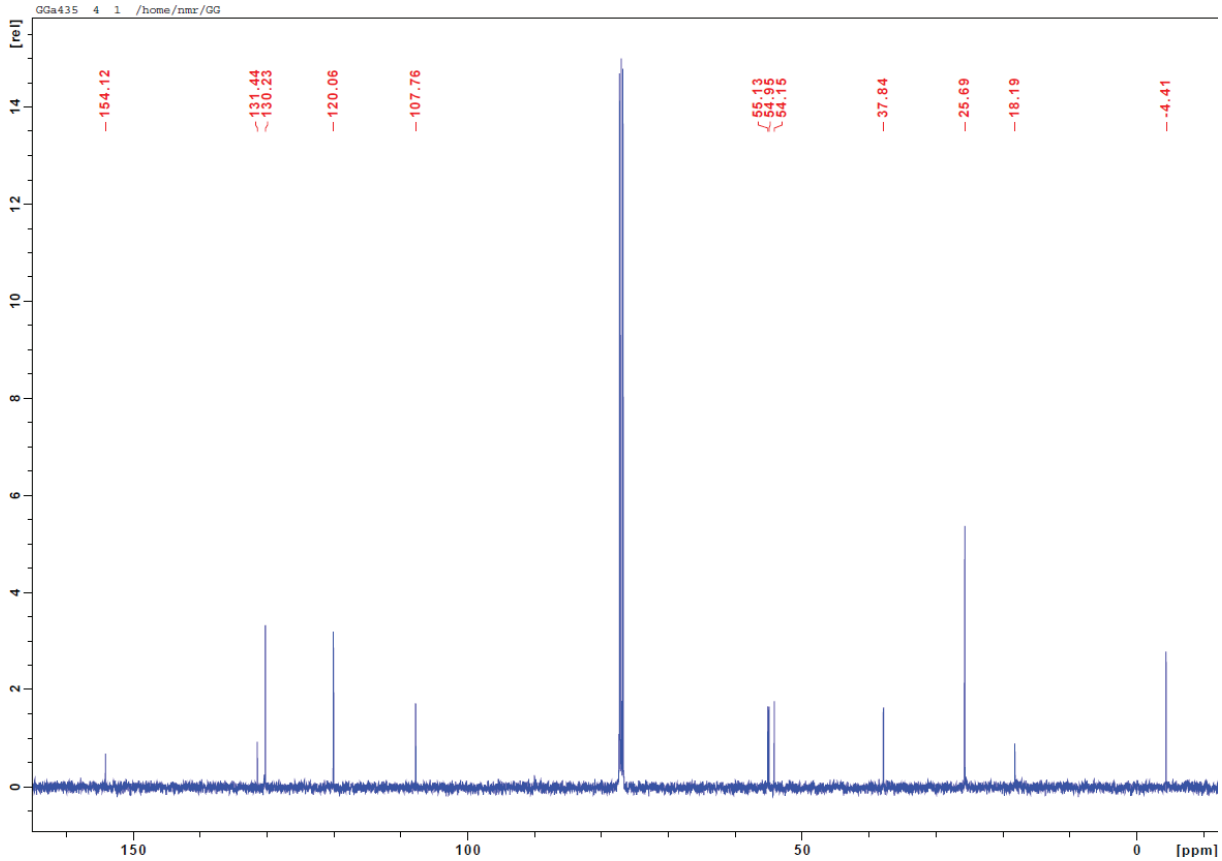




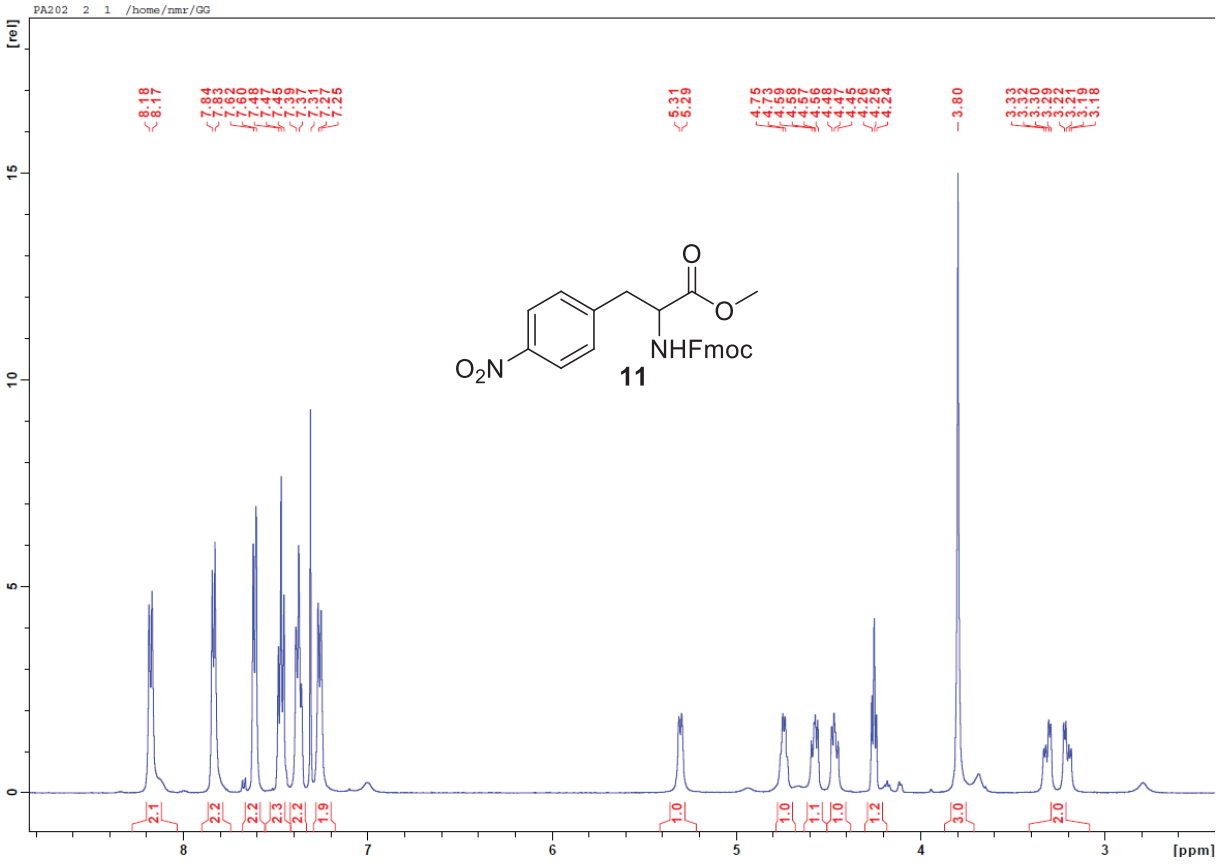

GGa434 2 1 /home/nmr/GG
[rel]
7.35
7.33
7.31
7.29
7.08
7.06
6.78
6.77
5.11
5.08
5.05
5.03
4.96
4.94
4.17
4.16
4.14
4.08
4.07
3.44
3.39
2.89
2.88
2.86
2.77
2.76
2.74
2.73
1.00
0.24
5.8
1.9
2.1
2.1
0.9
1.2
0.9
5.7
1.0
0.9
9.3
5.3
OMe
OMe
NHCbz
TBSO
9
[ppm]

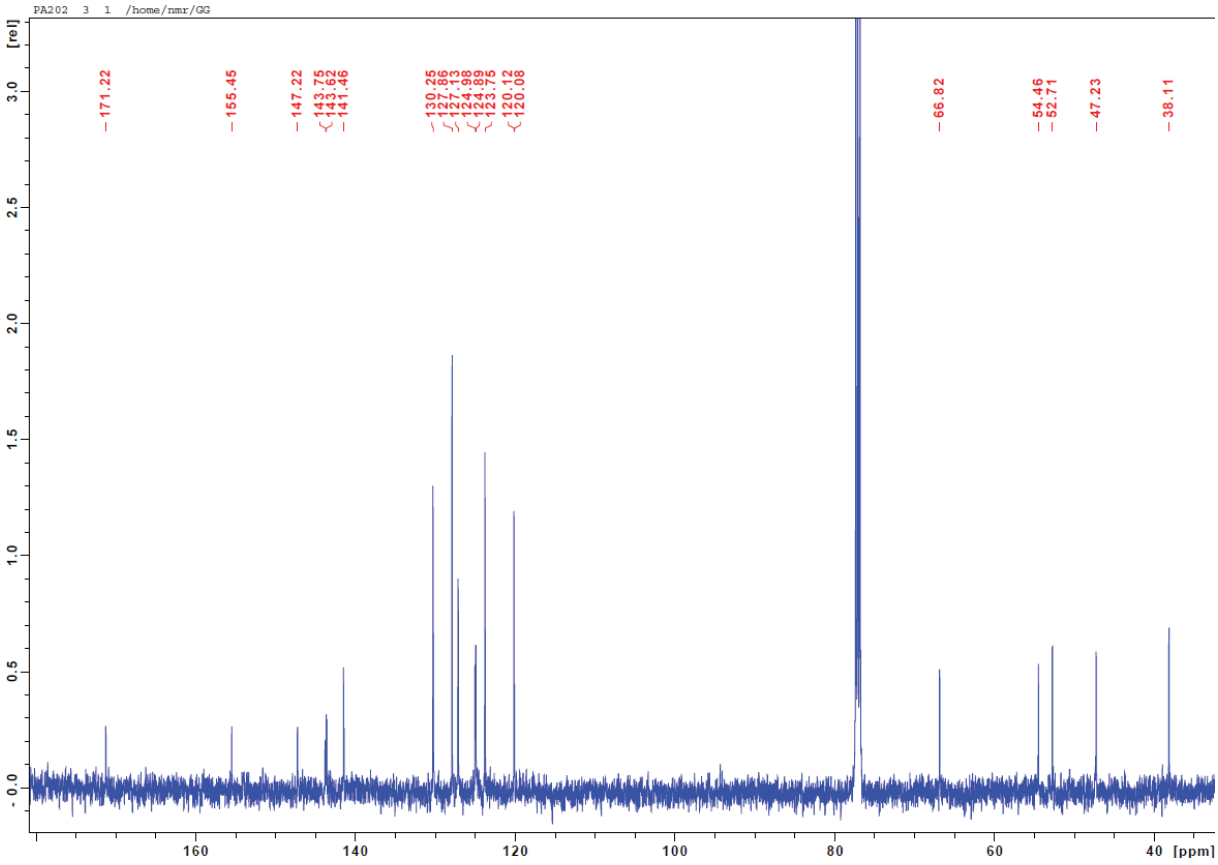

GGa434 5 1 /home/nmr/GG
[rel]
156.10
154.18
136.62
130.40
130.24
128.47
127.99
127.92
120.04
104.71
66.59
55.70
55.58
53.53
35.31
25.70
18.20
-4.41
[ppm]


GGa435 2 1 /home/nmr/GG
[rel]
7.10
7.08
6.81
6.79
4.10
4.09
3.72
3.46
3.45
3.12
3.11
3.10
3.09
3.08
2.93
2.92
2.90
2.89
2.49
2.47
2.46
2.45
1.00
0.21
OMe
OMe
NH2
TBSO
10
2.1
2.0
0.9
0.3
5.6
1.0
1.0
1.0
9.1
5.7
[ppm]

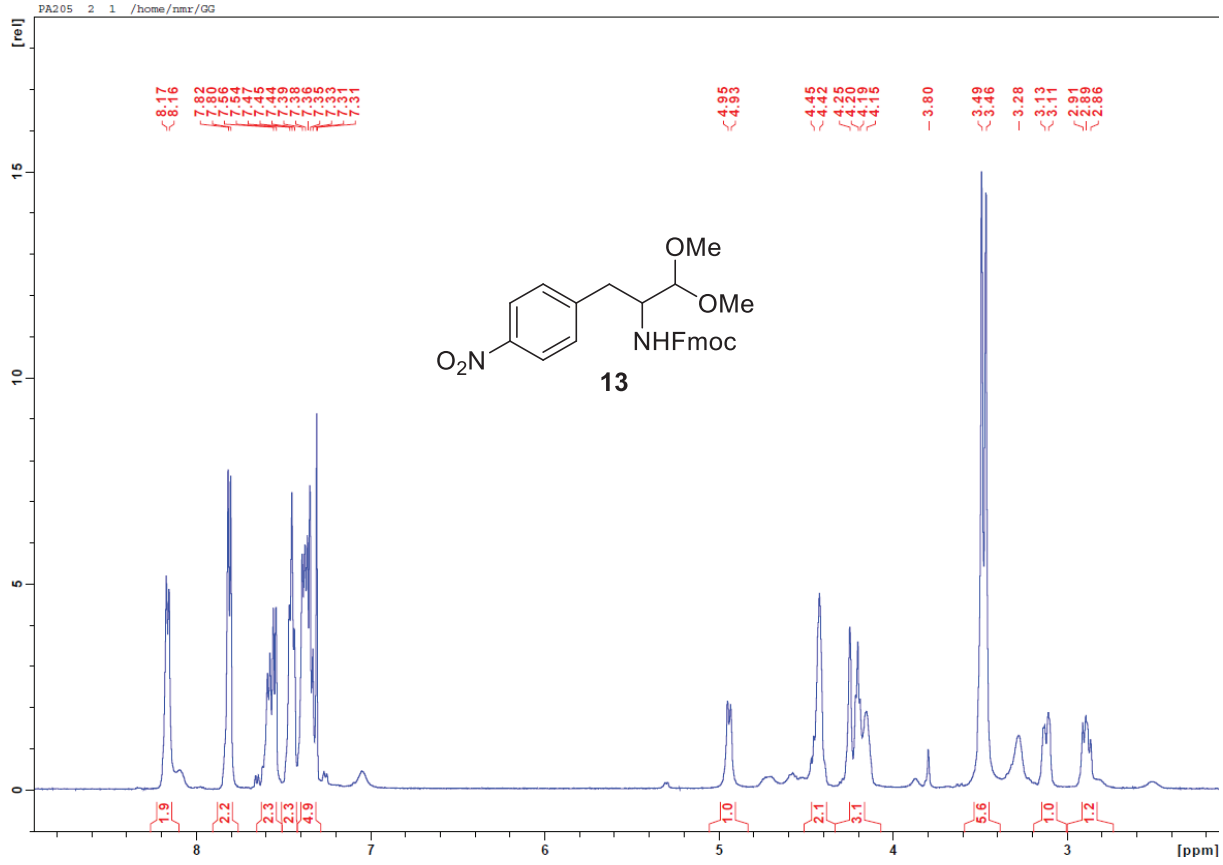

GGa435 4 1 /home/nmr/GG
[rel]
154.12
131.44
130.23
120.06
107.76
55.13
54.95
54.15
37.84
25.69
18.19
-4.41
[ppm]

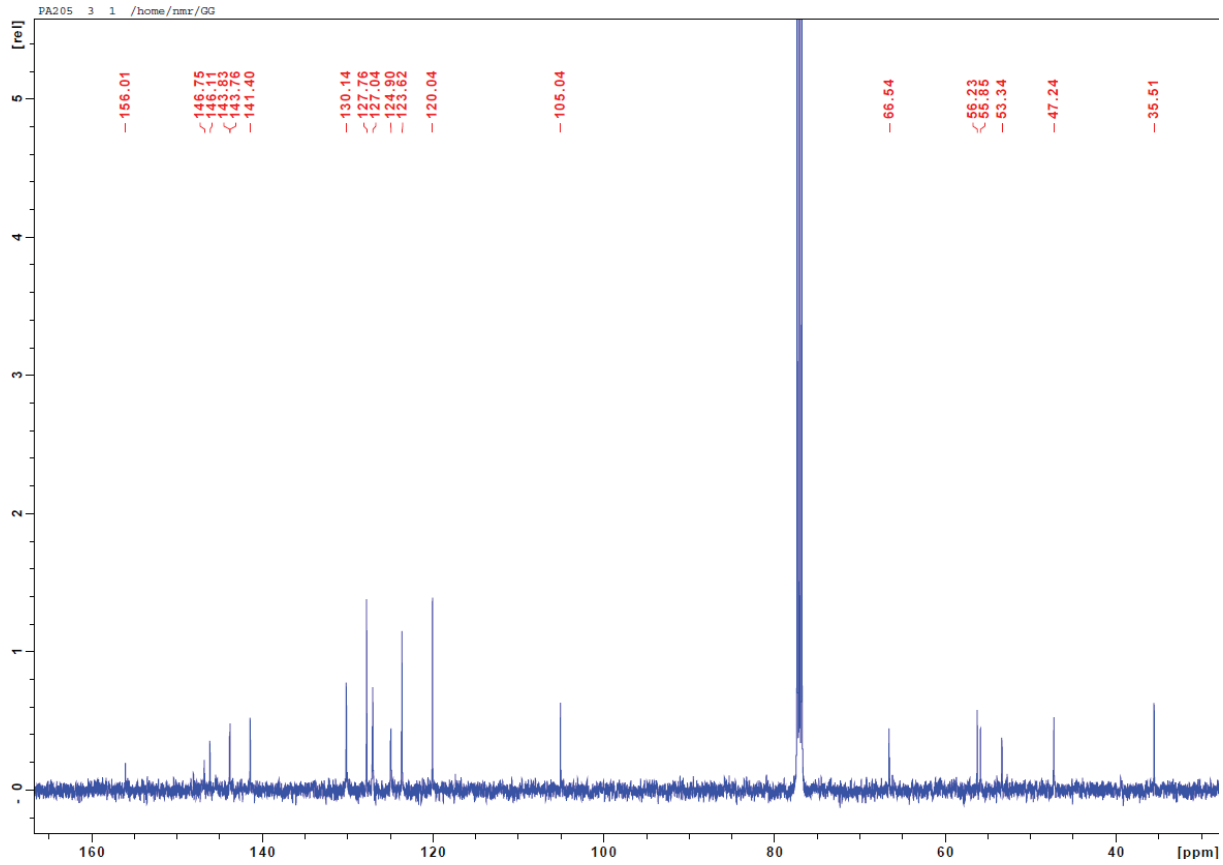
S 21


PA202 2 1 /home/nmr/GG
[rel]
8.18
8.17
7.84
7.83
7.62
7.60
7.48
7.47
7.45
7.39
7.37
7.31
7.27
7.25
5.31
5.29
4.75
4.73
4.59
4.58
4.57
4.56
4.47
4.45
4.26
4.25
4.24
3.80
3.33
3.32
3.30
3.29
3.22
3.21
3.19
3.18
O
O
NHFmoc
O2N
11
2.1
2.2
2.2
2.3
2.2
1.9
1.0
1.0
1.1
1.0
1.2
3.0
2.0
[ppm]

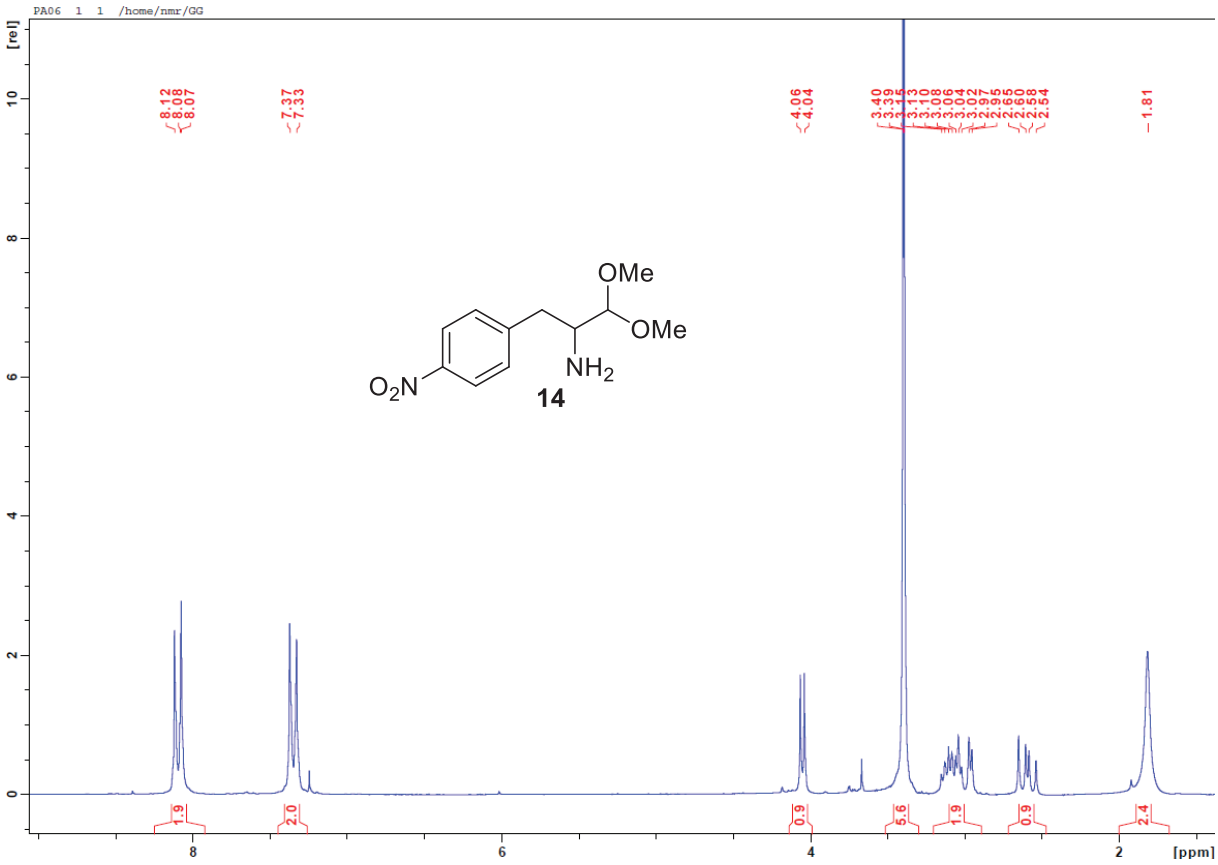

PA202 3 1 /home/nmr/GG
[rel]
171.22
155.45
147.22
143.75
143.62
141.46
130.25
127.86
127.13
124.98
124.69
123.75
120.12
120.08
66.82
54.46
52.71
47.23
38.11
[ppm]

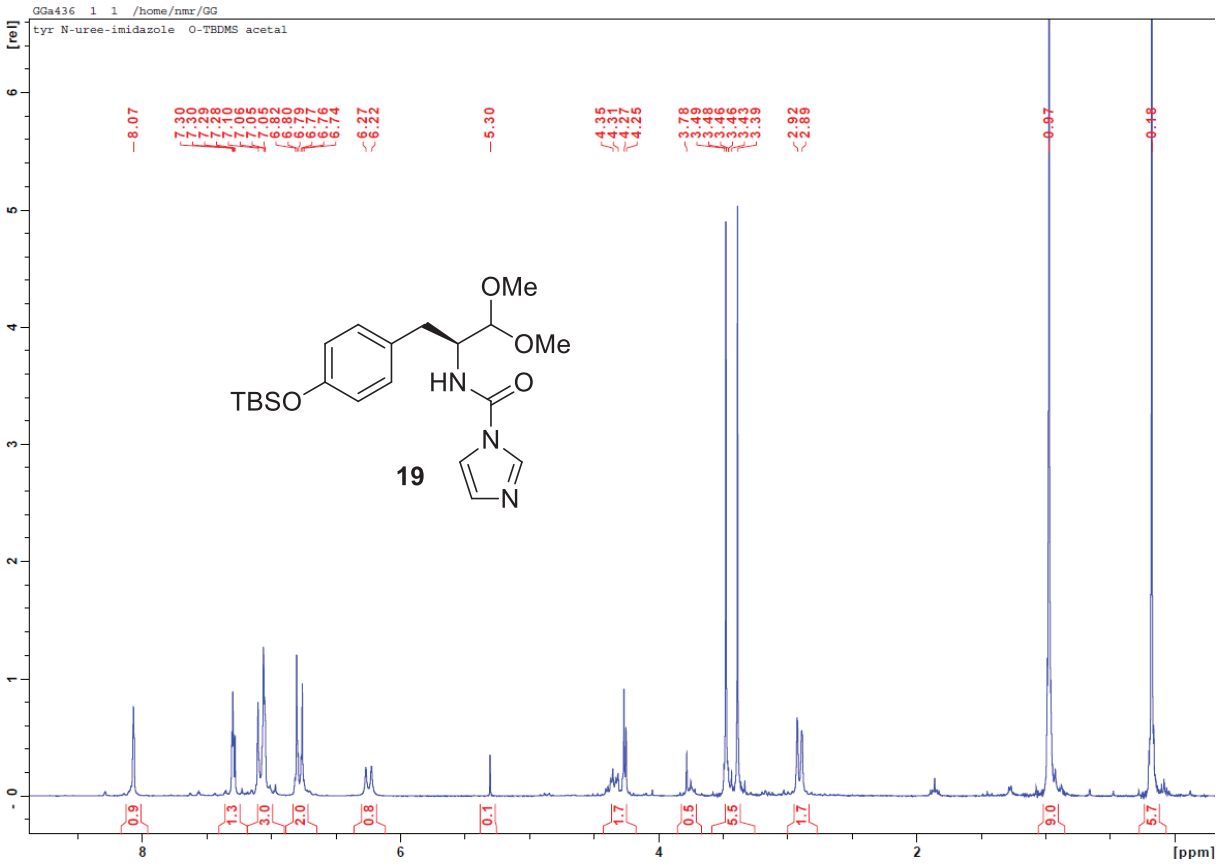
S 22

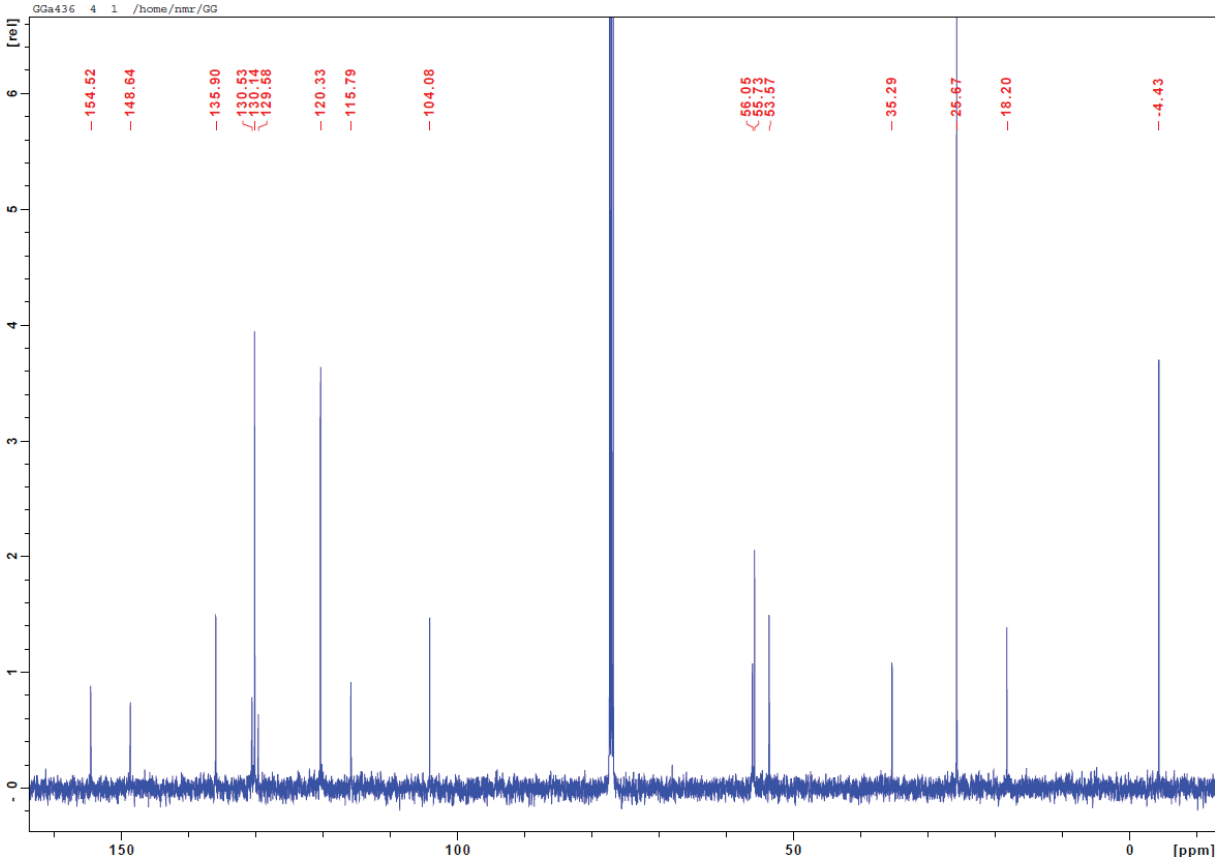

PA205 2 1 /home/nmr/GG
[rel]
8.17
8.16
7.82
7.80
7.56
7.54
7.47
7.45
7.44
7.39
7.38
7.36
7.35
7.33
7.31
4.95
4.93
4.45
4.42
4.25
4.20
4.19
4.15
3.80
3.49
3.46
3.28
3.13
3.11
2.91
2.89
2.86
OMe
OMe
NHFmoc
O2N
13
1.9
2.2
2.3
2.3
4.9
1.0
2.1
3.1
5.6
1.0
1.2
[ppm]

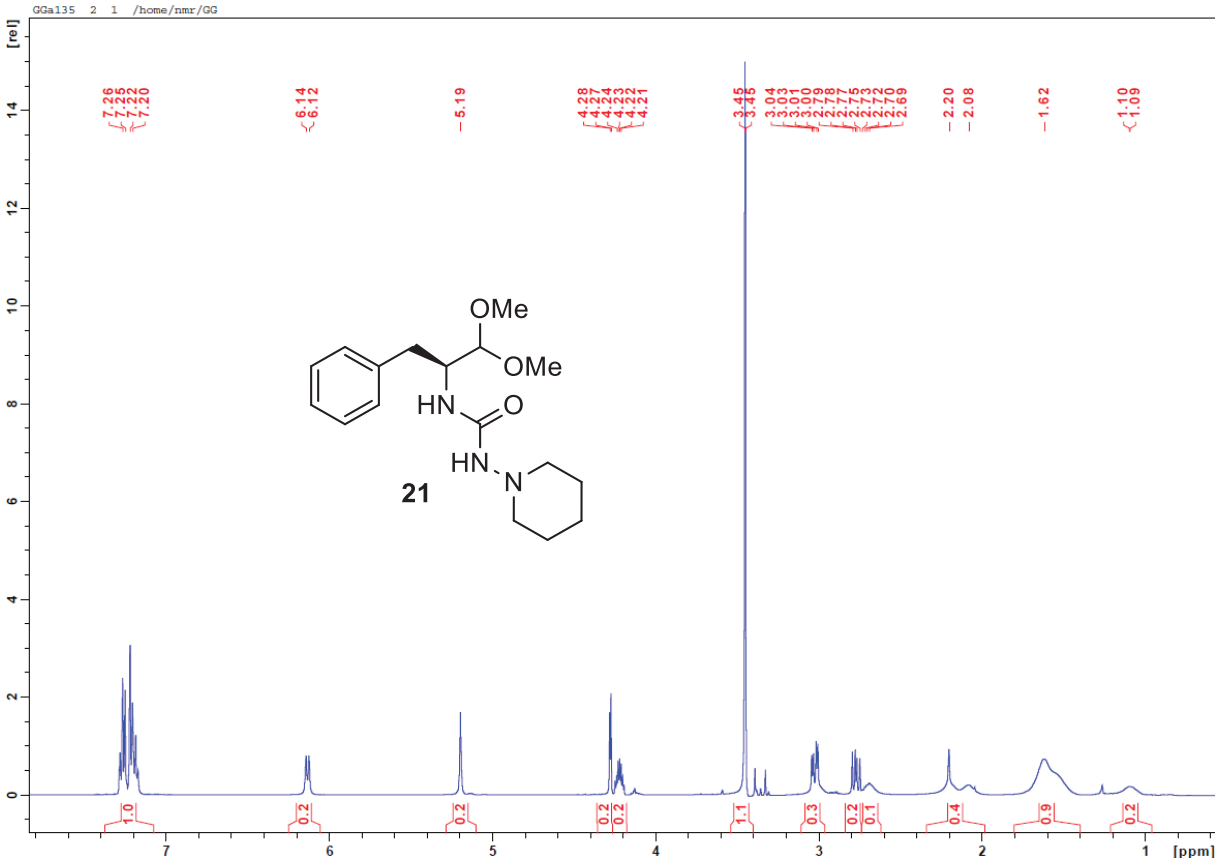

PA205 3 1 /home/nmr/GG
[rel]
156.01
146.75
146.11
143.83
143.76
141.40
130.14
127.76
127.04
124.90
123.62
120.04
105.04
66.54
56.23
55.85
53.34
47.24
35.51
[ppm]

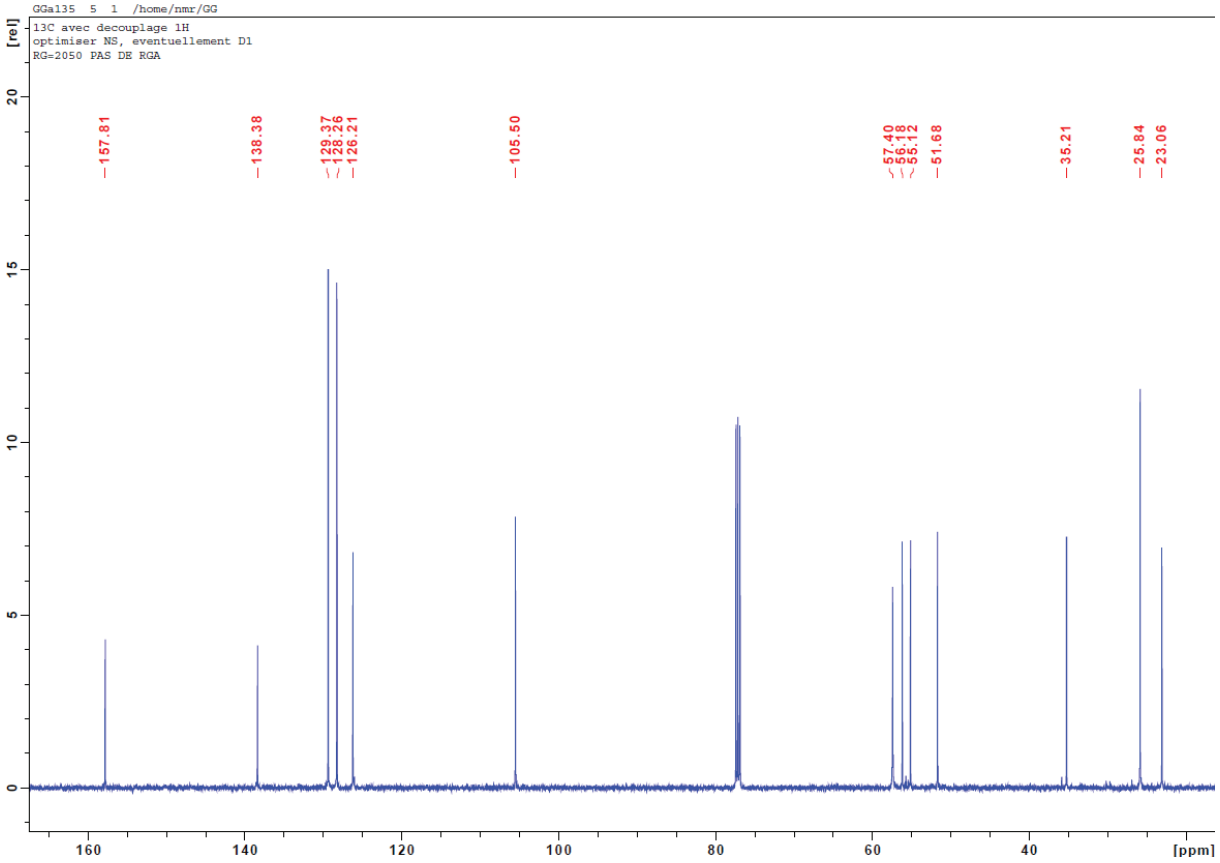
S 23

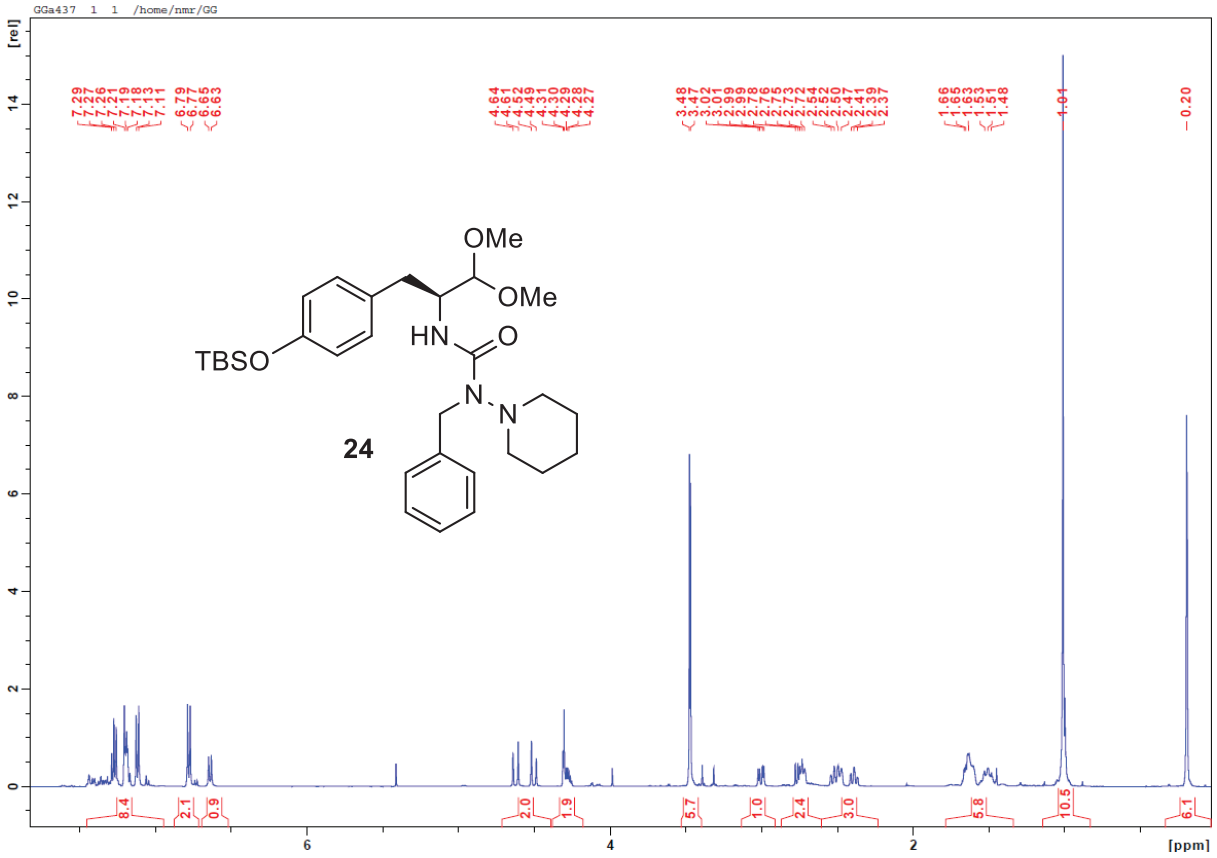

PA06 1 1 /home/nmr/GG
[rel]
8.12
8.08
8.07
7.37
7.33
4.06
4.04
3.40
3.39
3.15
3.13
3.08
3.06
3.04
3.02
2.97
2.95
2.60
2.58
2.54
1.81
OMe
OMe
NH2
O2N
14
1.9
2.0
0.9
5.6
1.9
0.9
2.4
8
6
4
2
[ppm]

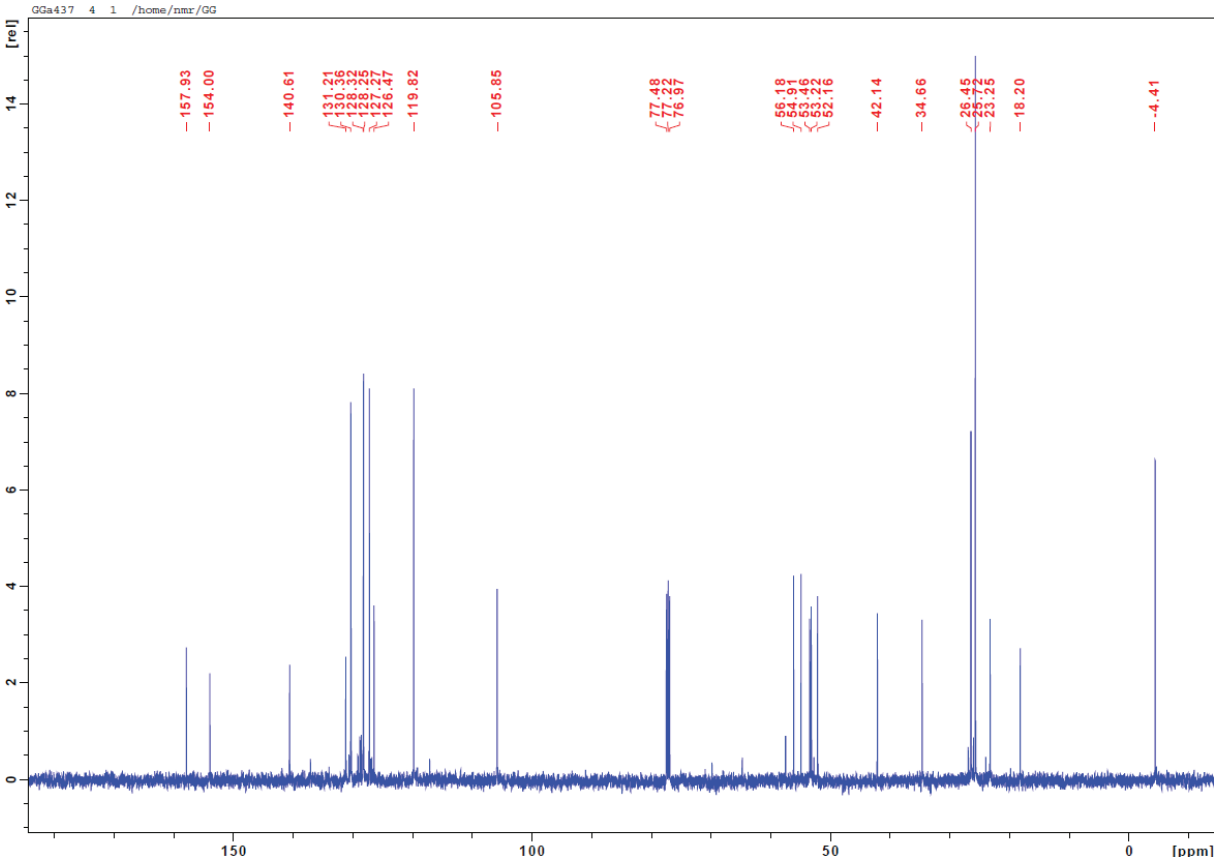

GGa436 1 1 /home/nmr/GG
tyr N-uree-imidazole O-TBDMS acetal
[rel]
[ppm]
OMe
OMe
HN
O
TBSO
N
N
19

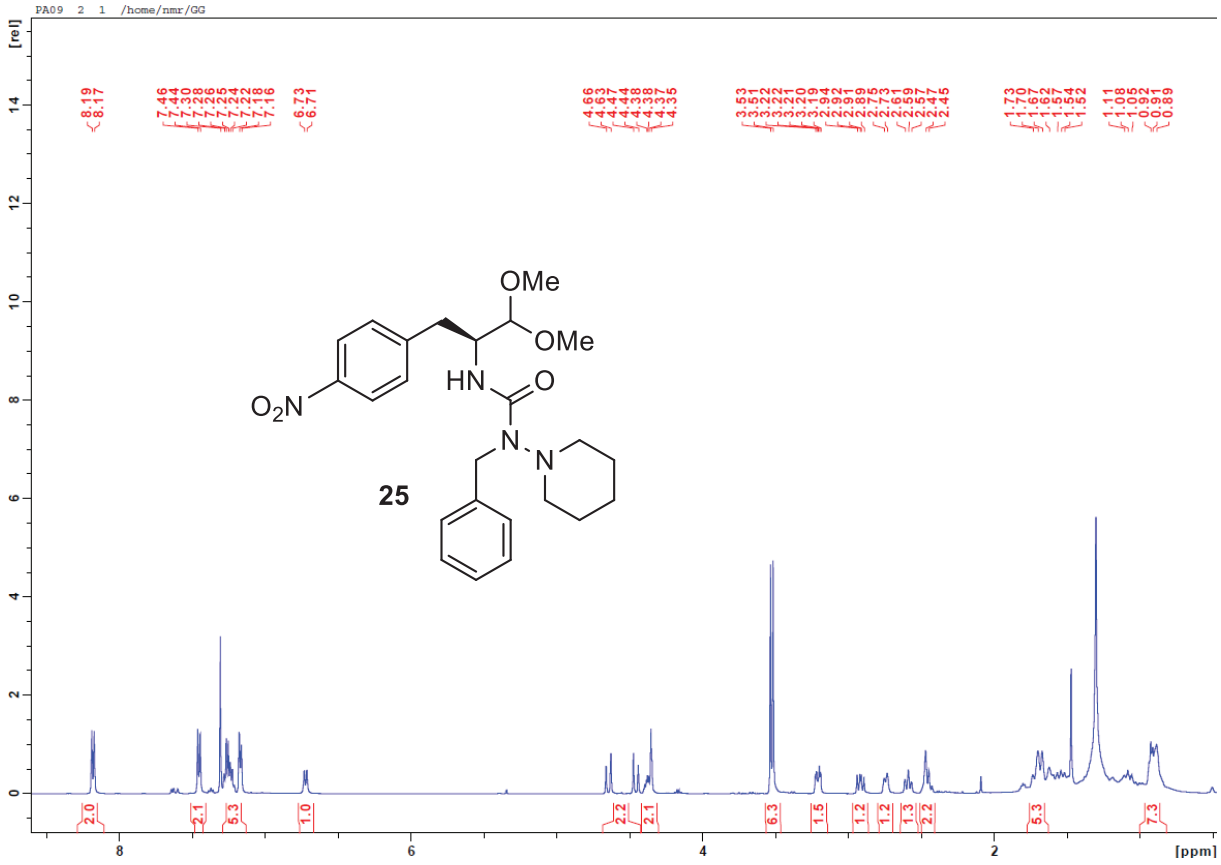

GGa436 4 1 /home/nmr/GG
[rel]
[ppm]

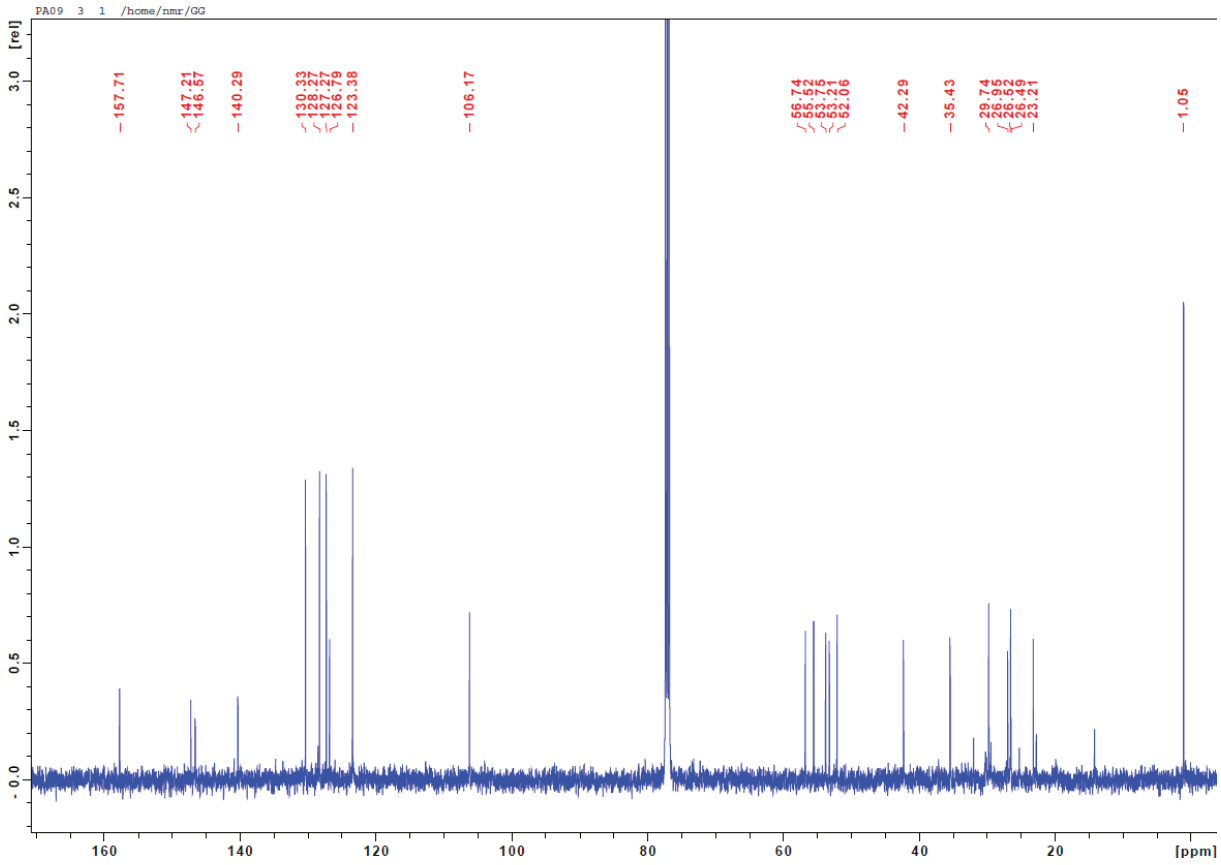

GGa135 2 1 /home/nmr/GG
[rel]
[ppm]
OMe
OMe
HN
O
HN
N
21


GGa135 5 1 /home/nmr/GG
13C avec decouplage 1H
optimiser NS, eventuellement D1
RG=2050 PAS DE RGA
[rel]
[ppm]

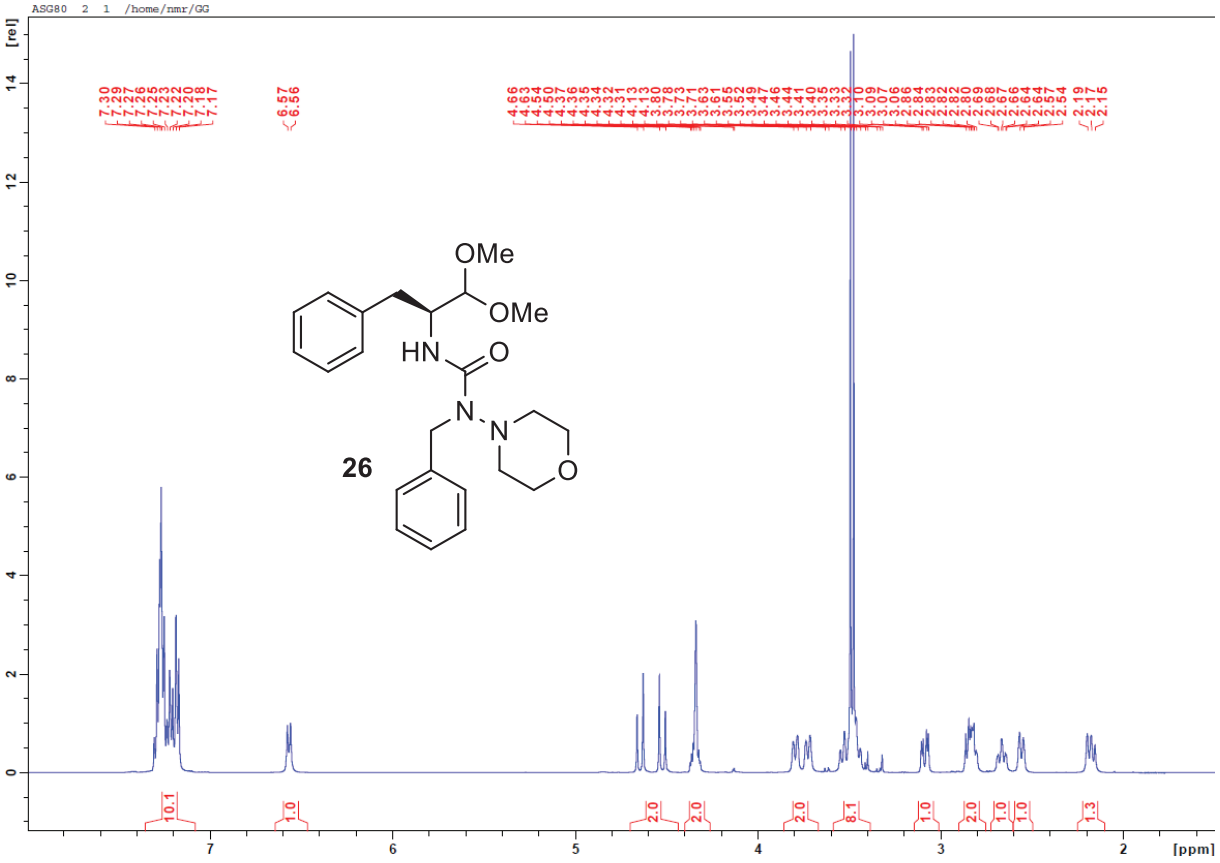

GGa437 1 1 /home/nmr/GG
[rel]
[ppm]
OMe
OMe
HN
O
TBSO
N
N
24

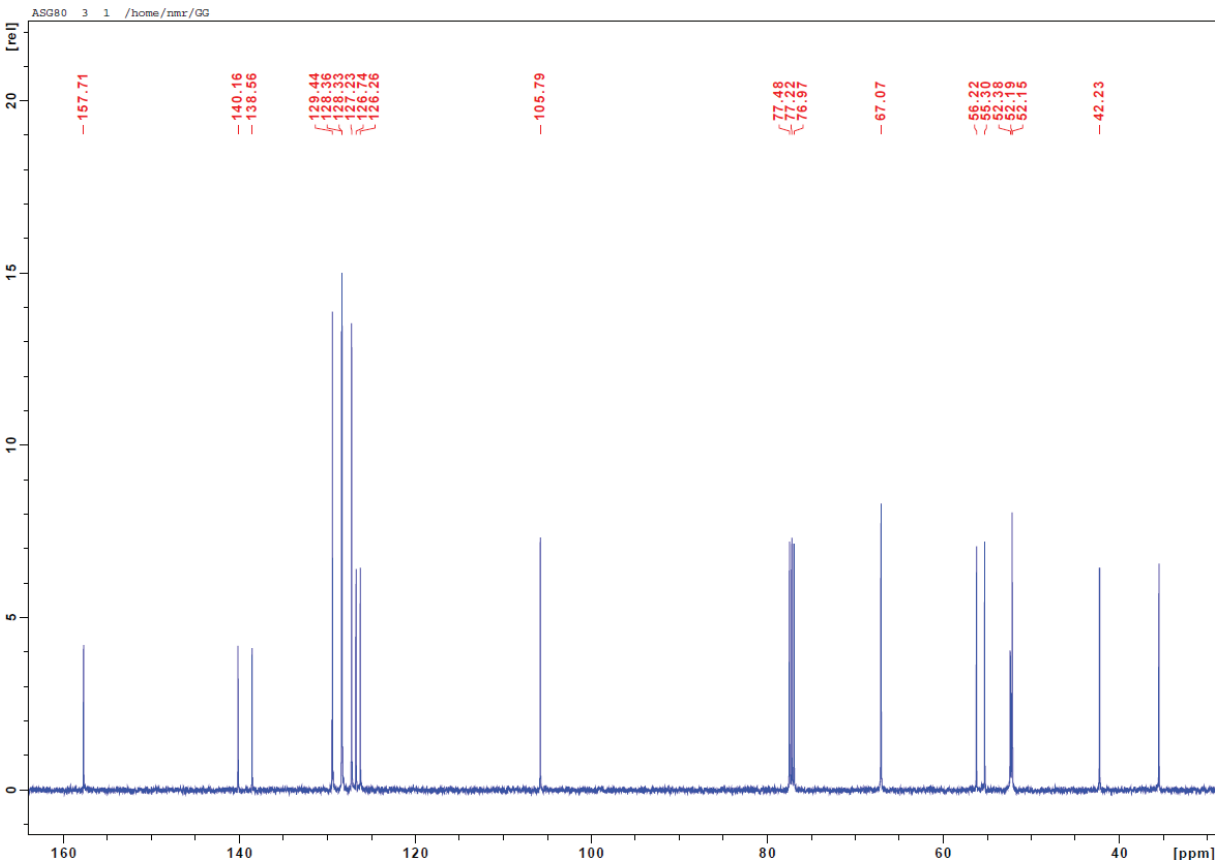

GGa437 4 1 /home/nmr/GG
[rel]
[ppm]

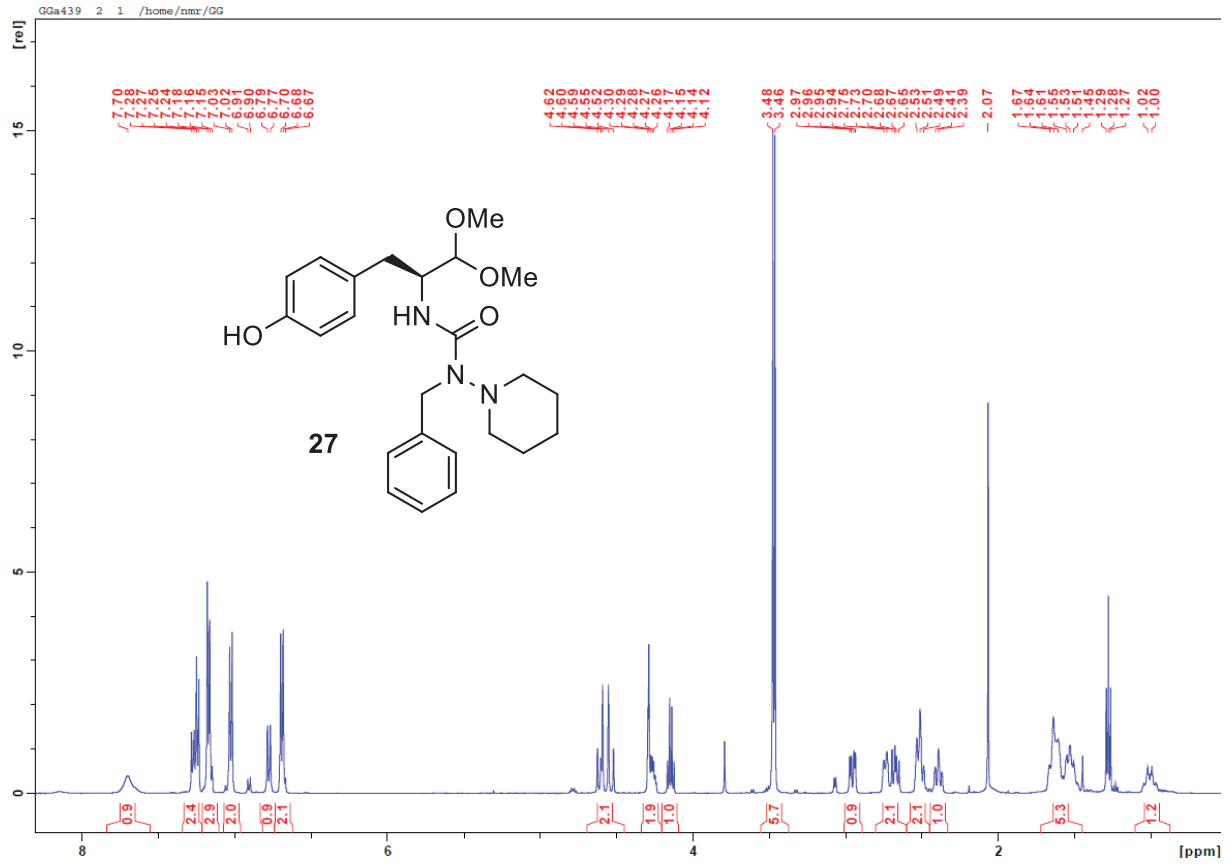

PA09 2 1 /home/nmr/GG
[rel]
[ppm]
OMe
OMe
HN
O
O_2N
N
N
25

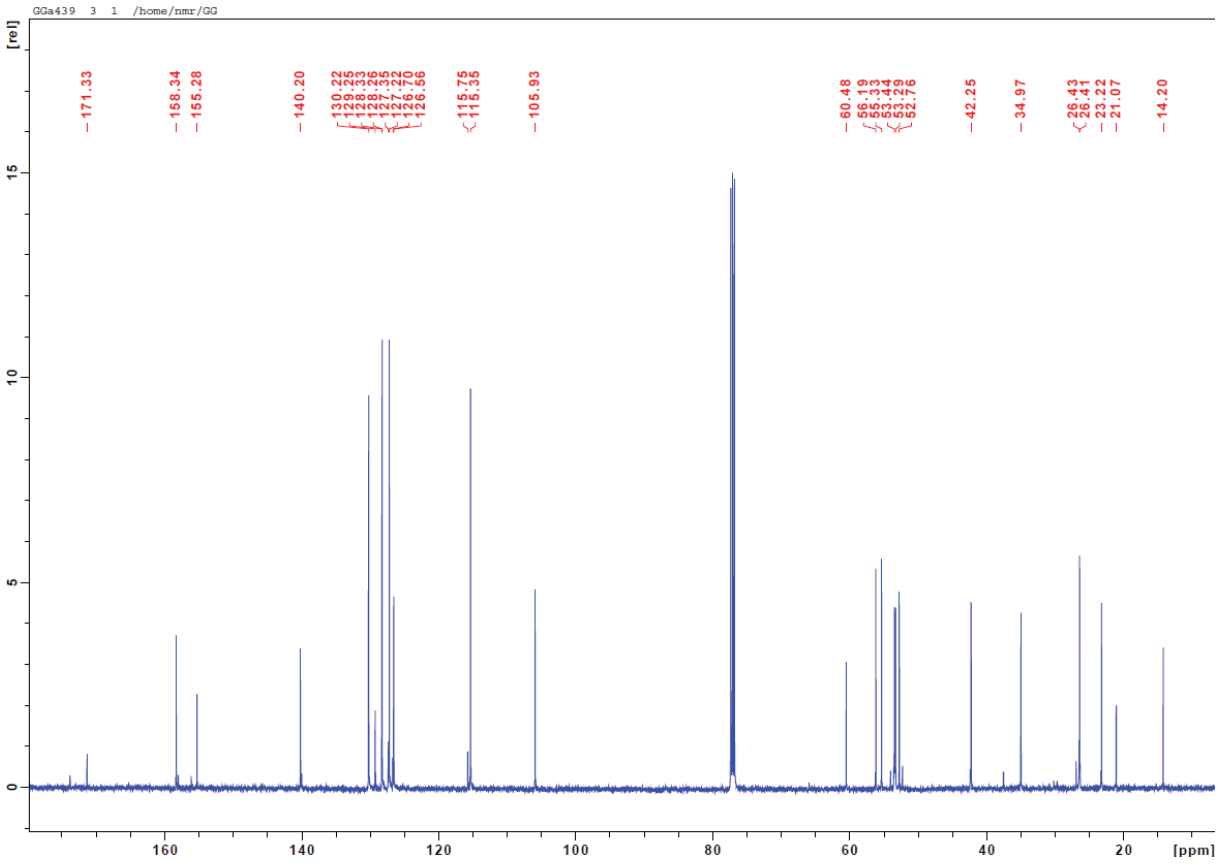

PA09 3 1 /home/nmr/GG
[rel]
[ppm]

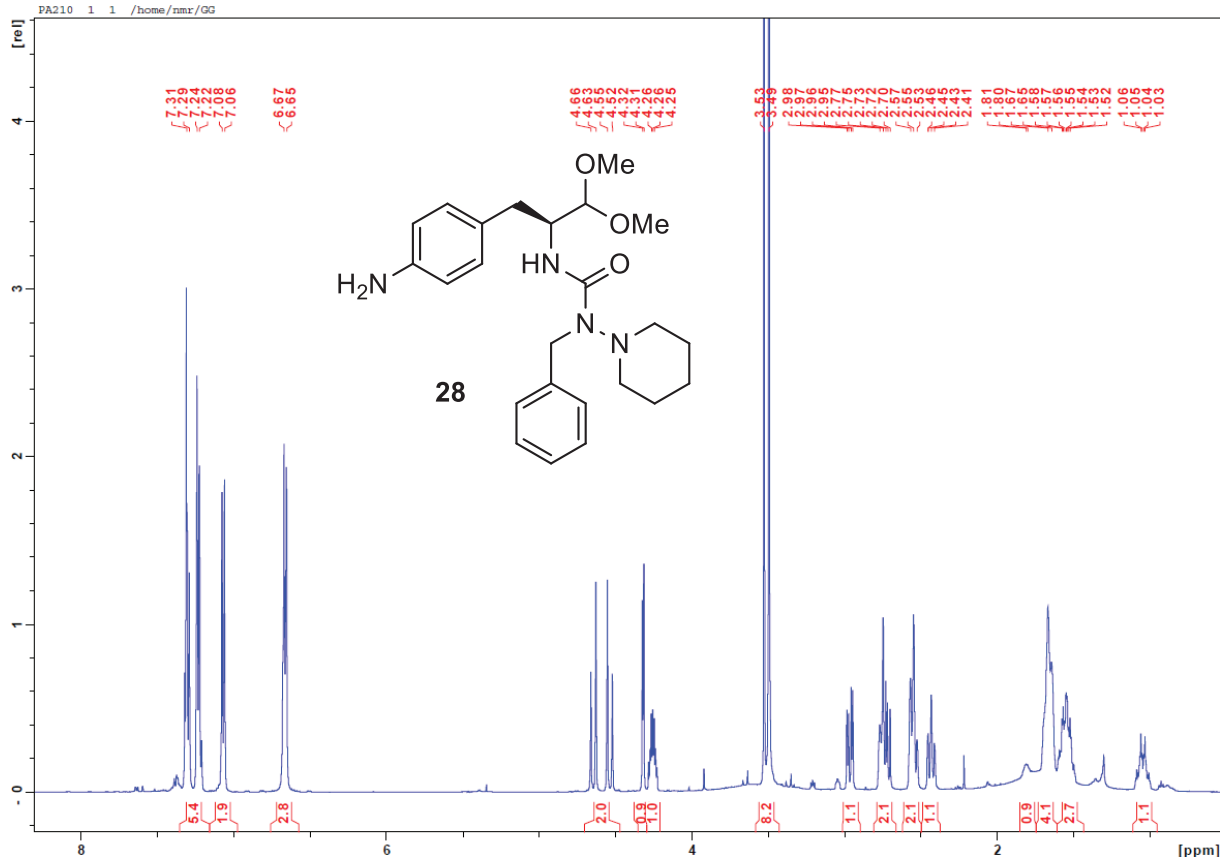
S 28

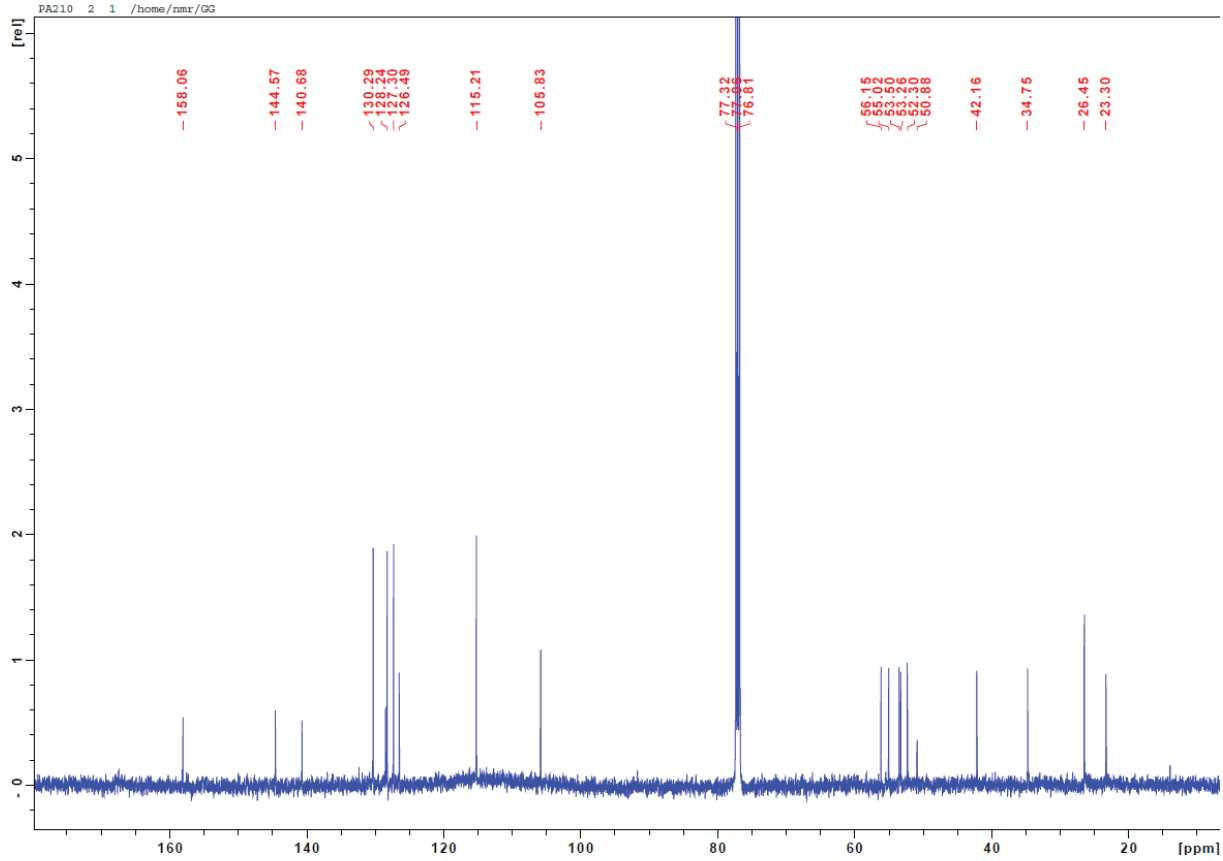

ASG80 2 1 /home/nmr/GG
[rel]
[ppm]
OMe
OMe
HN
O
N
N
O
26

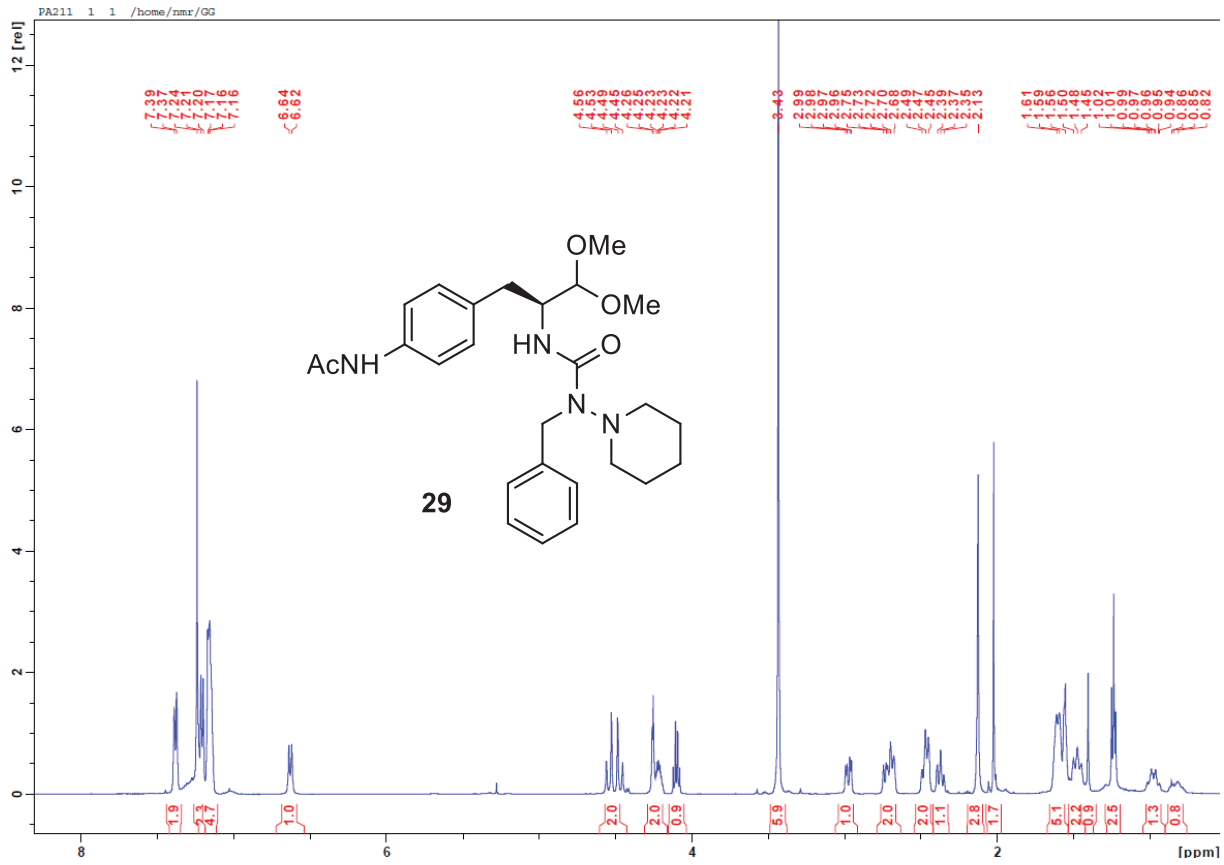

ASG80 3 1 /home/nmr/GG
[rel]
[ppm]

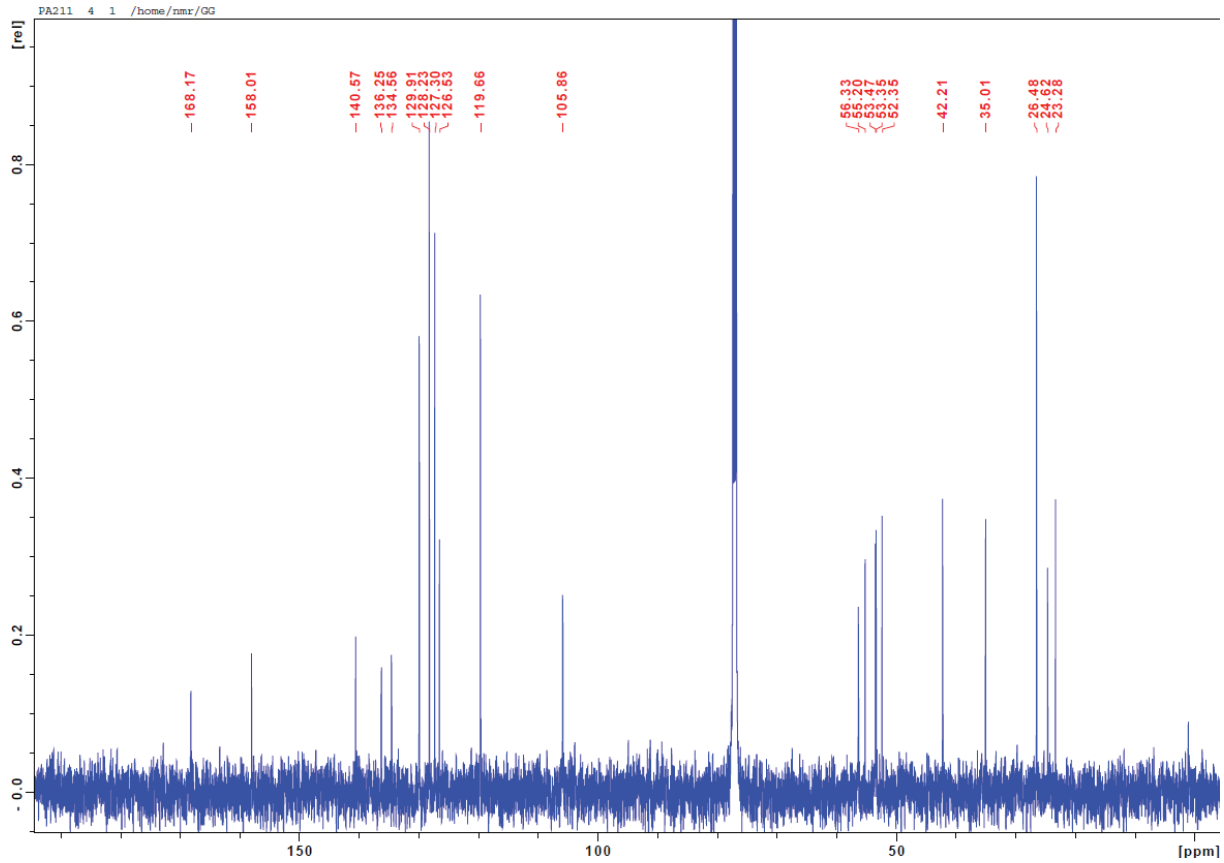

GGa439 2 1 /home/nmr/GG
[rel]
[ppm]
OMe
OMe
HN
O
HO
N
N
27

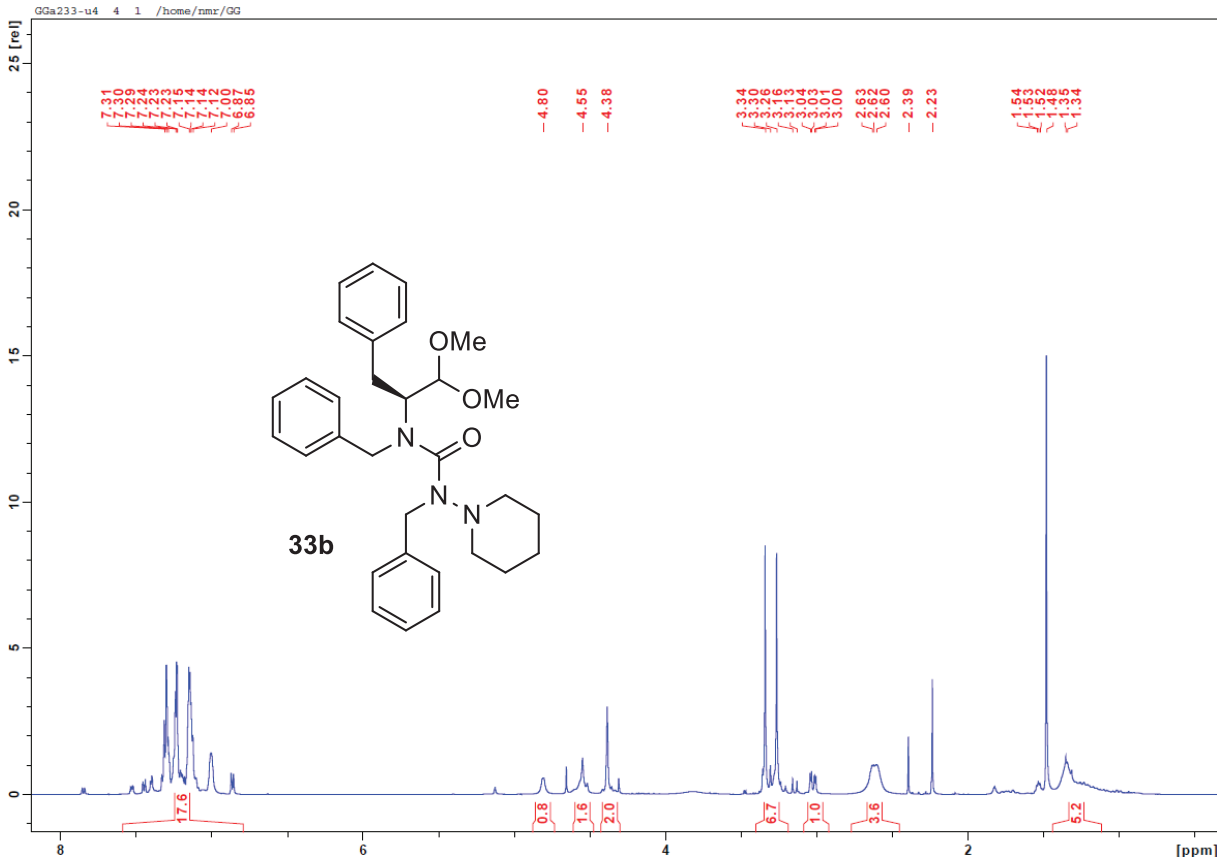

GGa439 3 1 /home/nmr/GG
[rel]
[ppm]

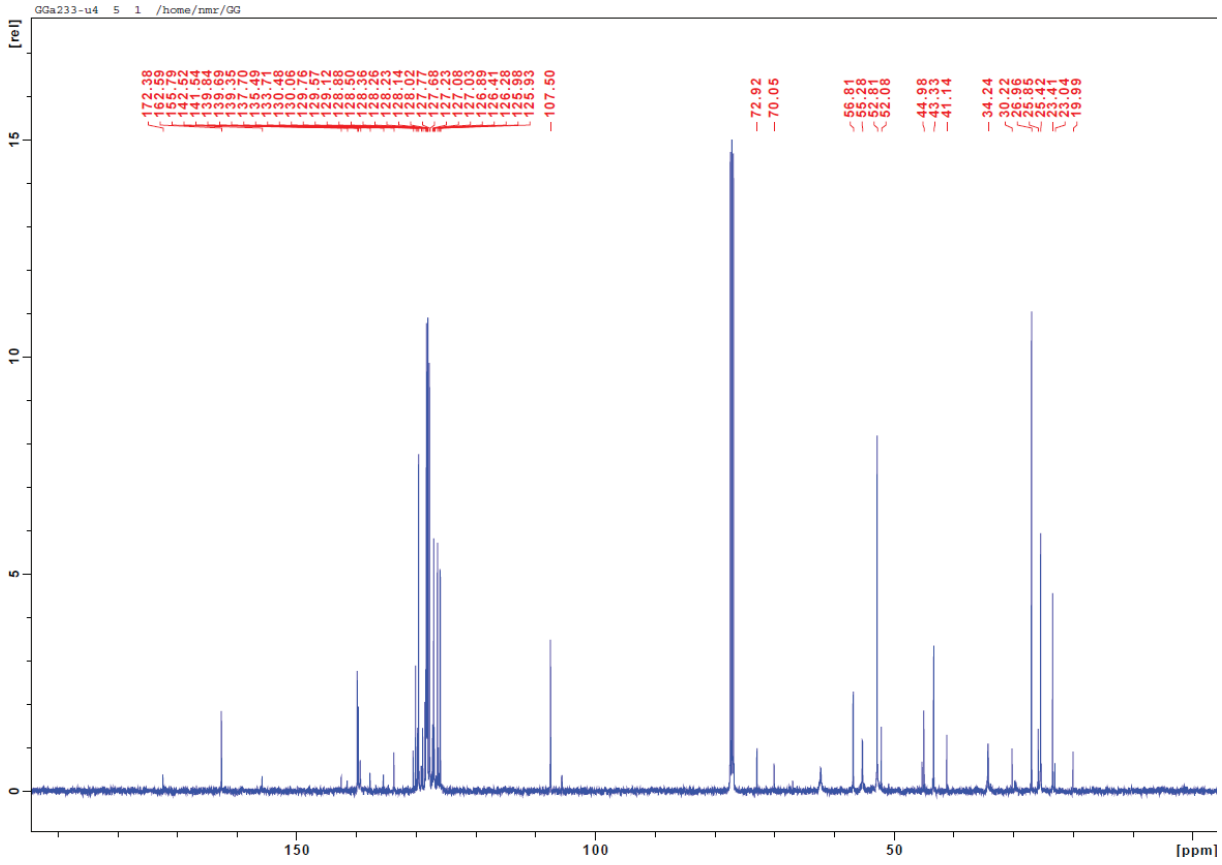

PA210 1 1 /home/nmr/GG
[rel]
[ppm]
OMe
OMe
HN
O
N
N
H_2N
28

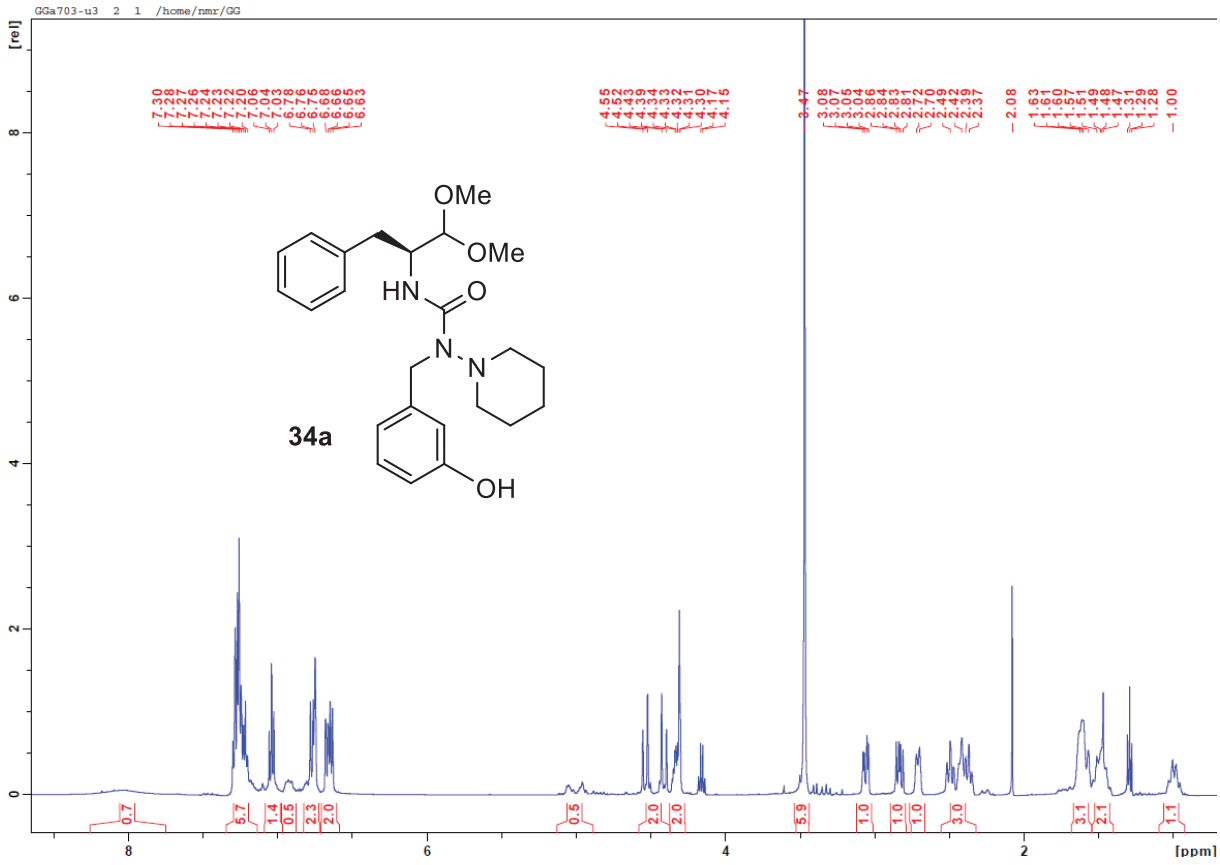

PA210 2 1 /home/nmr/GG
[rel]
[ppm]

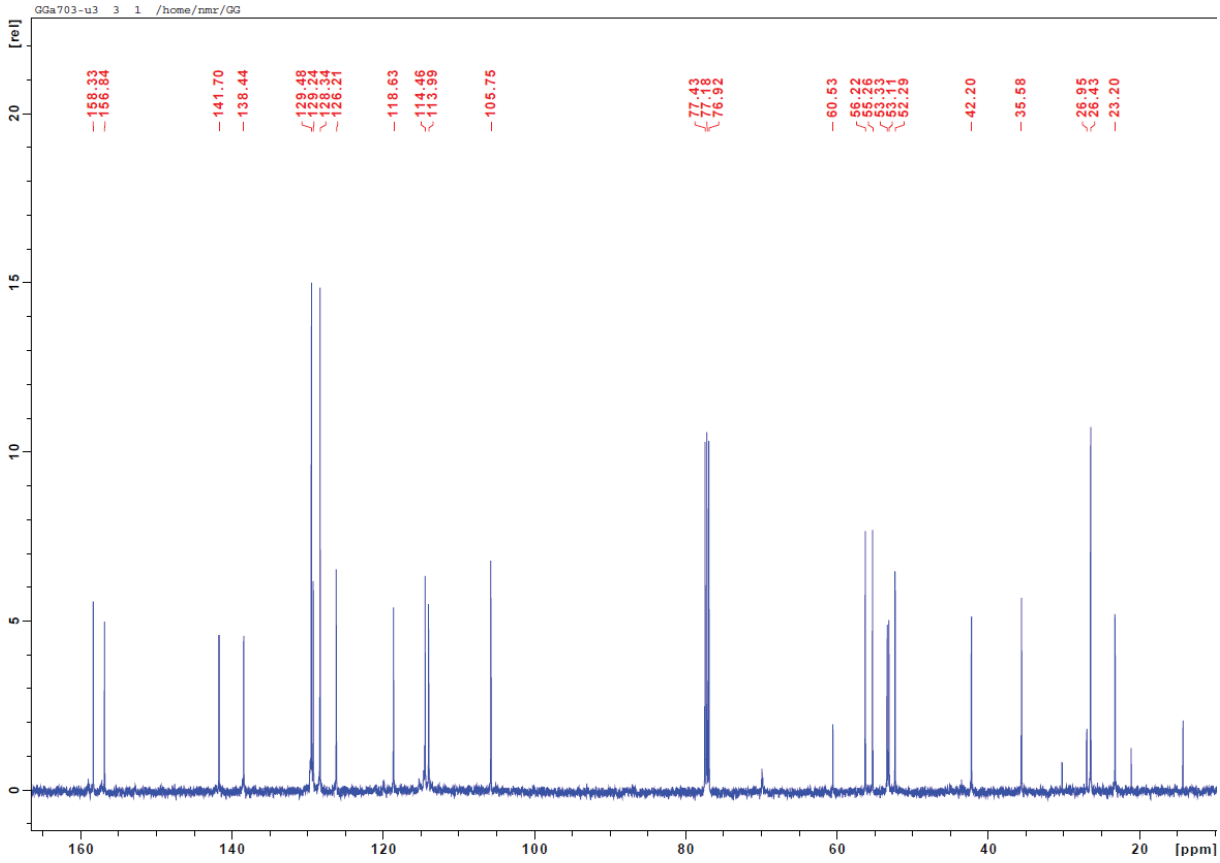

PA211 1 1 /home/nmr/GG
OMe
OMe
HN
O
AcNH
N
N
29
[ppm]

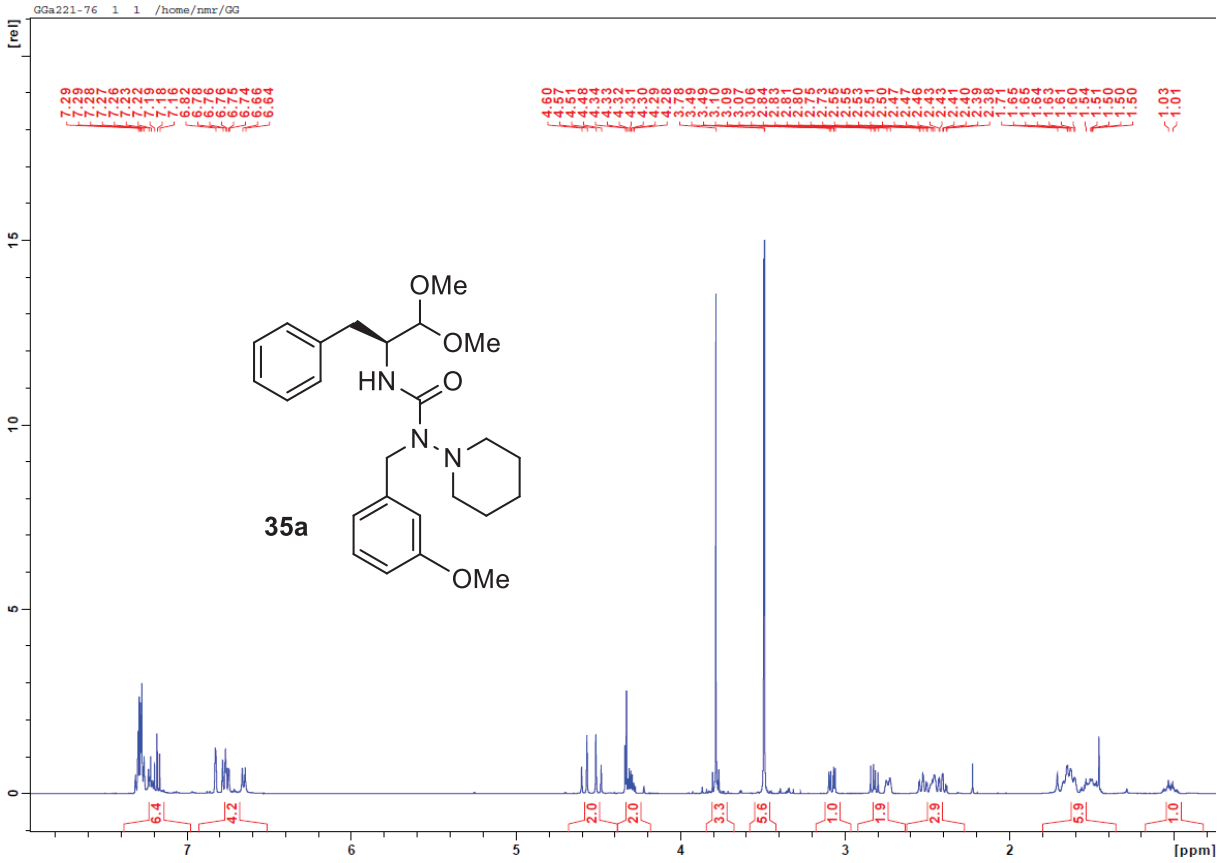

PA211 4 1 /home/nmr/GG
[rel]
168.17
158.01
140.57
136.25
134.56
129.91
128.23
127.30
126.53
119.66
105.86
56.33
55.20
53.47
53.35
52.35
42.21
35.01
26.48
24.62
23.28
[ppm]

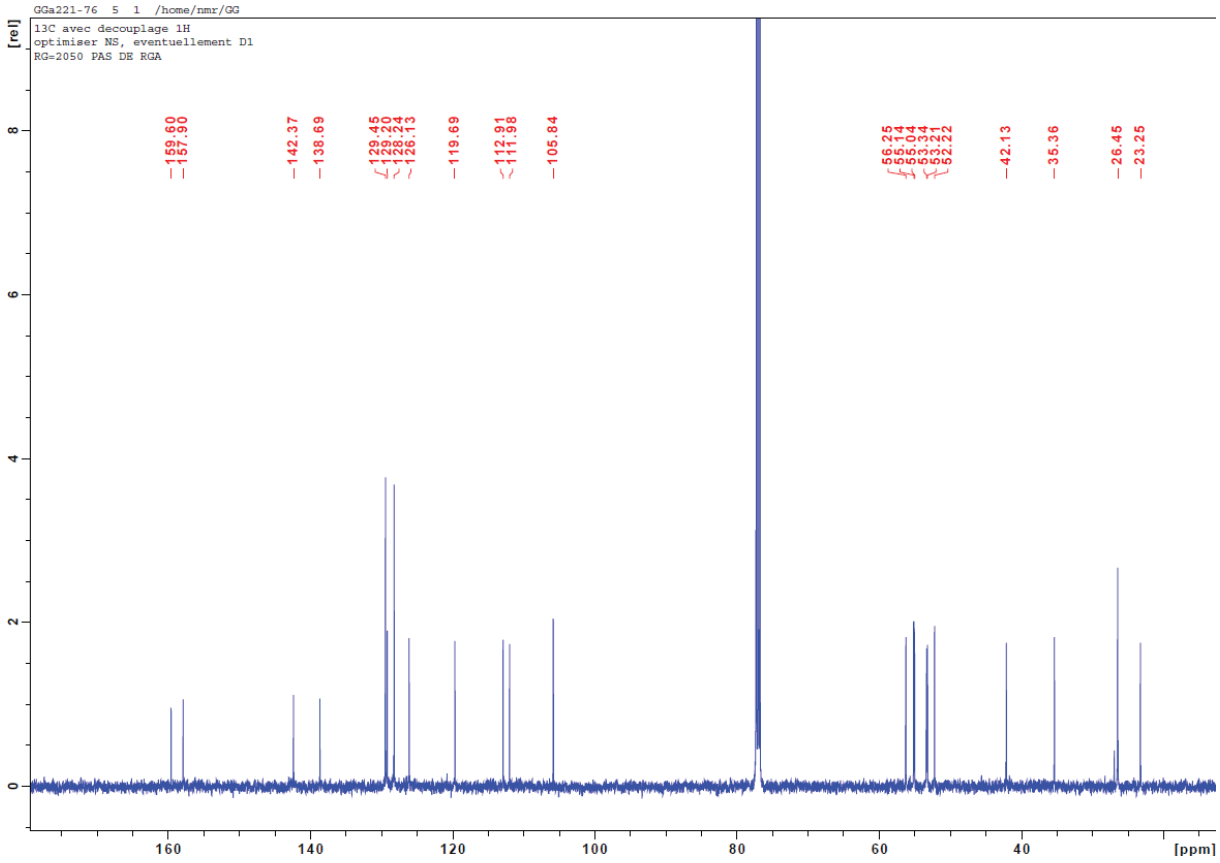

GGa233-u4 4 1 /home/nmr/GG
OMe
OMe
N
O
N
N
33b
[ppm]

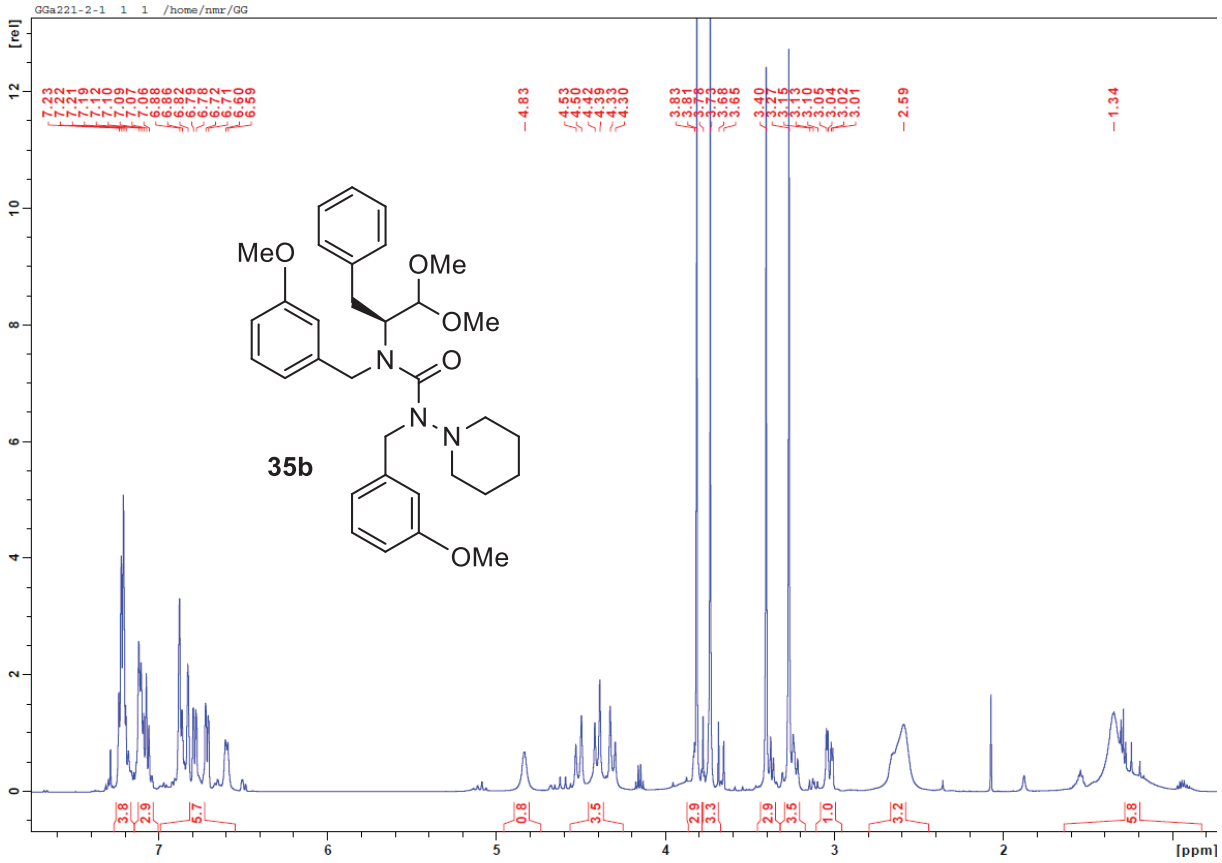

GGa233-u4 5 1 /home/nmr/GG
[rel]
[ppm]

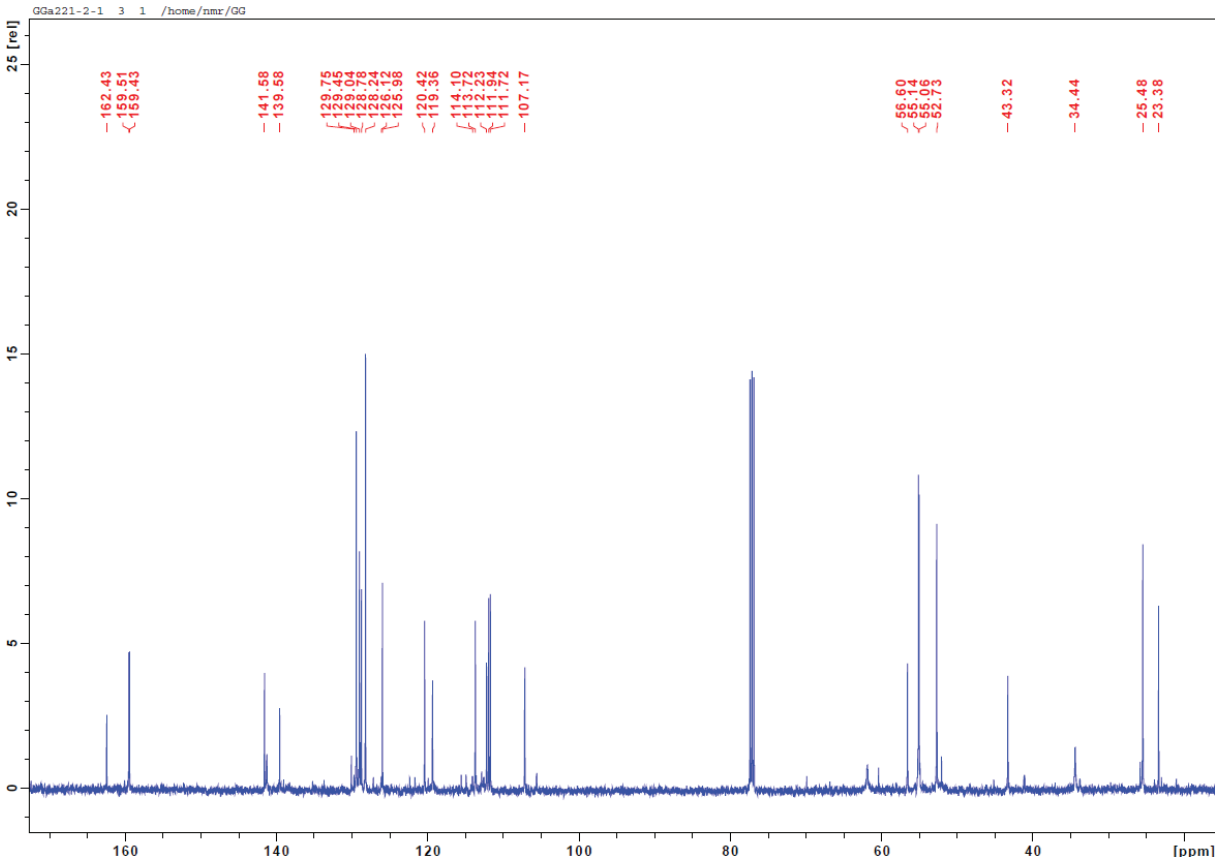

GGa703-u3 2 1 /home/nmr/GG
[rel]
[ppm]
OMe
OMe
HN
O
N
N
OH
34a

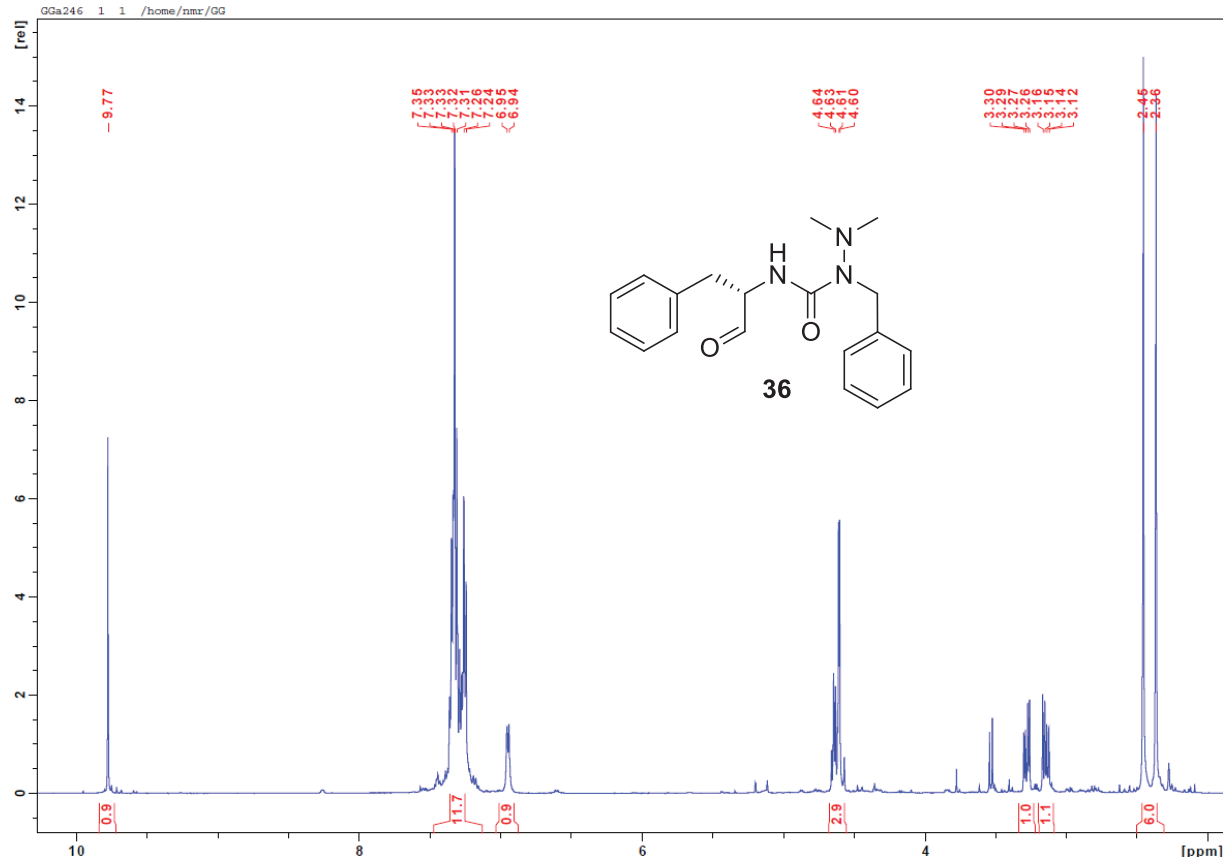

GGa703-u3 3 1 /home/nmr/GG
[rel]
[ppm]

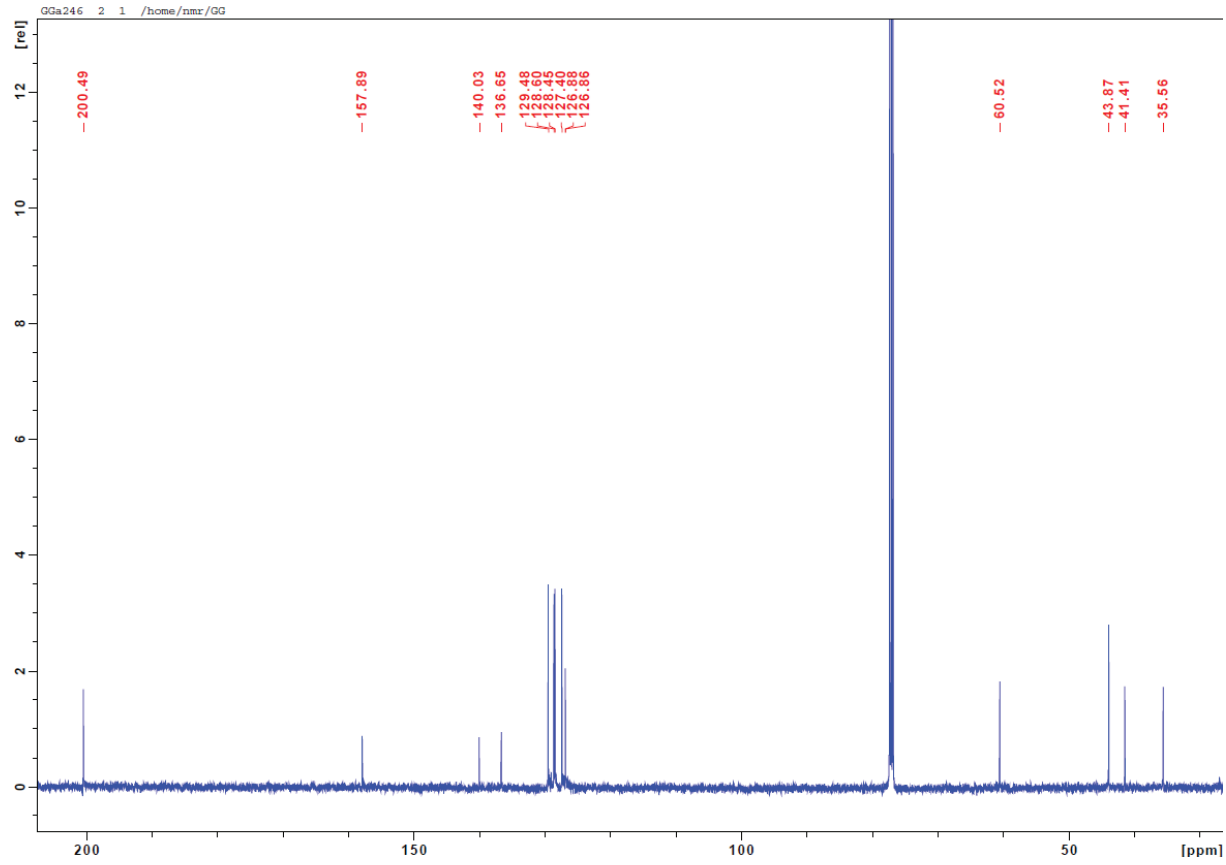

GGa221-76 1 1 /home/nmr/GG
[rel]
[ppm]
OMe
OMe
HN
O
N
N
OMe
35a

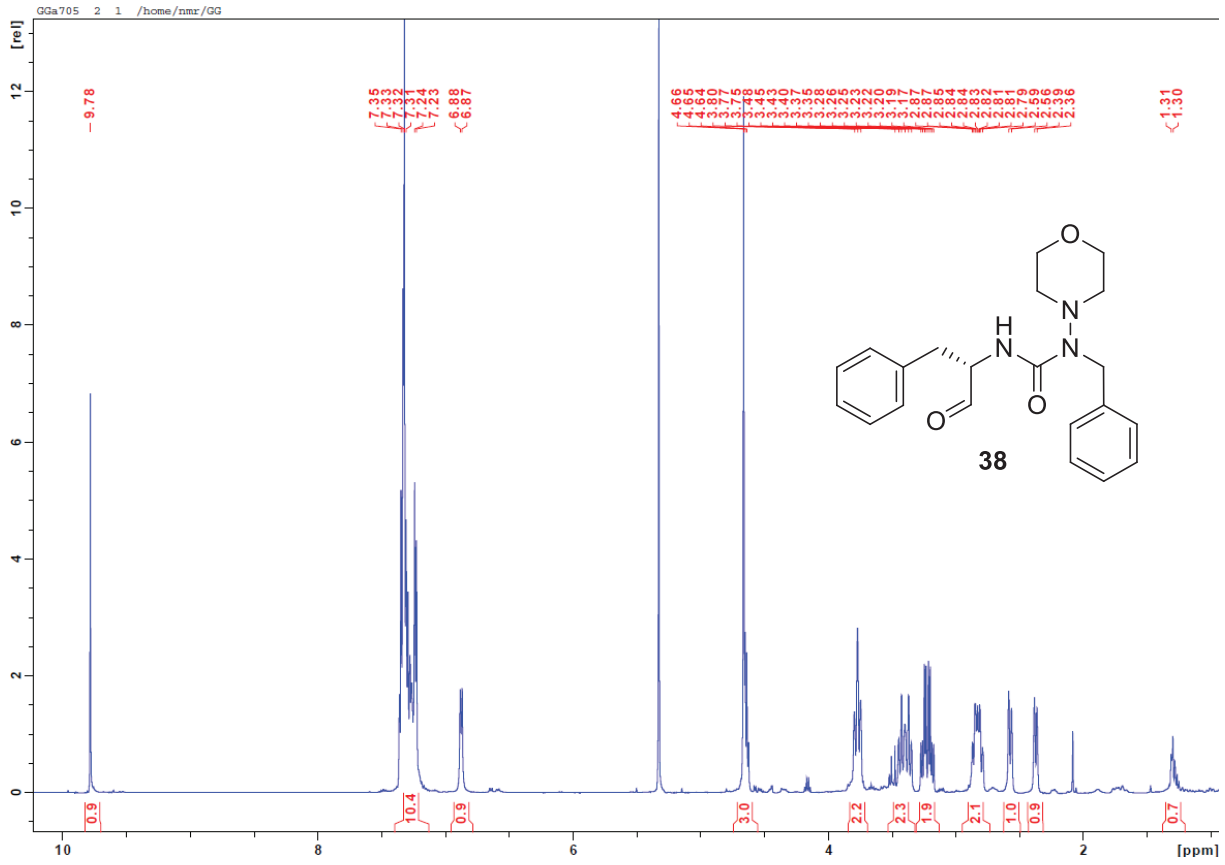

GGa221-76 5 1 /home/nmr/GG
13C avec decouplage 1H
optimiser NS, eventuellement D1
RG=2050 PAS DE RGA
[rel]
[ppm]


GGa221-2-1 1 1 /home/nmr/GG
[rel]
7.23
7.22
7.21
7.19
7.12
7.10
7.09
7.07
7.06
6.88
6.86
6.82
6.78
6.72
6.71
6.60
6.59
4.83
4.53
4.50
4.42
4.39
4.33
4.30
3.83
3.81
3.78
3.78
3.65
3.40
3.27
3.15
3.13
3.06
3.04
3.02
3.01
2.59
1.34
MeO
OMe
OMe
N
O
N
N
35b
OMe
3.8
2.9
5.7
0.8
3.5
2.9
3.3
2.9
3.5
1.0
3.2
5.8
[ppm]

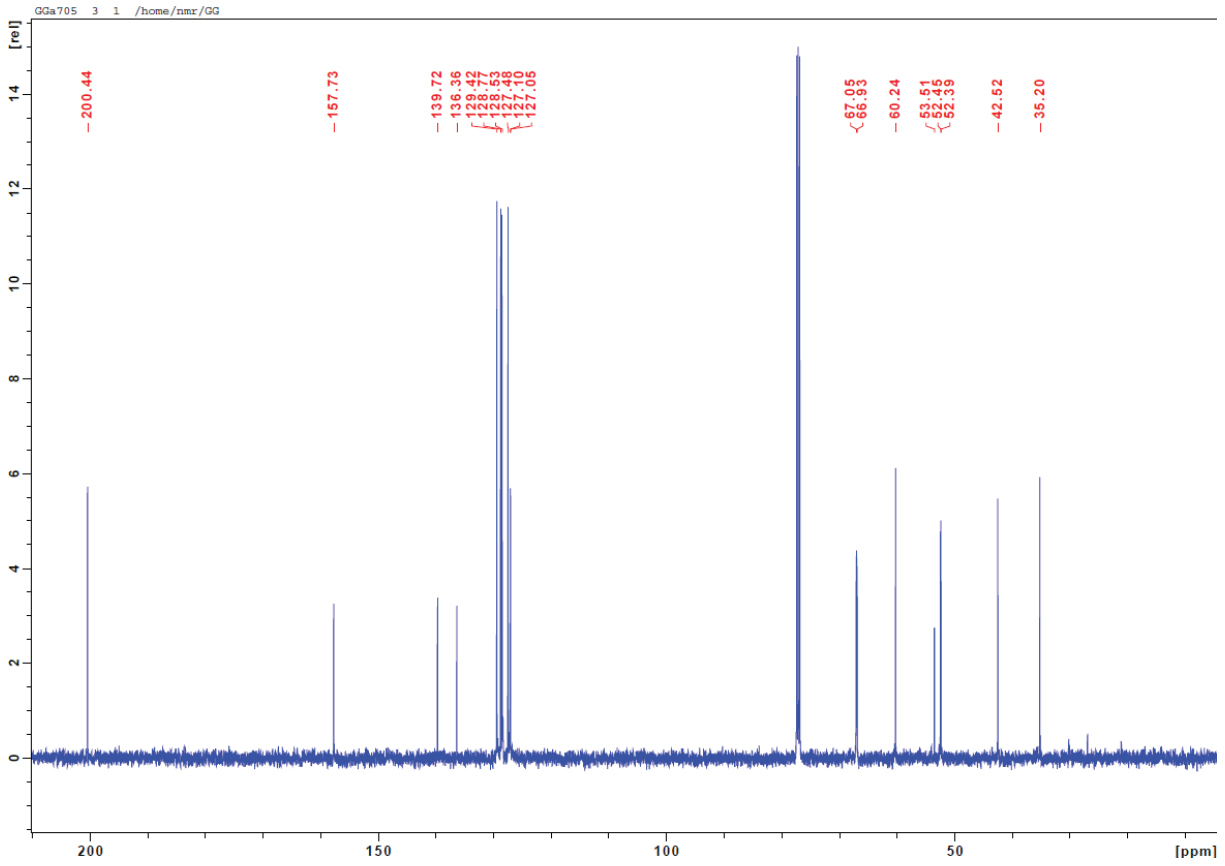

GGa221-2-1 3 1 /home/nmr/GG
[rel]
162.43
159.51
159.43
141.58
139.58
129.75
129.45
129.04
128.78
128.24
126.98
120.42
119.36
114.10
113.72
112.23
111.94
111.72
107.17
56.60
55.14
55.06
52.73
43.32
34.44
25.48
23.38
[ppm]


GGa246 1 1 /home/nmr/GG
[rel]
9.77
7.35
7.33
7.32
7.31
7.26
7.24
6.95
6.94
4.64
4.63
4.61
4.60
3.30
3.27
3.26
3.16
3.15
3.14
3.12
2.45
2.36
0.9
11.7
0.9
2.9
1.0
1.1
6.0
[ppm]

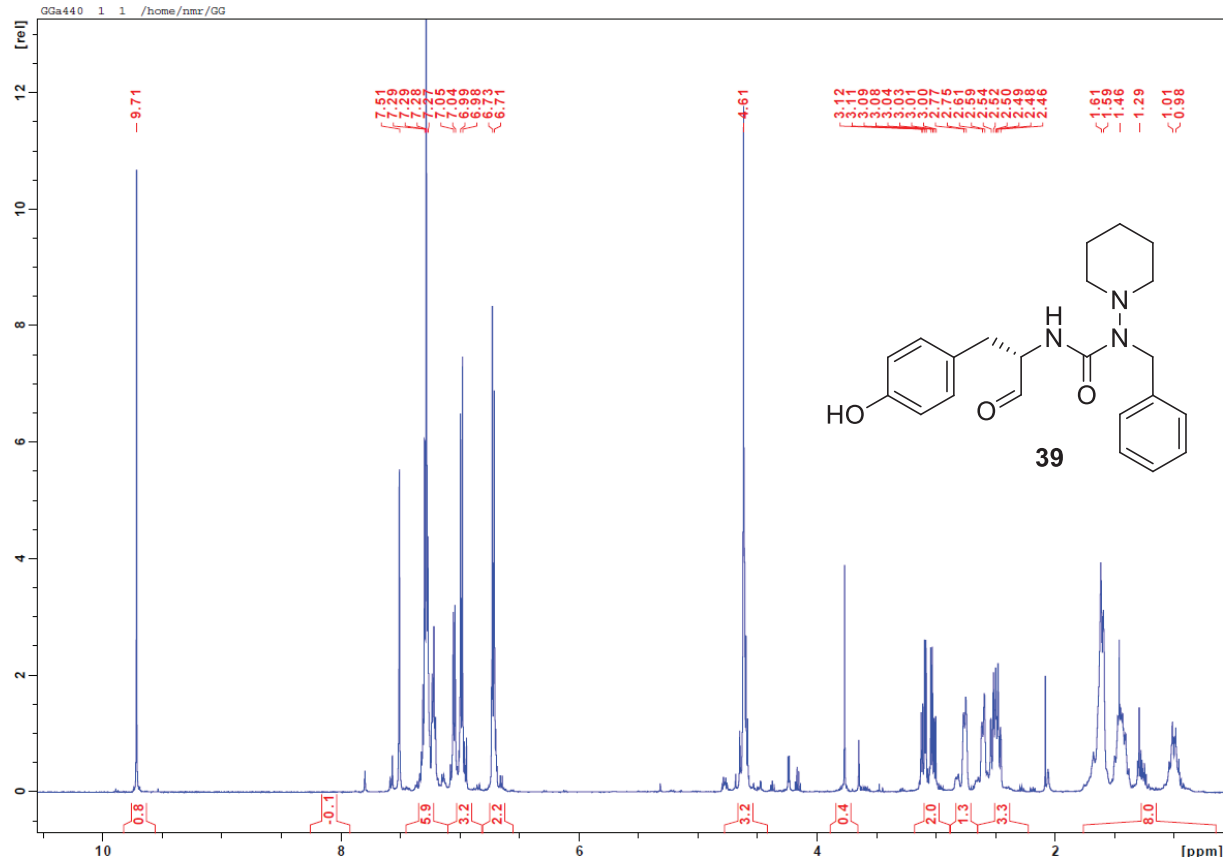

GGa246 2 1 /home/nmr/GG
[rel]
200.49
157.89
140.03
136.65
129.48
128.60
128.45
127.40
126.88
126.86
60.52
43.87
41.41
35.56
[ppm]

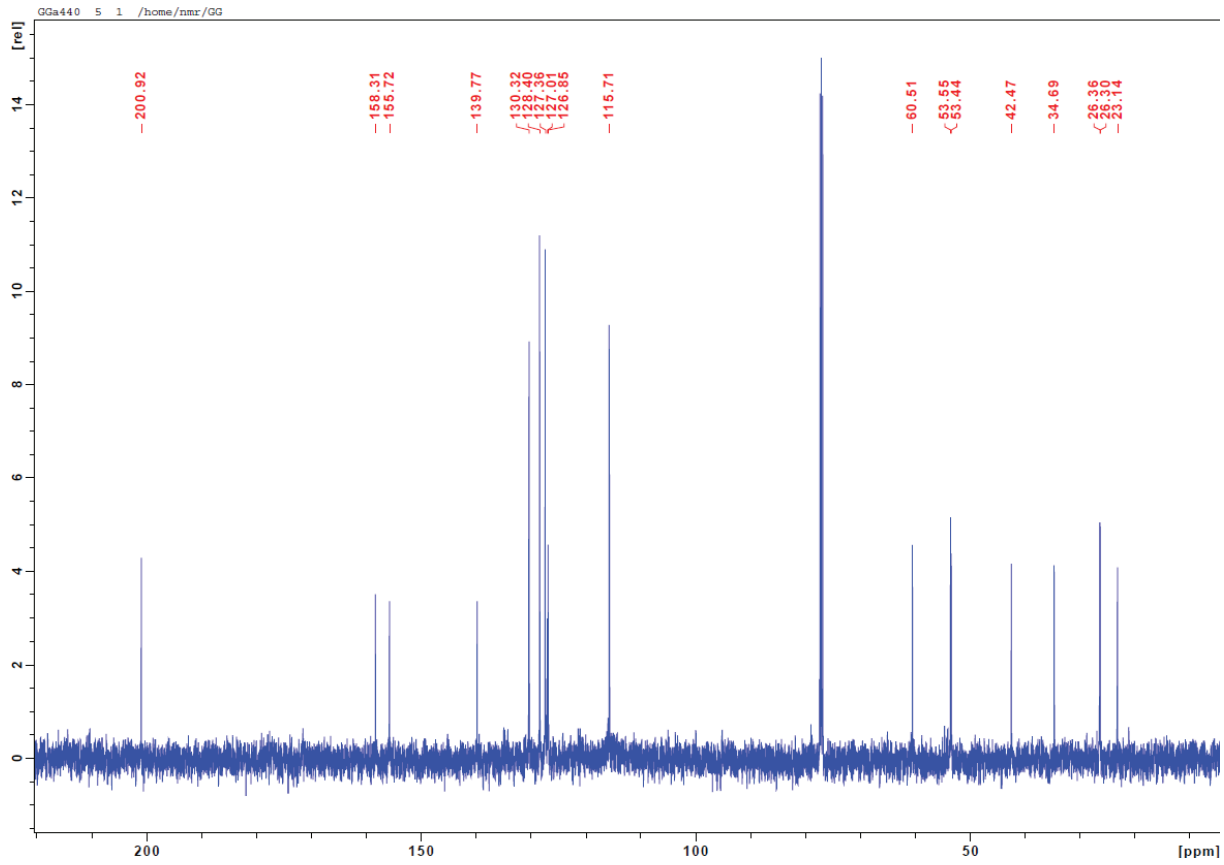

GGa705 2 1 /home/nmr/GG
[rel]
9.78
7.35
7.33
7.32
7.31
7.24
7.23
6.88
6.87
4.66
4.65
4.64
4.63
3.77
3.75
3.48
3.45
3.43
3.40
3.37
3.35
3.28
3.26
3.25
3.23
3.22
3.20
3.19
3.17
2.87
2.85
2.84
2.83
2.82
2.81
2.61
2.59
2.56
2.39
2.36
1.31
1.30
0.9
10.4
0.9
3.0
2.2
2.3
1.9
2.1
1.0
0.9
0.7
[ppm]

38

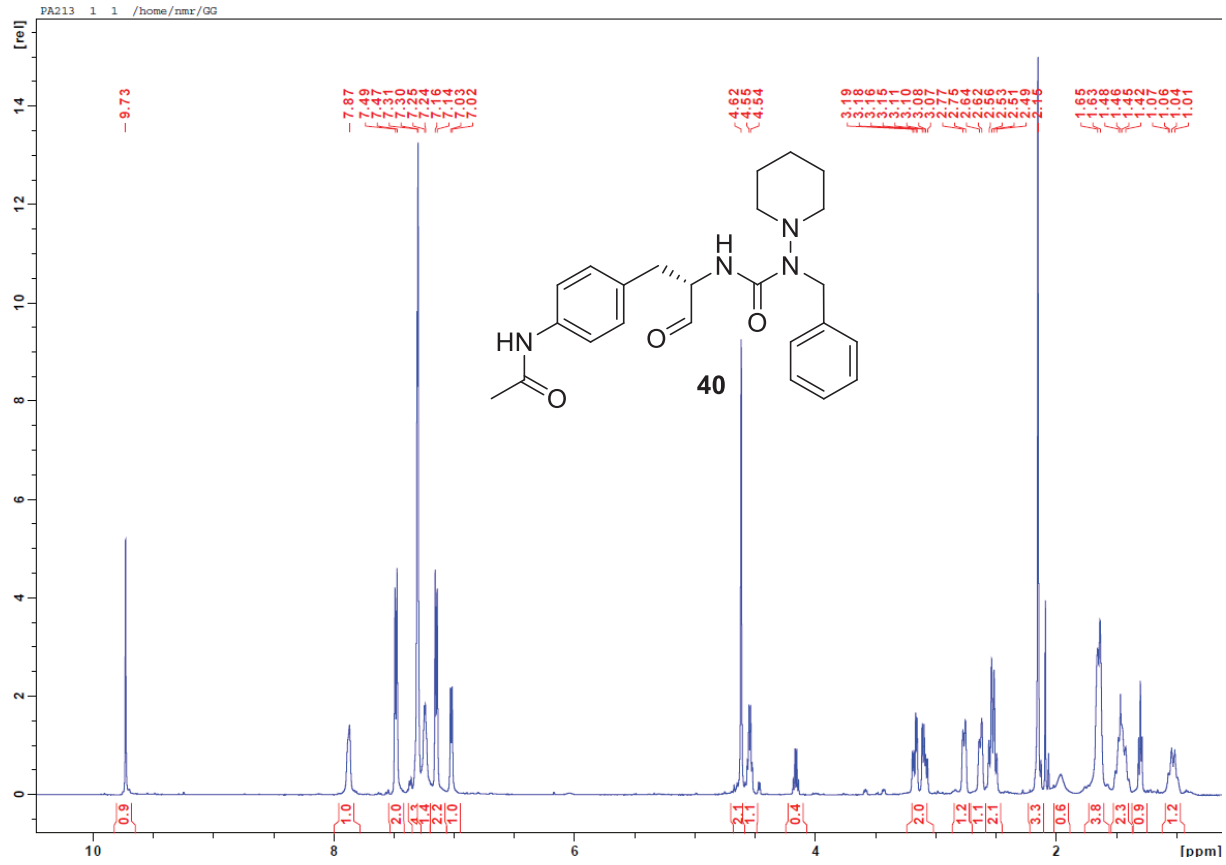

GGa705 3 1 /home/nmr/GG
[rel]
200.44
157.73
139.72
136.36
129.42
128.77
128.53
127.48
127.10
127.06
67.05
66.93
60.24
53.51
52.45
52.39
42.52
35.20
[ppm]

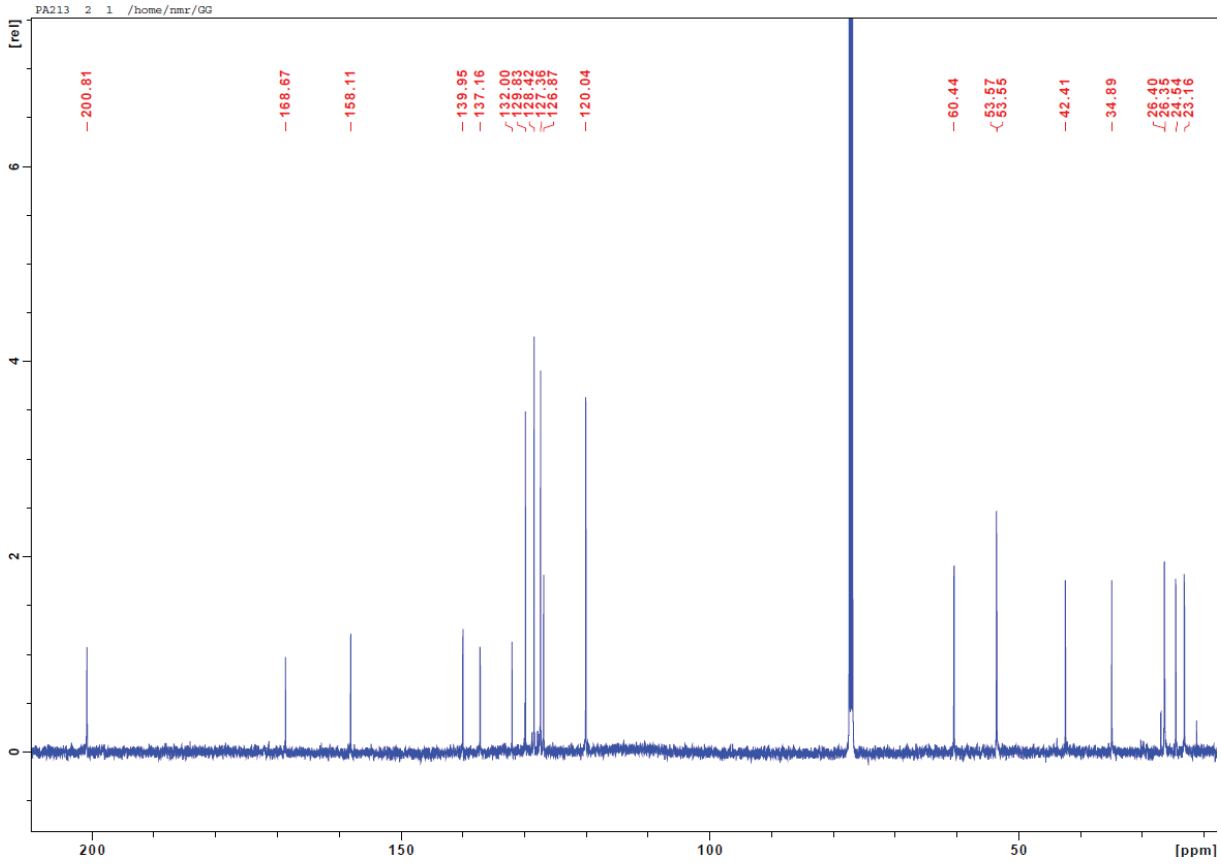
S 38


GGa440 1 1 /home/nmr/GG
[rel]
9.71
7.51
7.29
7.29
7.28
7.25
7.04
6.99
6.98
6.73
6.71
4.61
3.12
3.11
3.09
3.08
3.04
3.03
3.01
3.00
2.77
2.75
2.61
2.59
2.54
2.52
2.50
2.49
2.48
2.46
1.61
1.59
1.46
1.29
1.01
0.98
0.8
0.1
5.9
3.2
2.2
3.2
0.4
2.0
1.3
3.3
8.0
[ppm]
HO
O
H
N
N
N
O
39

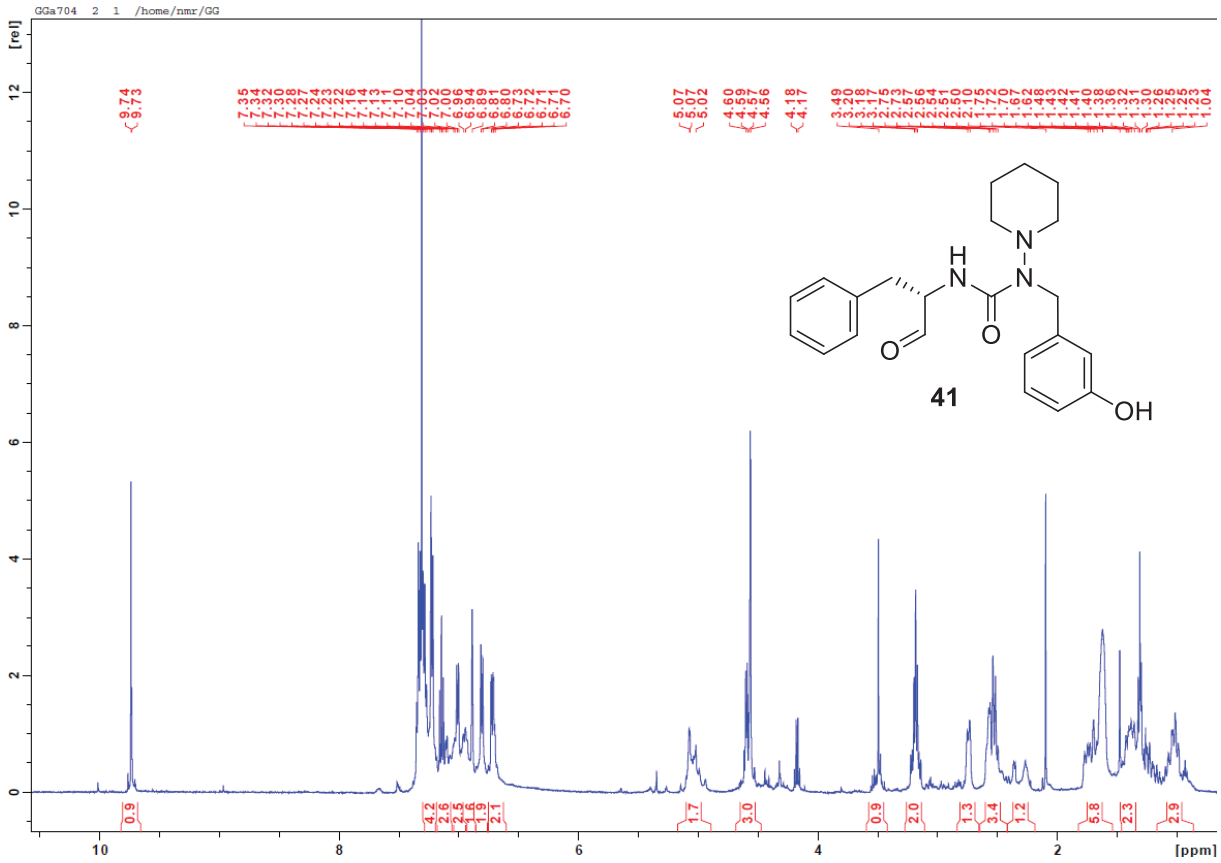

GGa440 5 1 /home/nmr/GG
[rel]
200.92
158.31
155.72
139.77
130.32
129.40
127.36
127.01
126.85
115.71
60.51
53.55
53.44
42.47
34.69
26.36
26.30
23.14
[ppm]

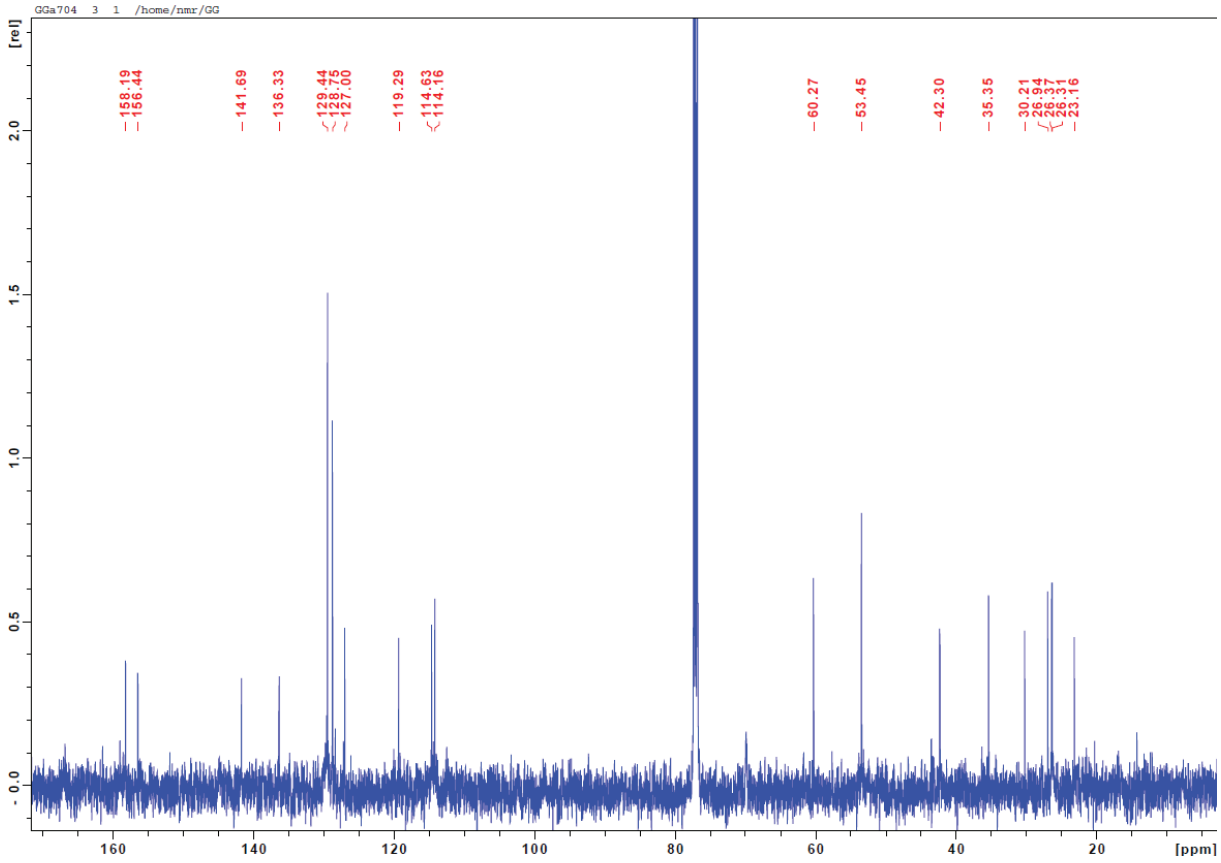
S 39

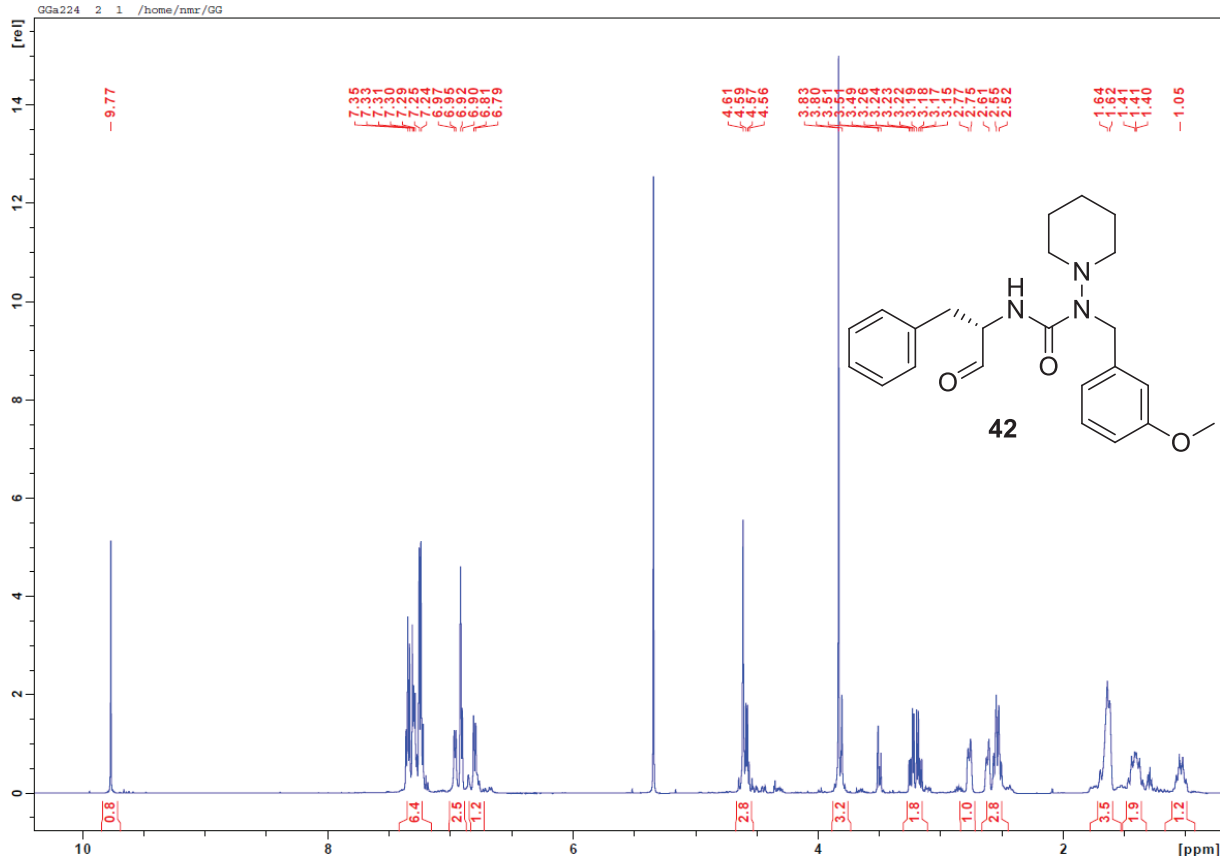


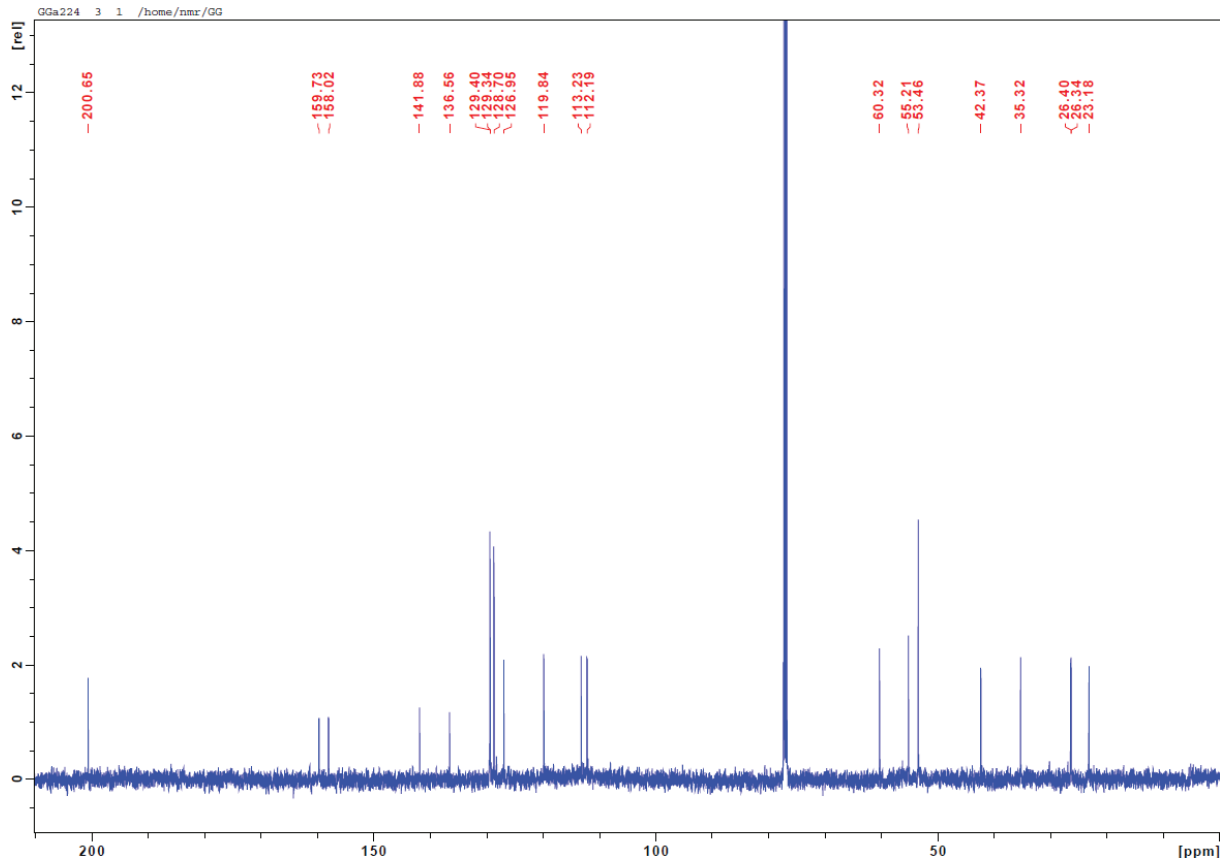


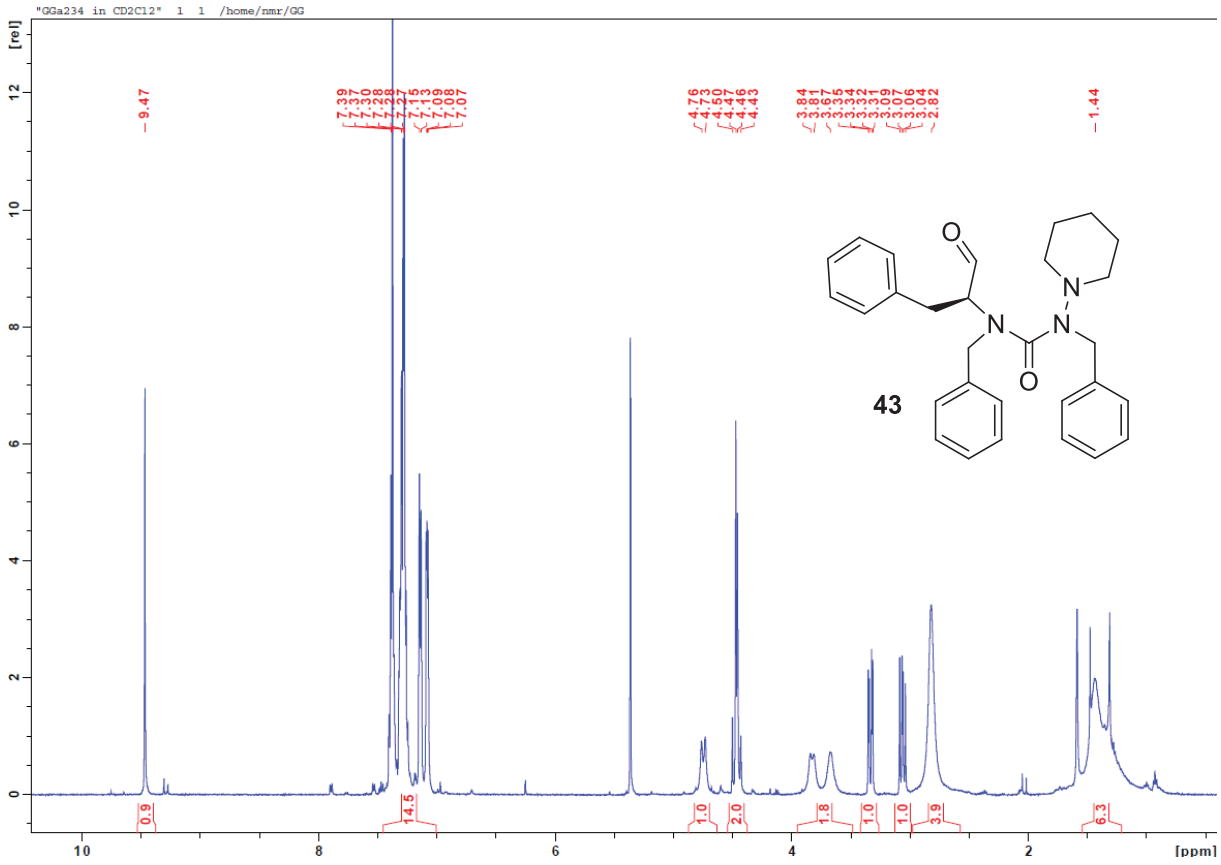
S 40

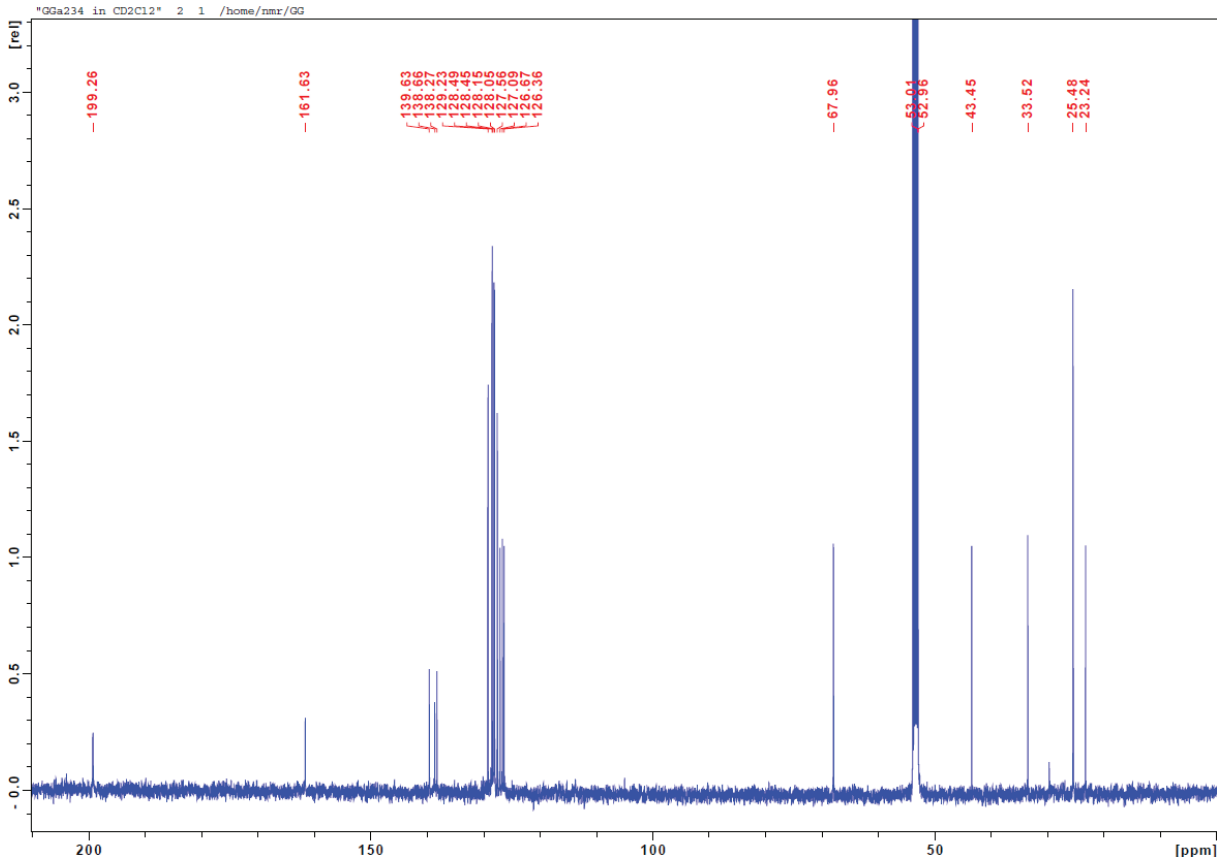

GGa704 2 1 /home/nmr/GG
[rel]
[ppm]
41
OH


GGa704 3 1 /home/nmr/GG
[rel]
[ppm]


GGa224 2 1 /home/nmr/GG
[rel]
[ppm]
42


GGa224 3 1 /home/nmr/GG
[rel]
[ppm]


"GGa234 in CD2Cl2" 1 1 /home/nmr/GG
[rel]
9.47
7.39
7.37
7.30
7.28
7.27
7.15
7.13
7.09
7.08
7.07
4.76
4.73
4.50
4.47
4.46
4.43
3.84
3.81
3.67
3.35
3.34
3.32
3.31
3.09
3.07
3.06
3.04
2.82
1.44
0.9
14.5
1.0
2.0
1.8
1.0
1.0
3.9
6.3
[ppm]
43


"GGa234 in CD2Cl2" 2 1 /home/nmr/GG
[rel]
199.26
161.63
139.63
138.66
138.07
129.23
128.49
128.45
128.15
128.05
127.56
127.09
126.67
126.36
67.96
53.01
52.96
43.45
33.52
25.48
23.24
[ppm]

S 43

Conclusion

Developing multifunctional drugs with dual-function behavior may be an interesting solution to overcome efflux problem because they have two actions in an only molecule, avoiding drug-drug interactions. First dual- or multi-function agents described were inhibitors which simultaneously blocked several transporters, such as cyclosporin A, elacridar and tariquidar. However, modulation of multiple transporters continues to be debated within the scientific community. Since the discovery that P-gp can be reversed by verapamil in 1981, scientists has spent a considerable time to find a nontoxic, highly selective inhibitor that has not been successful until now. Knowing these efflux pumps are flexible with a considerable substrate binding cavity, it does not seem so logical to design selective compounds for a “promising” target.

Here we attempted to establish other dual agents: inhibitors of both HIV-1 protease and P-gp, non-substrate of P-gp. In the first approach, our aim was to establish a reliable *in silico* protocol to screen the laboratory library in order to find the molecules. One of the major problems identified was the reliability of the docking program (Sybyl molecular-modeling) whose results were hardly reproducible. With the same parameters, the score list varied in a significant manner. In that context, AuPosSOMS was used to clarify results. Two molecules (CT1347 and CT1353) were predicted in different groups according to the analytical method (score or AuPosSOMS). *In cellulo* and FRET assays showed that neither of them have the capacity to inhibit the protease. Further research could integrate the use of other docking programs, such as AutoDock-vina, to compare it prediction with Score. In this manner, new molecules with high possibly of activity could be synthesized.

In the second approach, only three molecules were found to inhibit the HIV protease without being transported by P-gp or ABCG2. As developing molecules which does not interact with these pumps is a substantial work, these results are of great interest. Combined with X-ray crystallography of complexes of these compounds with HIV protease or P-gp, these results may lead to the identification of new molecules with increased inhibition activity of the protease of HIV.

CHAPTER II: STRUCTURAL STUDY OF P-GP

Context

The resolution of P-gp structure, either in several states of the transport cycle, or in complex with a substrate, is essential to understand its inner workings. Therefore, we decided to crystallize this protein in collaboration with Professor Geoffrey Chang (UCSD, San Diego California, USA).

Part I: P-gp in complex

In order to crystallize P-gp, several substrates (daunorubicin, hoechst33342, mitoxantrone and rhodamine123), inhibitors of our laboratory (CT1347, CT1357, CT1336, CT1364…) and HIV-protease inhibitors (dauronavir, amprenavir, indinavir, and the new PIs of the publication I: compounds numbered 36 to 43) were used. The expression, purification and crystallization condition were applied as described in Materials and Methods section.

Between all the conditions tested for each compound (0.1 M Hepes pH 7–8, 50 mM lithium sulfate, 10 mM EDTA, and 24–29.5 % w/v PEG 600 Da at 4 °C) the best one is show in Figure 58. The broad majority of the crystals were obtained at pH: 7–8 and PEG 600: 25.5–28 %.

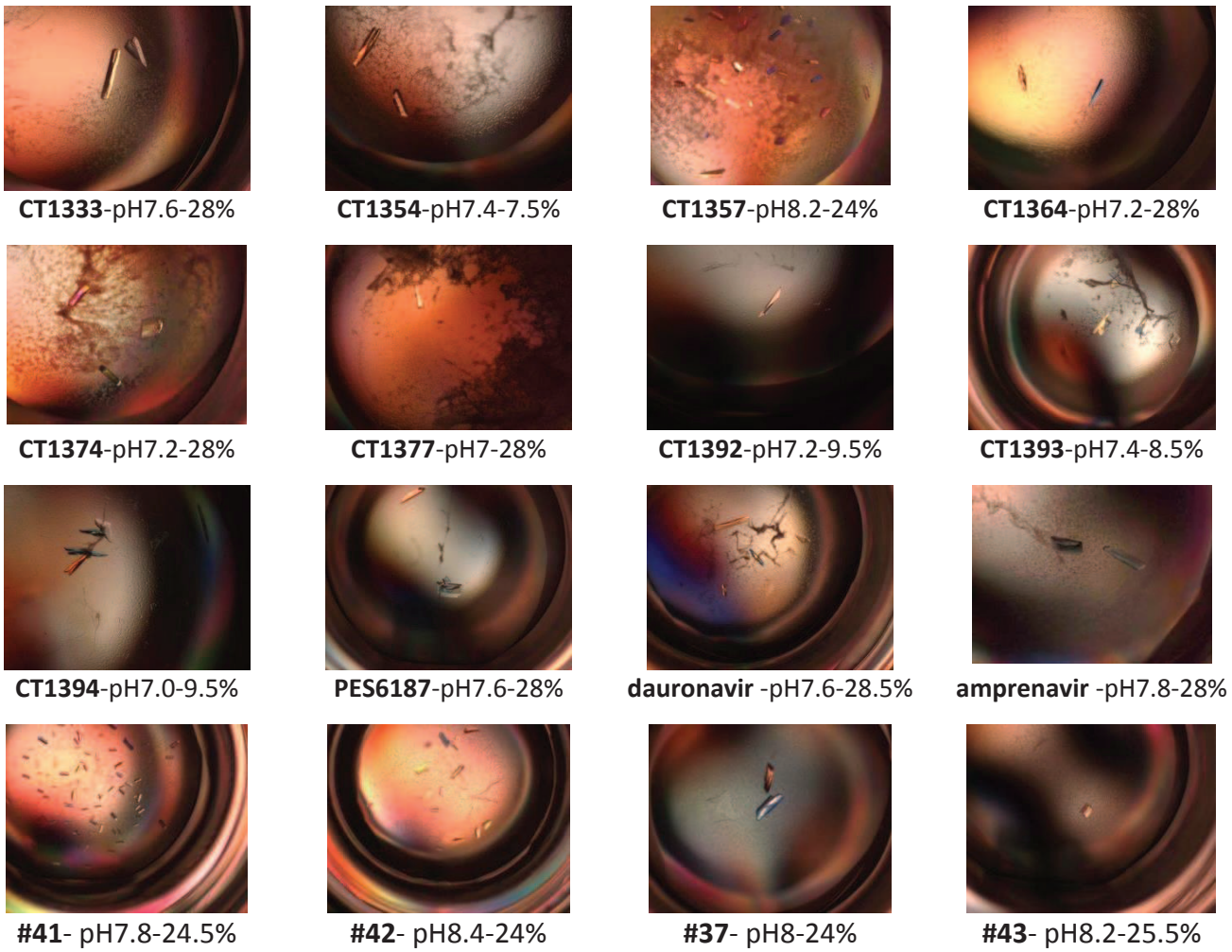
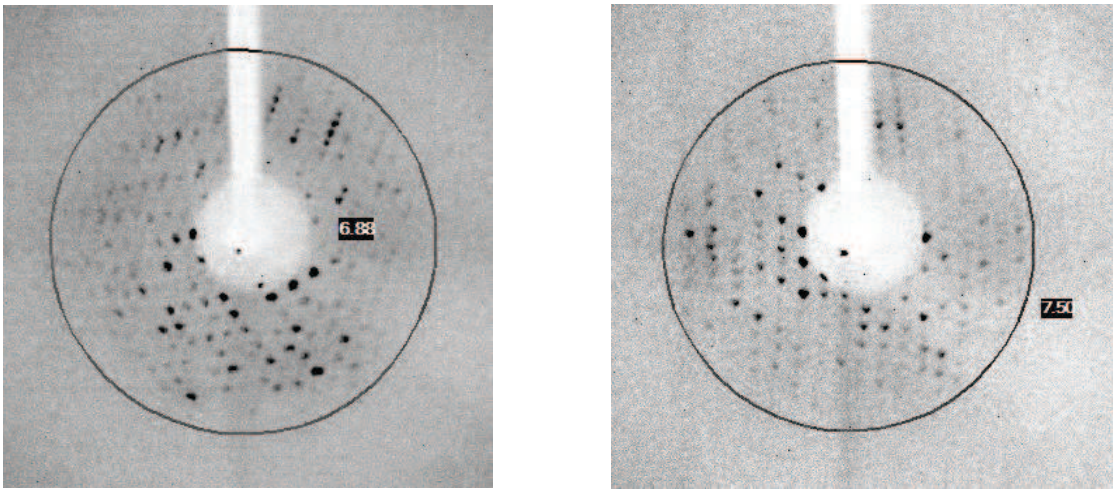


Figure 58. Crystals obtained.
Name of the compound and best crystallization condition (pH and % of PEG 600 Da).

Despite successful crystal formation with different morphologies (Figure 58), the resolution (> 6 Å, Figure 59) of these trials was never high enough to localize the compounds. A resolution between 3 and 6 Å can lead to mistakes in interpreting the electron density map. From 3 Å resolution, the shape of the polypeptide backbone can be traced, small molecules such as ADP or ATP are visible. However, the individual atoms are not distinguished and the side chains are hardly identified.



CT1394-pH 7, 29.5 % PEG 600 **CT1374**-pH 7.0, 29.5 % PEG 600

Figure 59. Pattern of X-ray diffraction of two crystals.
Obtained at D8 VENTURE crystal X-ray diffraction -Bruker

Part II: P-gp in the outward-facing conformation

A P-gp construction was designed as P-gp-Linkerless with a deletion of 40 amino acid between NBD1 and TM7 (Dr. Ina Urbatsch, TTUHSC, Texas, USA), as illustrated in red in Figure 60A. After purification (as above), the protein migrated on SDS PAGE (Figure 60B) and coomassie stained shows the presence of two bands (instead of one at 144 kDa) at 58 and 55 kDa. While the anti-histidine western blot shows only one band at 144 kDa for the WT P-gp and at 55 kDa for the P-gp-Linkerless (Figure 60C). Two halves of P-gp are thus expressed and only one half had the His tag. This construction was hypothesized to improve the resolution because the linker region in P-gp structures is highly flexible. This construction was also used to trap P-gp in the outward-facing conformation, because without the linker NBD's could interact more easily and the protein may be more organized and give a better resolution.

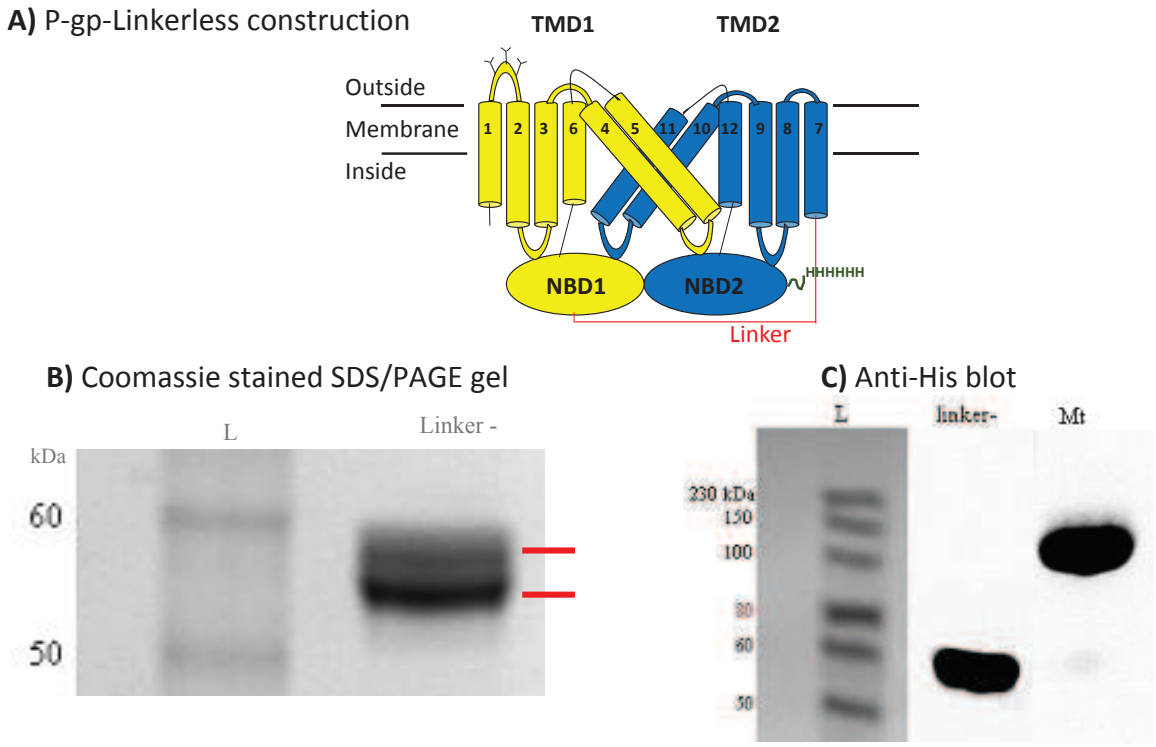


Figure 60. P-gp-Linkerless.
L: Protein molecular weight marker in kDa; Linker -: P-gp-Linkerless membrane fraction; Mt: P-gp wild-type membrane fraction.

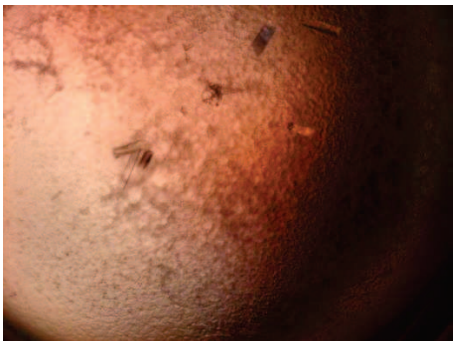
Trials with P-gp-Linkerless in the previously conditions (0.1 M Hepes pH 7–8, 50 mM lithium sulfate, 10 mM EDTA, and 24–29.5 % w/v PEG 600 Da at 4 °C) were set-up. Furthermore, a completed screen of almost 576 condition (Table 17) were set-up in co-work with Rupak Doshi (Post-Doc), Mark Villaluz (staff research associated) and Paul Szewczyk (PhD student). These conditions use either non hydrolysable ATP-analogues (e.g., AMPPNP) or ATP, in presence of magnesium to favor the catalysis of ATP. Sodium orthovanadate (Na_3VO_4) is a phosphate-analogue that traps an ADP•Pi•Vi complex in a post-hydrolysis state. This results in an outward-facing conformation with dimerized NBDs. Palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE) was used to improve the catalytic activity of P-gp in presence of verapamil.

Protein condition	precipitating agent and %	pH	Salt
P-gp-Linkerless + $MgSO_4$ + AMPPNP	w/ 350 MME 15-30% w/ 250 DME 15-30% w/ PEG 600 15-30% w/ PEG 400 15-30% w/ PEG 1000 10-25% w/ PEG 1500 10-25% Δ % = 0.5	0.1 M Hepes pH 6-8.2, Δ pH = 0.2	50 mM of $(NH_4)_2SO_4$ or NaCl
P-gp-Linkerless + $MgSO_4$ + Sodium Orthovanadate (Na_3VO_4) + ATP			
P-gp-Linkerless + POPE + verapamil + $MgSO_4$ + Na_3VO_4 + ATP			

Table 17. Crystallization conditions tested for P-gp-linkerless.
POPE: palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine; 350 MME : Polyethylene glycol 350 monomethyl ether; 250 DME: Polyglycol DME 250; PEG 600: Polyethylene glycol 600; PEG 400: Polyethylene glycol 400; PEG 1000: Polyethylene glycol 1000; PEG 1500: Polyethylene glycol 1500; *Final concentration:* $MgSO_4$, Na_3VO_4 and ATP were at 10 mM; POPE was at 0.1 mg/ml; verapamil was at 0.1 mM.

Crystals in most conditions (P-gp- Linkerless alone, in presence of either AMPPNP or ATP + Na_3VO_4) were obtained. Unfortunately, we only could conclude that the crystals of P-gp-linkerless in presence with AMPPNP has a different symmetry (Monoclinic, P2) compare to the P-gp- Linkerless alone (Orthorhombic, $P2_12_12_1$) because the resolution crystal didn't allow us to go further (Figure 61).

A) Crystals obtained



B) Pattern of X-ray diffraction

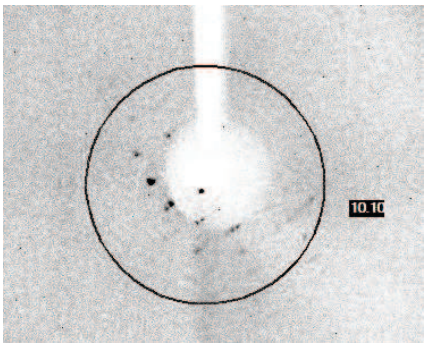


Figure 61. Best condition for P-gp-linkerless.
10 mM AMPPNP, 10 mM $MgSO_4$, 50 mM $AmSO_4$, 18 % PEG 1500, 0.1 M HEPES pH 6.2.

Part III: P-gp validation model

The mouse P-gp 3D-structure was resolved at 4 Å which is a moderate diffraction resolution. In order to validate this structure, 63 single Cys-mutants were expressed, purified and crystallized. Once the crystal was formed, it was soaked in ethyl-mercury chloride. This molecule produces an anomalous signal (show in red mesh in Figure 62) allowing to locate with a high precision the cysteine residue in the 3D-structure. P-gp has 7 cysteine (show in green spheres in Figure 62), thus the presence of the eighth signal is the residue sought (show in blue spheres in Figure 62).

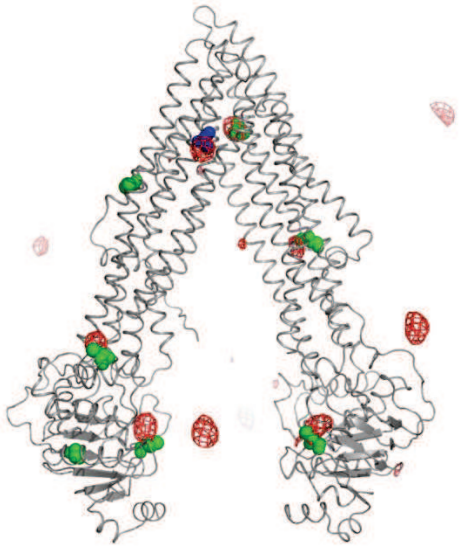


Figure 62. Cysteine mutants show Hg sites.
The seven cysteine of P-gp are show in spheres, colored in green. The residue sought is show in spheres, colored in blue. Anomalous signal is show in red mesh.

Thus, 24 position were validated (Figure 1B, publication II): 17 in this study (highlight in blue in Figure 63) and 7 from the previous published inward-facing x-ray structures of mouse P-gp [166].

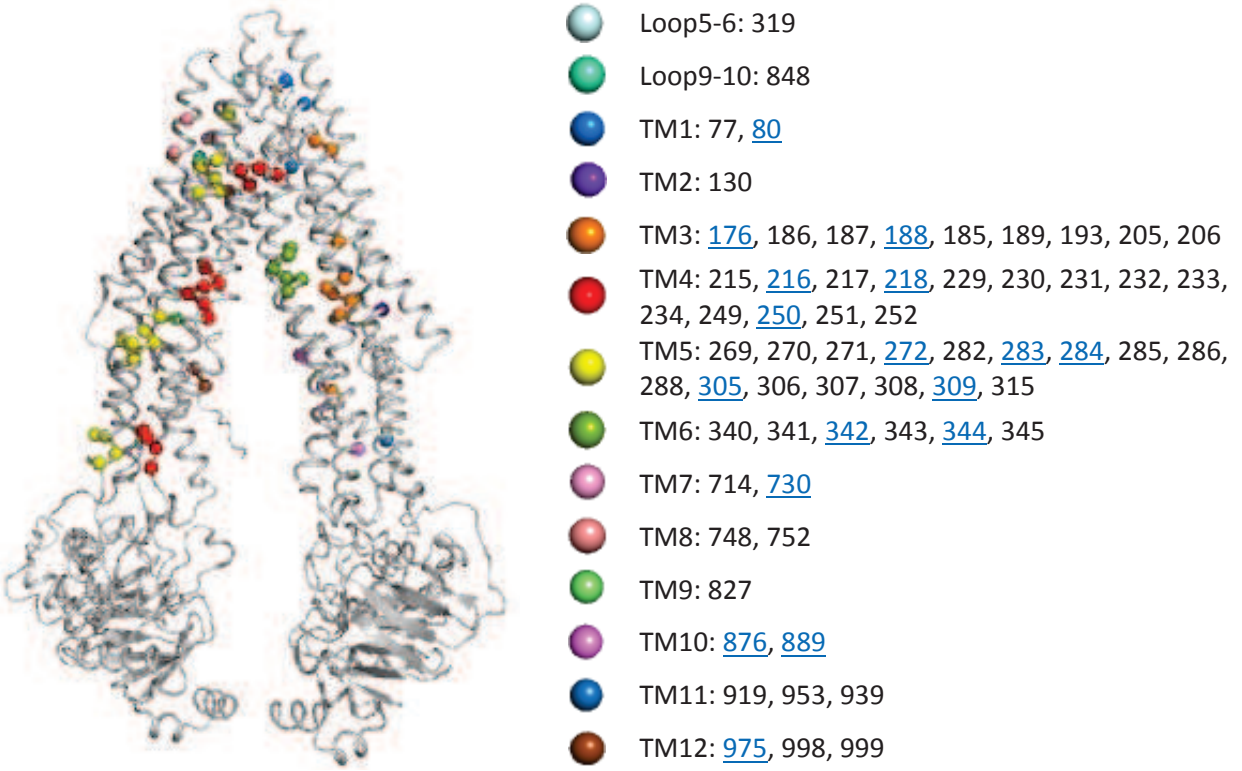


Figure 63. Composite view of all mutants realized in mouse P-gp.
Spheres represent C-α atoms of residues in the TM. Highlight in blue 17 residues were validated in publication II.

This series of crystallization and crystallography experiments led to the identification of three new inward-facing conformations of mouse P-gp. Two of them were derived from the same crystal form. In general, the overall protein fold and topologies of these structures were wider (the distance between NBD1 and NBD2) from those published in 2009 (Figure 64A), but comparable to other published structures of ABC exporters [167, 170, 171]. Like the original P-gp structures, the linker (residues 627-683), sites important for phosphorylation, between NBD1 and TM7 is not well observed and presumed to be disordered.

Another 3D-structure of P-gp was solved in complex with a camel nanobody (Nb), called Nb592 (Figure 64B and Figure 2 in publication II). It binds selectively to the first NBD of P-gp and stabilizes the corresponding protein conformation. It is also a very potent inhibitor of the ATPase activity of P-gp that prevents dimerization of NBD, essential for the hydrolysis of nucleotides and transporting drugs (Figure 3, publication II). Interestingly, this Nb592 only interacts with NBD1 and not with NBD2 although the high protein sequence similarity between both NBDs (Figure 3, Supporting Information of the publication II). Its capacity to prevent the ATP-hydrolysis makes this nanobody a good candidate to inhibit P-gp, providing that the nanobody enters into cells. For example, a mix of the Nb with nanoparticles or peptide/chemical tags may improve cell entry.

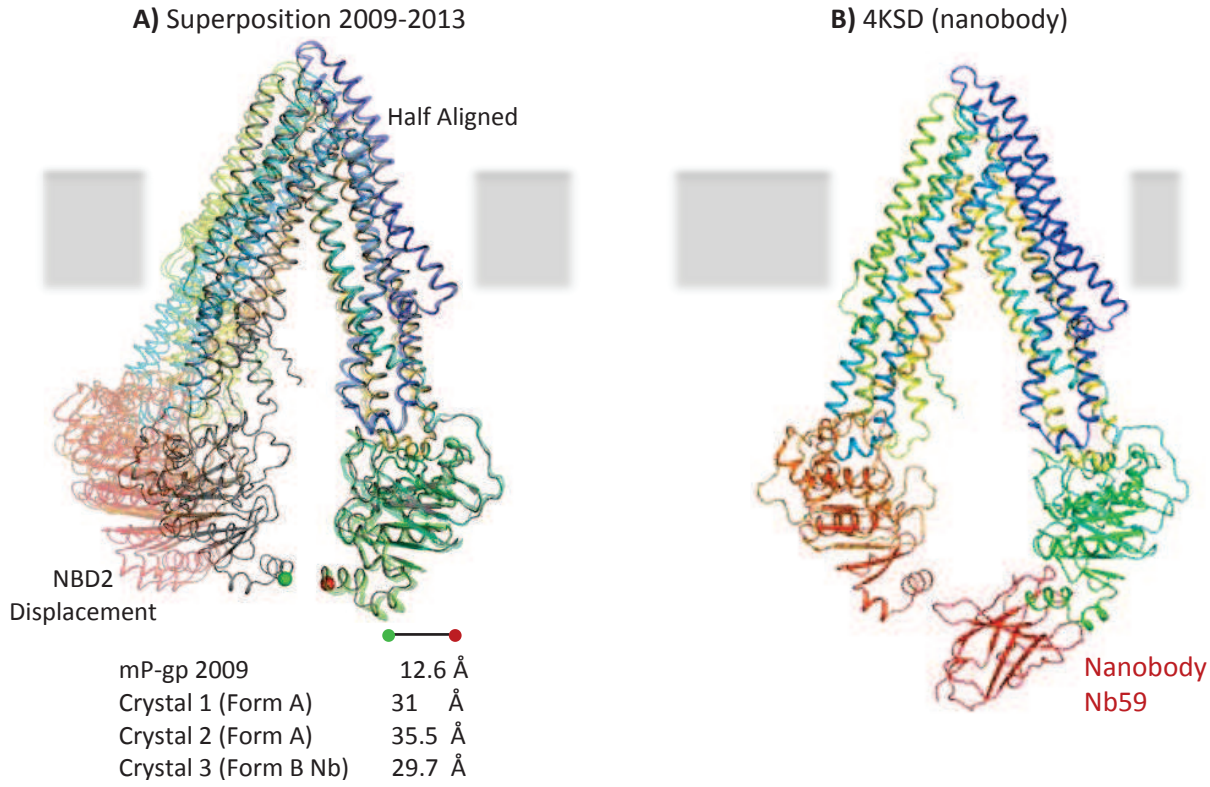


Figure 64. Different P-gp structures from Chang's group.
The Apo P-gp structure from 2009 (PDB code: 3G5U) is show in gray. The new P-gp structures are show in cartoon colored by rainbow (N-terminal in blue, C-terminal in red)
"Half Aligned": X-ray structures aligned by using residues in TMD1 and NBD1 (residues 33–209, 852–961, and 320–626); The distance between C-term of NBD1 (residue 626 in red) to C-term of NBD2 (residue 1271 in green) are indicated.

These results (part III) were published in PNAS and were presented in the following conference:

- Seminaries "Bases Moléculaires et Structurales des Systèmes Infectieux" (BMSSI) spring 2014 (February 2014, Lyon, France). Oral presentation: Conformational flexibility of the multidrug resistance P-glycoprotein. Martinez L.

Part IV: P-gp's structures

Related to this, Li *et al.*, [169] published a set of corrected mouse P-gp structures using single-wavelength anomalous (SAD) phasing, instead of multiwavelength anomalous dispersion (MAD) phasing as was done for the 2009 structures. The improved experimental electron density map, allow to build a better model of P-gp. Figure 65 displays the electron density map of 2009 mouse structure and the corrected one. We can observe that the new map is better defined, revealing new details. Furthermore, four amino-acid (A32, A216, A244, A280 and A302) assignments were confirmed by mercury labeling.

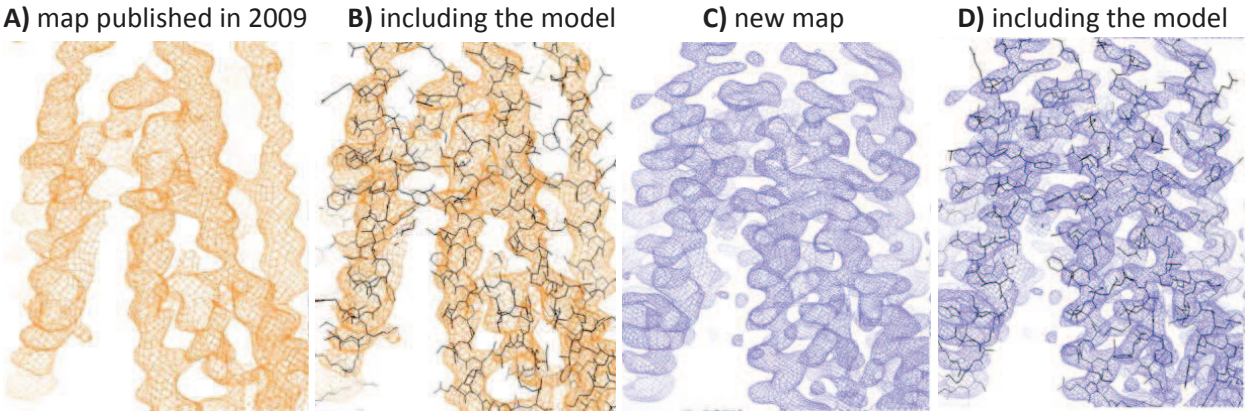


Figure 65. Experimental electron density maps.
The models are shows in gray sticks
According to [169]

These structures and the 2013 Chang's group structures are similar to the *Cænorhabditis elegans* P-gp structure. Is the 2009 mouse structures misinterpreted? I will present here differences in transmembrane helix 3, 4, 5 and coupling helix two (CH2) between all the P-gp's structures published so far.

Since the structures are in different 3D-coordinates, structural alignment has been done to facilitate the comparison. Thus, *C. elegans*, mouse and corrected mouse P-gp structures (PDB code: 4F4C, 4KSC and 4M2T respectively) were aligned toward the C-ter moety of mouse P-gp (PDB code: 3G61) by using residues in TMD1 and NBD1 (residues 33–206, 320-626, and 852-961 for the mouse P-gp's and 60-234+351-664+897-1004 for the *C. elegans* P-gp). Excepting for the comparison of TM4-CH2-TM5 region, for which those structures (4F4C, 4KSC and 4M2T) were aligned toward the CH2 of mouse P-gp (PDB code: 3G61) by using residues 243-281 for the mouse P-gp's and 271-309 for the *C. elegans* P-gp.

Figures 66-69 displays the region compared in magenta (panels A). The amino acid sequence alignment of human (H), mouse (M) and *C. elegans* (C) P-gp of the compared zone (panels B). Where the compared residues were highlighted in gray. Residues located with heavy atoms (mercury or selenium) were show in balls and underlined: colored in green and firebrick were mercury-labelled cysteines while in orange were selenium-labelled methionines, excepting for the residue 284 which was mercury-labelled cysteine. The panels C-G show the comparison of *C. elegans* (PDB code: 4F4C) *vs* mouse P-gp 2009 (PDB code: 3G61), *C. elegans* (PDB code: 4F4C) *vs* mouse P-gp 2013 (PDB code: 4KSC), *C. elegans* (PDB code: 4F4C) *vs* mouse P-gp corrected 2009 (PDB code: 4M2T), mouse P-gp 2009 (PDB code: 3G61) *vs* mouse P-gp corrected 2009 (PDB code: 4M2T) and mouse P-gp 2013 (PDB code: 4KSC) *vs* mouse P-gp corrected 2009 (PDB code: 4M2T), respectively. The color codes are: *C. elegans* P-gp structure (PDB code 4F4C) is shown in light orange; original mouse P-gp structure (PDB code 3G61) is shown in blue and corrected mouse P-gp structure (PDB code 4M2T) is shown in firebrick.

A detailed comparison of *C. elegans* and mouse P-gp structures was done by Jin *et al.*, [167]. Both structures have a low resolution 3.4 and 3.8 Å respectively. However the *C. elegans* P-gp

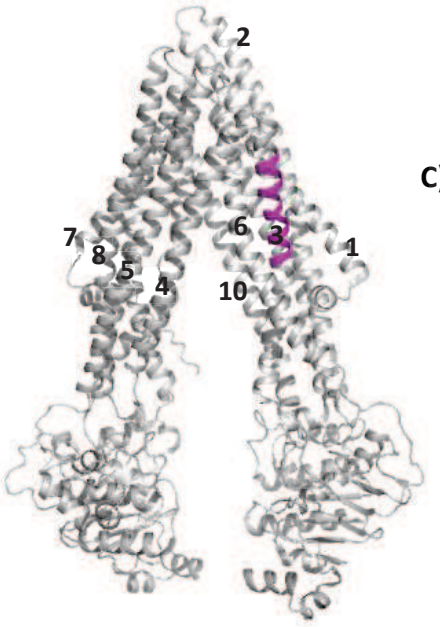
structure has 34 selenium-labelled methionines and three mercury-labelled cysteines, while mouse P-gp structure only has seven labeled residues with mercury. These heavy metal are helpful to validate the individual residue position. According to Jin *et al.*, the mouse P-gp structure displays significant differences in the orientation of the transmembrane helices 3, 4 and 5 (involved in drug−substrate binding): a shift of one amino acid in the TM3 helix, four amino acids in TM4 helix and three amino acids in TM5 helix.

The panel C of Figure 66 and Figure 67 confirm these differences between mouse and the *C. elegans* P-gp structure. Comparing the position of the phenylalanine 200 (mouse) with phenylalanine 228 (*C. elegans*) is slightly shifted (Figure 66C) and the position of the serine 218 (mouse) with threonine 246 (*C. elegans*) is move by almost one turn (Figure 67C). However, the panel C Figure 69 did not show a shift of three amino acid in TM5 helix as affirmed Jin, *et al*., comparing the position of the asparagine 275 (mouse) with threonine 303 (*C. elegans*).

Li *et al.*, [169] realized an in-depth comparison between the corrected mouse P-gp structures and the 2009 P-gp structures. A significant amount of changes were done through all P-gp. The authors also remark some mistakes in the *C. elegans* P-gp structure concerning the salt bridge, which allow the contact between the NBDs and coupling helix.

Panels D and E of Figures 66-69 show the similarities between the new mouse structures (PDB codes: 4KSC and 4M2T) and *C. elegans* P-gp structure. For example, a rotation of the TM3 (corresponding to one residue), a rotation of the TM4 (corresponding to four residues), remodeling coupling helix 2 and a rotation of TM5. Panels F and G of Figures 66-69 show the changes of the corrected mouse P-gp towards the original mouse P-gp, and the similarities and slight differences with the 2013 mouse P-gp, respectively. However, Li *et al.,* [169] found discrepancies between the new mouse structures (PDB codes: 4KSC) and the mouse P-gp corrected 2009 (PDB code: 4M2T) in other regions: coupling helix 1, TM8, TM12 and the N-terminal elbow helix.

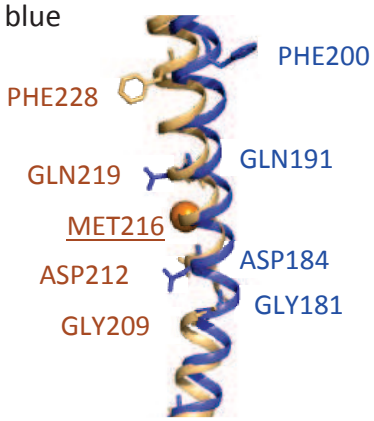
A) Compared zone colored in magenta in mouse P-gp (3G61)



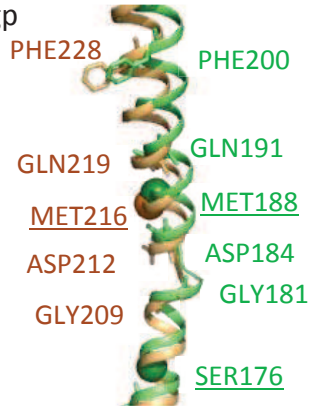
B) Amino acid sequence alignment of human (H), mouse (M) and *C. elegans* (C) P-gp.

```
H 176 TDDVSKINEGIGDKIGMFFQSMATFFTGFIVGFTRGW
M 172 TDDVSKINEGIGDKIGMFFQAMATFFGGFIIGFTRGW
C 200 FDNLERVKEGTGDKIGMAFQYLSQFITGFIVAFTHSW
```

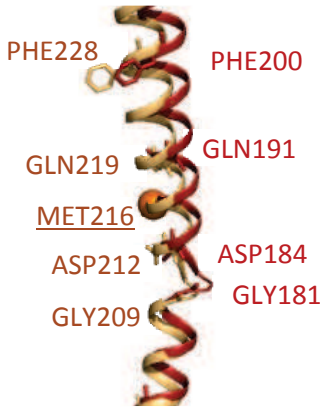
C) *C. elegans* (4F4C) in light orange *vs* mouse P-gp (3G61) in blue



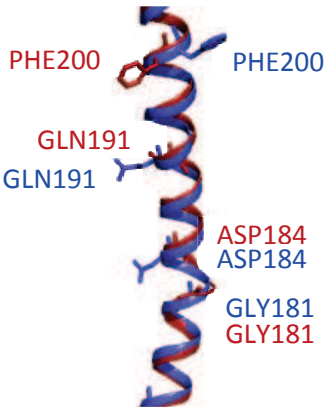
D) *C. elegans* (4F4C) in light orange *vs* new mouse P-gp (4KSC) in lime



E) *C. elegans* (4F4C) in light orange *vs* corrected mouse P-gp (4M2T) in firebrick



F) mouse P-gp (3G61) in blue vs corrected mouse P-gp (4M2T) in firebrick



G) new mouse P-gp (4KSC) in lime vs corrected mouse P-gp (4M2T) in firebrick

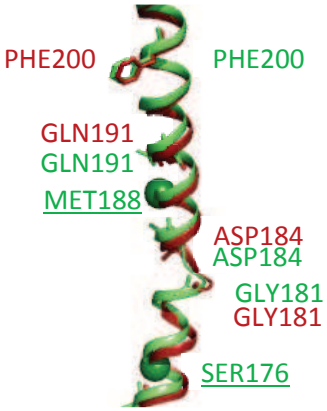
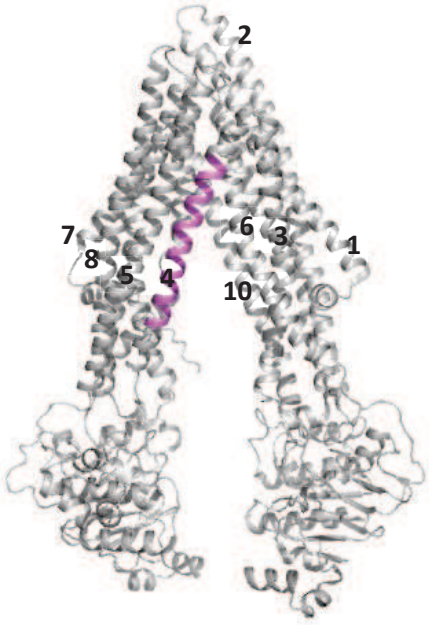


Figure 66. Structural comparison of helix TM3.

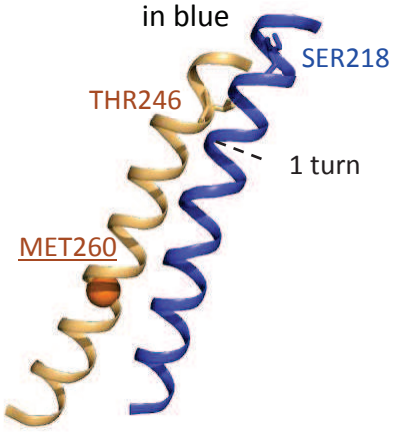
A) Compared zone colored in magenta in mouse P-gp (3G61)



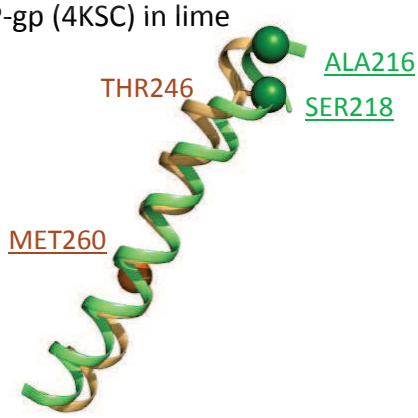
B) Amino acid sequence alignment of human (H), mouse (M) and *C. elegans* (C) P-gp.

```
H 219 LAISPVLGLSAAVWAKILSSFT
M 215 LAISPVLGLSAGIWAKILSSFT
C 243 LAVTPIQALCGFAIAKSMSTFA
```

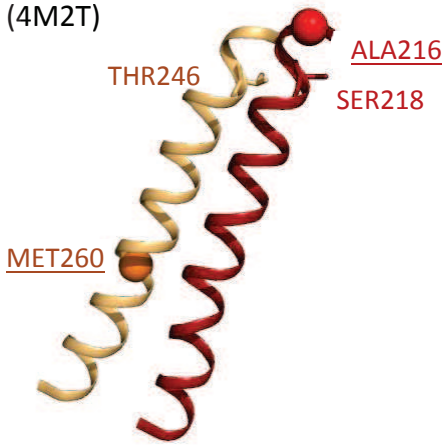
C) *C. elegans* (4F4C) in light orange *vs* mouse P-gp (3G61) in blue



D) *C. elegans* (4F4C) in light orange *vs* new mouse P-gp (4KSC) in lime



E) *C. elegans* (4F4C) in light orange *vs* corrected mouse P-gp (4M2T) in firebrick



F) mouse P-gp (3G61) in blue vs corrected mouse P-gp (4M2T) in firebrick

G) new mouse P-gp (4KSC) in lime vs corrected mouse P-gp (4M2T) in firebrick

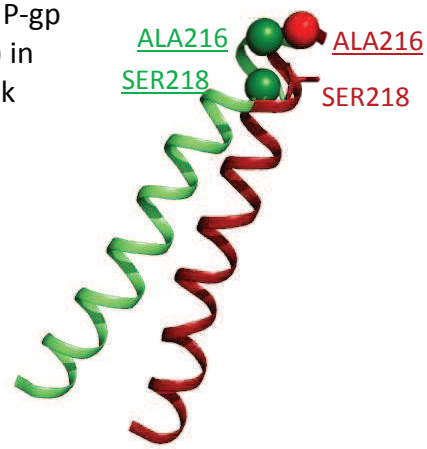
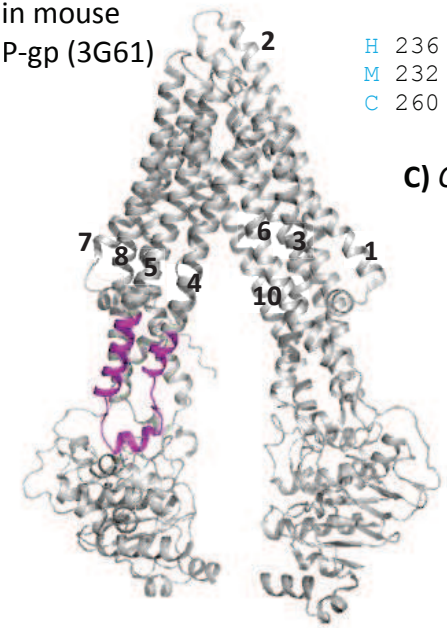


Figure 67. Structural comparison of helix TM4.

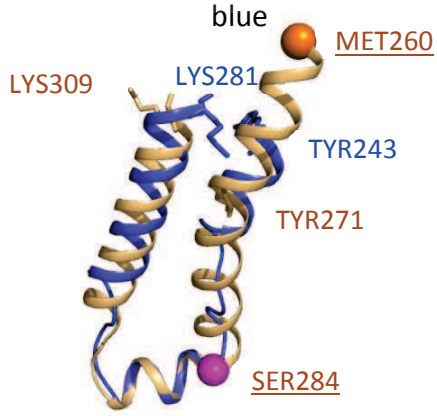
A) Compared zone colored in magenta in mouse P-gp (3G61)



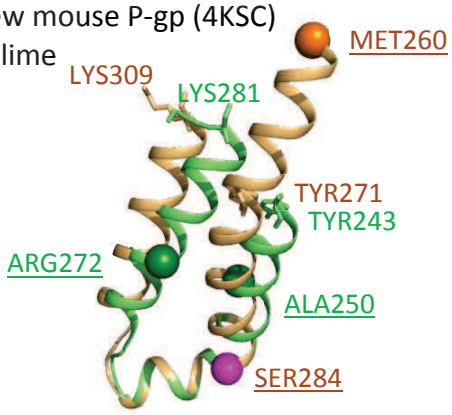
B) Amino acid sequence alignment of human (H), mouse (M) and *C. elegans* (C) P-gp.

```
H 236 LSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKELERYNKNLEEAK
M 232 LSSFTDKELHAYAKAGAVAEEVLAAIRTVIAFGGQKKELERYNNNLEEAK
C 260 MSTFAIRETLRYAKAGKVVEETISSIRTVVSLNGLRYELERYSTAVEEAK
```

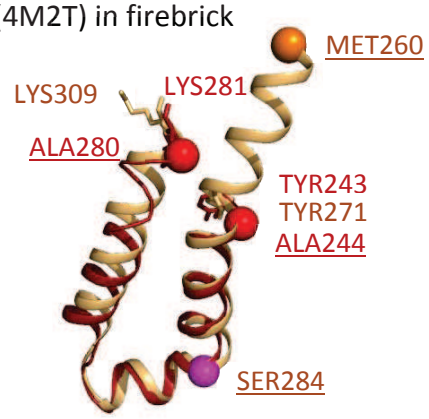
C) *C. elegans* (4F4C) in light orange *vs* mouse P-gp (3G61) in blue



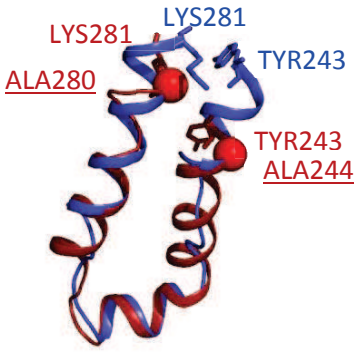
D) *C. elegans* (4F4C) in light orange *vs* new mouse P-gp (4KSC) in lime



E) *C. elegans* (4F4C) in light orange *vs* corrected mouse P-gp (4M2T) in firebrick



F) mouse P-gp (3G61) in blue vs corrected mouse P-gp (4M2T) in firebrick



G) new mouse P-gp (4KSC) in lime vs corrected mouse P-gp (4M2T) in firebrick

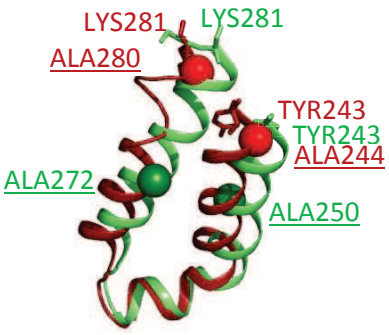
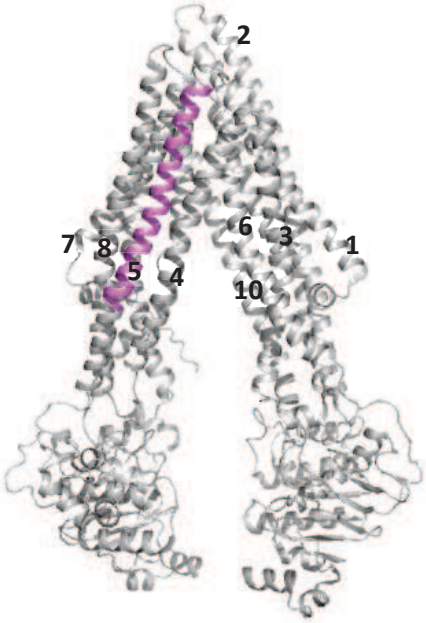


Figure 68. Structural comparison of TM4-CH2-TM5 region.

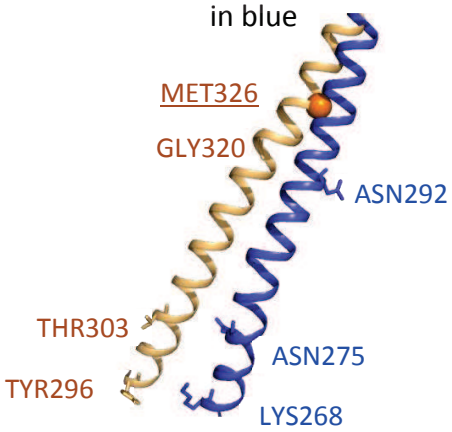
A) Compared zone colored in magenta in mouse P-gp (3G61)



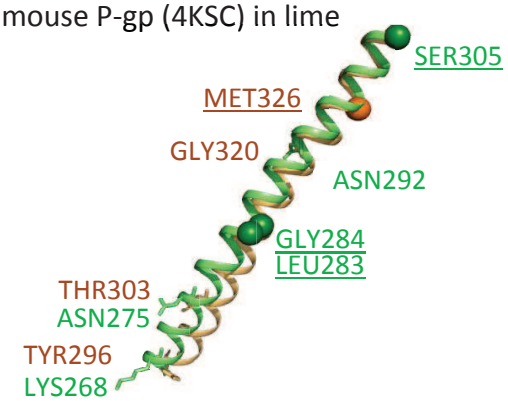
B) Amino acid sequence alignment of human (H), mouse (M) and *C. elegans* (C) P-gp.

```
H 272 KELERYNKNLEEAKRIGIKKAITANISIGAAFLLIYAS
M 268 KELERYNNNLEEAKRLGIKKAITANISMGAAFLLIYAS
C 296 YELERYSTAVEEAKKAGVLKGLFLGISFGAMQASNFIS
```

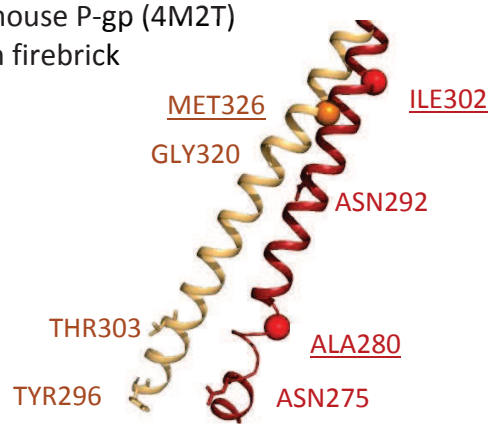
C) *C. elegans* (4F4C) in light orange *vs* mouse P-gp (3G61) in blue



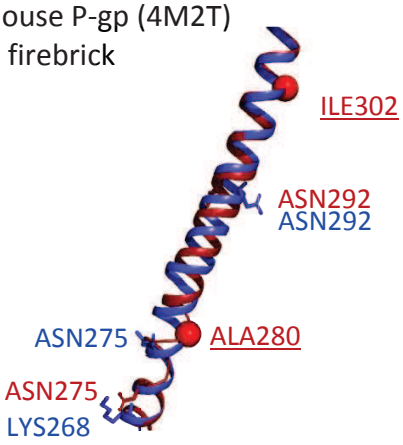
D) *C. elegans* (4F4C) in light orange *vs* new mouse P-gp (4KSC) in lime



E) *C. elegans* (4F4C) in light orange *vs* corrected mouse P-gp (4M2T) in firebrick



F) mouse P-gp (3G61) in blue vs corrected mouse P-gp (4M2T) in firebrick



G) new mouse P-gp (4KSC) in lime vs corrected mouse P-gp (4M2T) in firebrick

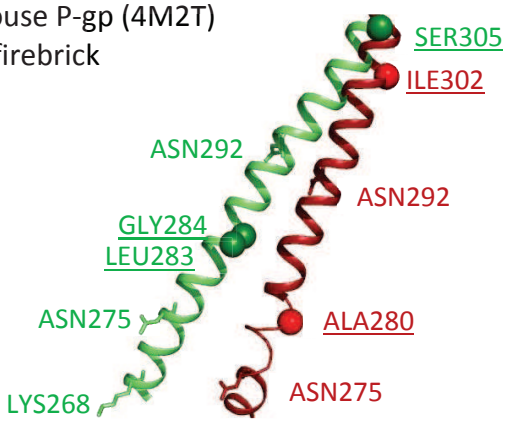


Figure 69. Structural comparison of helix TM5.

Publication II

Structures of P-glycoprotein reveal its conformational flexibility and an epitope on the nucleotide-binding domain

Andrew B. Ward[a,1], Paul Szewczyk[b,1], Vinciane Grimard[c], Chang-Wook Lee[d], Lorena Martinez[e], Rupak Doshi[d], Alexandra Caya[a], Mark Villaluz[d], Els Pardon[f,g], Cristina Cregger[d], Douglas J. Swartz[h], Pierre Guy Falson[e], Ina L. Urbatsch[h], Cedric Govaerts[c], Jan Steyaert[f,g], and Geoffrey Chang[d,i,2]

[a]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037; [b]Division of Biological Sciences, [d]School of Pharmacy and Pharmaceutical Sciences, and [i]Department of Pharmacology, School of Medicine, University of California at San Diego, La Jolla, CA 92093; [c]Laboratory of Structure and Function of Biological Membranes, Université Libre de Bruxelles, 1050 Brussels, Belgium; [e]Drug Resistance Mechanism and Modulation Laboratory, Bases Moléculaires et Structurales des Systémes Infectieux Unité Mixte de Recherche 5086, Centre National de la Recherche Scientifique–Lyon 1 University, Institute of Biology and Chemistry of Proteins, 69367 Lyon, France; [f]Structural Biology Research Centre and [g]Structural Biology Brussels, Vrije Universiteit Brussel, 1050 Brussels, Belgium; and [h]Cell Biology and Biochemistry, Texas Tech University Health Sciences Center, Lubbock, TX 79430

Edited by Hiroshi Nikaido, University of California, Berkeley, CA, and approved July 3, 2013 (received for review May 17, 2013)

P-glycoprotein (P-gp) is one of the best-known mediators of drug efflux-based multidrug resistance in many cancers. This validated therapeutic target is a prototypic, plasma membrane resident ATP-Binding Cassette transporter that pumps xenobiotic compounds out of cells. The large, polyspecific drug-binding pocket of P-gp recognizes a variety of structurally unrelated compounds. The transport of these drugs across the membrane is coincident with changes in the size and shape of this pocket during the course of the transport cycle. Here, we present the crystal structures of three inward-facing conformations of mouse P-gp derived from two different crystal forms. One structure has a nanobody bound to the C-terminal side of the first nucleotide-binding domain. This nanobody strongly inhibits the ATP hydrolysis activity of mouse P-gp by hindering the formation of a dimeric complex between the ATP-binding domains, which is essential for nucleotide hydrolysis. Together, these inward-facing conformational snapshots of P-gp demonstrate a range of flexibility exhibited by this transporter, which is likely an essential feature for the binding and transport of large, diverse substrates. The nanobody-bound structure also reveals a unique epitope on P-gp.

membrane protein structure | ABC transporter | nanobody-transporter complex

The mammalian adenosine triphosphate (ATP)-binding cassette (ABC) transporter P-glycoprotein (P-gp) is present in many tissues (1) and can detoxify cells by pumping xenobiotics across the membrane (2). A hallmark feature of this transporter is its ability to bind and transport an array of structurally diverse molecules ranging in size from 100 to 4,000 daltons (Da) (molecular mass) (3). In fact, P-gp was discovered in a cell line where it prevented the permeation of several different drugs (4). Although its physiological function is not fully understood, the well-recognized role of P-gp in mediating multidrug resistance (MDR) in many types of cancers has made it a valid therapeutic target (5). P-gp also plays a significant role in the pharmacokinetics and bioavailability of drugs by mediating their transport in the liver, intestines, and across the blood–brain barrier (1, 2, 6).

P-gp is a ∼170-kDa molecule comprised of two pseudosymmetric halves, each containing a nucleotide-binding domain (NBD) and a transmembrane domain (TMD). Currently, the "alternating access" model is the most widely accepted paradigm explaining the mechanics of transport by ABC transporters (7, 8). According to this model, binding of ATP at the NBDs drives conformational changes in the TMDs and switches the transporter's overall conformation from inward-facing to outward-facing (inward/outward refer to the opening of the drug-binding pocket relative to the cell). This ATP-driven switch results in the vectorial transport of substrates out of the cell. The hydrolysis of ATP and release of Pi/ADP are essential for resetting the transporter back to the inward-facing conformation.

The overall protein topology of the inward-facing conformation of P-gp has been established (9) and is reiterated in the X-ray structures of related bacterial and mammalian ABC transporters found in the ABCB subfamily (8, 10, 11), including the human ABCB10 [Protein Data Bank (PDB) ID codes 3ZDQ, 4AYX, 4AYT, and 4AYW] (12). Furthermore, the inward-facing conformation has been biochemically validated in bacterial ABC exporters, using electron paramagnetic resonance on multicopy suppressor of Htrb mutations (MsbA) (13) and LmrA (14), cysteine cross-linking of MsbA (15), and hydrogen/deuterium exchange coupled to mass spectrometry of BmrA (16). The inward-facing conformation is a key intermediate in the alternating access mechanism (7), as it allows the transporter to scan the inner leaflet of the membrane for substrates. This notion is supported by the inward-facing X-ray structures of P-gp, which have cyclic peptide inhibitors bound to the substrate-binding pocket (9).

Owing to its role in MDR, several small molecule inhibitors of P-gp have been developed during the last four decades (2, 17), but none have yet been approved for clinical use (18). In recent years, antibody-derived therapeutics have become increasingly popular because of their specific affinities and relatively high tolerance in humans. Nanobodies (Nbs) are small (∼15 kDa) single-domain proteins that are derived from, and contain, the unique structural and functional properties of natural heavy chain-only antibodies found in camelids (19). Like conventional antibodies, Nbs achieve antigen-specificity by using three loops called complementarity determining regions (CDRs 1–3).

Nbs are mostly monomeric, encoded by single genes, highly soluble, and can be efficiently produced in prokaryotic and eukaryotic hosts, including bacteria and yeast. They are also

Author contributions: A.B.W., P.S., V.G., C.-W.L., I.L.U., C.G., J.S., and G.C. designed research; A.B.W., P.S., V.G., C.-W.L., L.M., A.C., M.V., E.P., C.C., D.J.S., I.L.U., C.G., and G.C. performed research; V.G., E.P., C.G., and J.S. contributed new reagents/analytic tools; A.B.W., P.S., V.G., R.D., P.G.F., I.L.U., and G.C. analyzed data; and A.B.W., P.S., V.G., R.D., P.G.F., I.L.U., and G.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 4KSB, 4KSC, and 4KSD).

[1]A.B.W. and P.S. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: g1chang@mail.ucsd.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1309275110/-/DCSupplemental.

capable of penetrating many tissues including the human gut (19). These characteristics provide Nbs with great promise because therapeutics against a number of diseases (19). Additionally, Nbs can bind to cryptic epitopes and potentially stabilize proteins in certain conformations. These properties make Nbs useful as chaperones for inducing lattice formation during protein crystallization (20).

Here, we present three unique X-ray structures of mouse P-gp (*ABCB1a*) with different inward-facing conformations, all wider than those originally published (9). One crystal form comprises P-gp complexed with a single Nb, called Nb592, bound to the first NBD (NBD1). Consistent with its binding location, we show that Nb592 is a potent inhibitor of P-gp's ATP hydrolysis activity. Taken together, these different structures of P-gp reveal a unique epitope and demonstrate the flexibility of the transporter to sample a range of inward-facing conformations.

Results

Structure Determination. To structurally map the conformational flexibility of P-gp in the absence of nucleotides and substrates, we determined three X-ray structures (Table S1) that crystallized in two different crystal forms (crystal forms A and B; Table S1). Depending on the crystal growth condition, crystal form A yielded crystals that had variations in unit cell dimensions (Table S1) with changes mostly in the c- and a-cell edges (crystal 1 and crystal 2). To reduce bias, we determined an experimentally derived electron density map of crystal form A by generating a heavy atom derivative using ethyl-mercury-chloride (Table S1, labeled "Experimental Map Crystal") and obtained protein phases by using the multiple anomalous dispersion technique (MAD). An initial model was built by using this electron density map to 4.5 Å (Fig. S1). A higher resolution native crystal (crystal 1; Fig. 1*A* and Table S1) later diffracted to 3.8 Å and was solved by the molecular replacement method. The structure revealed that the distance between the NBDs, as measured by the Cα atom of residues 626 (C-term of NBD1) and 1271 (C-term of NBD2), is ∼31 Å, which is much larger than the ∼13 Å observed in the previously published mouse P-gp structure (9). This represents a large (∼18 Å) range of displacement between the NBDs that P-gp can conformationally sample while in the inward-facing state.

A second native crystal of P-gp (crystal 2) diffracted to 4.0 Å and was determined by molecular replacement using our model derived from crystal 1. This structure of P-gp is the widest, with a distance of ∼36 Å between residues 626 and 1271. Crystals 1 and 2 represent the more extreme ranges of the variations observed for crystal form A. The two X-ray structures revealed distinct inward-facing conformations of P-gp, with the overall structure similar to the one published (9) and some smaller local changes. For example, a break in the helical structure of transmembrane helix (TM) 12 in the original more closed inward-facing conformation is fully helical in these structures. The linker (residues 627–683) between NBD1 and TM7 is not observed and presumed to be disordered. Residues 1–33 and 1272–1276 at the N- and C- termini are also not observed.

To further validate the models of P-gp (crystals 1–2) derived from crystal form A, which is critical at these moderate diffraction resolutions (3–4.5 Å), we made 17 single-site cysteine substitution mutants in P-gp and labeled them with ethyl mercury chloride. These mutants included residues S80C, S176C, M188C, A216C, S218C, A250C, R272C, L283C, G284C, S305C, A309C, G342C, A344C, K730C, S876C, S889C, and S975C distributed throughout the TM helices. Fig. 1*B* shows a composite view of all experimentally validated positions mapped on to the crystal 1 model, whereas Fig. S2 provides close-up shots of each mercury-labeled single-site mutant. Combined with seven wild-type cysteine residues of mouse P-gp previously identified (9) and

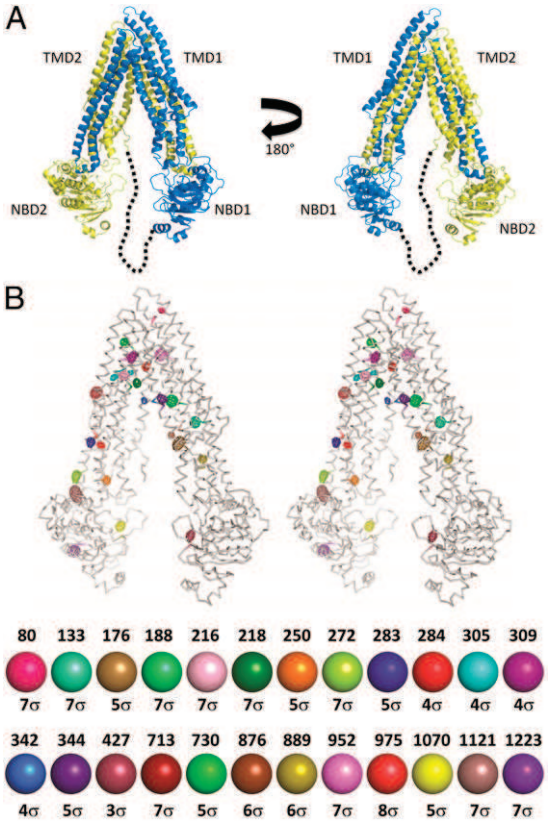


Fig. 1. Structure of inward-facing P-gp. (*A*) Open inward-facing conformation (crystal 1) with the NBDs far apart. The N-terminal half of the protein (blue) and the C-terminal half (yellow) are connected by a flexible linker region (black dashed line) that is disordered in the structure. (*B*) Stereo view of superposition of anomalous Fourier mercury peaks validating the position of knocked-in and wild-type cysteine residue positions. Relative peak positions were determined by aligning the corresponding TMD or NBD subdomains of the model derived from the mutant diffraction data (Table S2) with the model derived from crystal 1. The locations of the 24 anomalous mercury peaks (mesh, sigma values) were used to confirm the registration of amino acids in the structure.

observed in these data, we were able to experimentally validate the accuracy of our model by using a total of 24 positions.

A second, entirely different crystal form (crystal form B) of P-gp was obtained by cocrystallization of mouse P-gp with the nanobody Nb592 (Fig. 2 and Table S1). Size exclusion chromatography demonstrated that the P-gp and Nb592 form a complex that comigrates and elutes as a single peak (Fig. S3). Fractions from this sample peak were pooled together, concentrated, and crystallized. The X-ray structure of the P-gp–Nb592 complex (crystal 3) was determined to a resolution of 4.1 Å (Table S1) by molecular replacement using the model derived from crystal 1 and a model of Nb592 generated as described in *Materials and Methods*. The CDR loops of Nb592 were found to bind NBD1, thereby potentially precluding the NBDs from dimerizing. This inward-facing conformation is the least wide among the three structures presented here, with a distance of ∼30 Å between the Cα atoms of residues 626 and 1271.

The interface between Nb592 and NBD1 of P-gp is significant and constitutes ∼745 A^2 of the buried surface area. CDR3 (residues 98–107 of Nb592) penetrates into a pocket formed by three helices in NBD1 (Fig. 2*B*). The P-gp residues within 5 Å of Nb592 include S555, T559, E562, A563, H583, R584, L585, S586, T587, R589, H608, M612, F619, L621, V622, M623, T624, Q625, and T626 on NBD1 and K1260, F1264, S1265, and

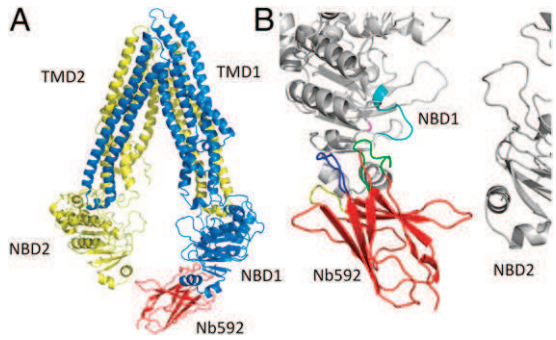


Fig. 2. X-ray structure of P-gp in complex with Nb592. (*A*) Overview of the entire structure. Nb592 (red) binds to the C terminus of NBD1. There are additional interactions with NBD2 (yellow). The binding site of the nanobody precludes the ABC domains from coming together, explaining its potent ATPase inhibition properties. (*B*) Close up view of Nb592 binding site on P-gp. The view is rotated 180° from *A*. The complementarity determining regions (CDR1: residues 25–33 for Nb592, blue; CDR2: residues 51–57 for Nb592, yellow; CDR3: residues 98–107, green) of the nanobody all interact with the C-terminal portion of NBD1. CDR3 inserts into a shallow pocket formed by three helices in NBD1. The walker-A motif (residues 423–430) located on NBD1 is colored in cyan. The conserved H583 is also shown in violet.

S1268 on NBD2 (Fig. S4*A*). T626 is the last resolved residue at the C terminus of NBD1 and is proceeding a flexible linker region presumed disordered but could also potentially interact with Nb592.

The Pgp-Nb592 structure (crystal 3) elucidates a unique epitope on the NBDs of P-gp. A protein sequence alignment between mouse and human P-gp reveals that the two proteins are similar in the epitope/Nb592-binding site (Fig. S4*B*). The only differences are at residues 623 and 624 (mouse numbering): mouse, Met-Thr; human, Thr-Met. Although these residues are within 5 Å of Nb592, T624 is pointing away from Nb592 and M623 is oriented toward the conserved portion of the Nb just before and after CDR2 (residues 551–557 of Nb592). Interestingly, Nb592 does not bind to mouse NBD2, which also has a high protein sequence similarity in this region (Fig. S4*B*).

Inhibition of P-gp ATPase by Nb592. The NBDs of ABC transporters are responsible for the binding and hydrolysis of ATP that drives drug transport through the TMDs. The structure of P-gp with Nb592 bound at NBD1 (Fig. 2) indicates that this Nb could significantly affect the catalytic function of P-gp. We tested this hypothesis by using two biochemical approaches. First, we measured the drug (verapamil)-stimulated ATPase activity of purified mouse P-gp, in the presence of increasing concentrations of Nb592. Our results revealed that Nb592 is a strong inhibitor of ATP hydrolysis with half-maximal inhibition achieved in the nanomolar range (IC_{50} of 520 ± 57 nM) (Fig. 3*A*). This IC_{50} value is comparable to or lower than those reported for most small molecule P-gp inhibitors (2, 9, 17, 21). Nb592 also potently inhibited basal ATPase activity (in the absence of drug) with 90 ± 5% inhibition seen at saturating amounts of 7.5 μM Nb592. Moreover, the inhibition was fully retained in the presence of 5 mM DTT (Fig. 3*B*), indicating that Nb592 is a robust protein that can withstand reducing agents at concentrations that are typical of the cytoplasm.

We made further measurements that suggested that Nb592 prevents dimerization of the NBDs. ATP hydrolysis requires intimate interactions of residues from both *cis* and *trans* NBDs with bound ATP to facilitate the hydrolytic attack on the γ-phosphate (22). Sodium orthovanadate (Vi) is an inorganic phosphate (Pi) analog that can trap ATP in the posthydrolysis state (ADP•Pi) (23), resulting in an outward-facing conformation with dimerized NBDs (8). Using 8-azido-[α^{32}P]-ATP, we show that 8-azido-[α^{32}P]-ADP is trapped in mouse P-gp by Vi, in the presence of the ATP hydrolysis activator, verapamil (Fig. 3*C*, lanes 1–3) (24). At increasing concentrations of Nb592, the hydrolysis of ATP is effectively inhibited as indicated by the reduced amount of trapped 8-azido-[α^{32}P]-ADP in the catalytic sites of P-gp (Fig. 3*C*, lanes 4–10).

Taken together, both biochemical experiments show that the binding of Nb592 to mouse P-gp inhibits its ATP hydrolysis activity by hindering the formation of an ATP hydrolysis-competent NBD dimer sandwich. These results complement the structure of the Pgp-Nb592 complex (Fig. 2), suggesting a molecular basis for the inhibition of ATPase activity. This unique mechanism of inhibiting P-gp by specifically targeting the NBDs contrasts the mode of inhibition by several small molecule inhibitors/drugs, which target the substrate-binding cavity in the TMDs (2, 9, 17).

Discussion

The overall topology and protein fold of the three unique P-gp X-ray structures described in this work are similar to those published (9), with some small localized differences. When the N-terminal halves of the P-gp models in this study were structurally aligned (residues 33–209, 852–961, and 320–626), the C-terminal halves showed a significant displacement relative to one another, most notably in the relative positions of NBD2 (Fig. 4). The overall change in the structures can be described as a hinge movement where a small angular change at the pivot point causes larger changes farther away toward the NBDs. The movement near the pivot or "TM-hinge region" on the extracellular side is smaller, and comprises Loop 3–4 (L3-4; residues 206–208) and Loop 5–6 (L5-6; residues 319–324), and the

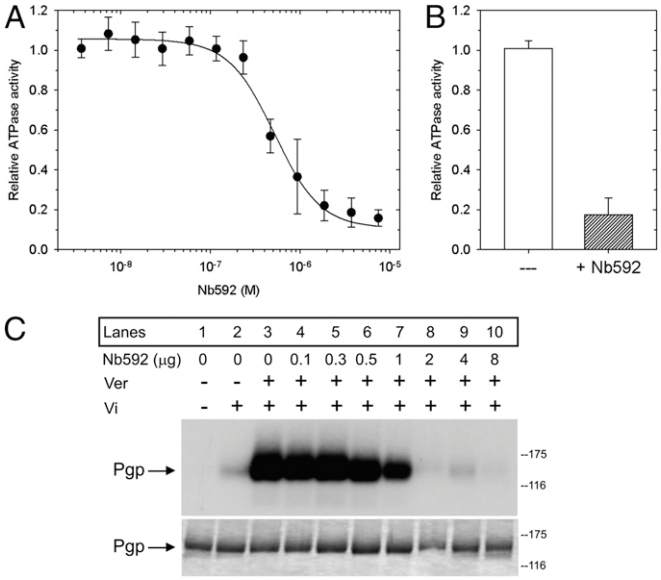


Fig. 3. Nb592 is a strong inhibitor of P-gp's ATPase activity. (*A*) Nb592 inhibits verapamil-stimulated ATPase activity with an IC_{50} of 520 ± 57 nM. Data points indicate the average activity ± SEM from four independent experiments, relative to P-gp's activity in the absence of Nb592. Lines represent nonlinear regression analysis of the data points; R^2 value for the fit was 0.95. (*B*) Inhibition of verapamil-stimulated ATPase activity in the presence of 5 mM DTT, without or with 7.5 μM Nb592 ($n = 4$). (*C*) Nb592 prevents vanadate-induced 8-azido-[α^{32}P]-ADP trapping in P-gp's catalytic sites. P-gp was incubated with 100 μM verapamil (VER) and 200 μM orthovanadate (Vi) in the 8-azido-[α^{32}P]-ATP hydrolysis/trapping reaction as indicated above the lanes. *Upper*, autoradiogram; *Lower*, Coomassie-stained SDS/PAGE gel showing the presence of P-gp (loading control).

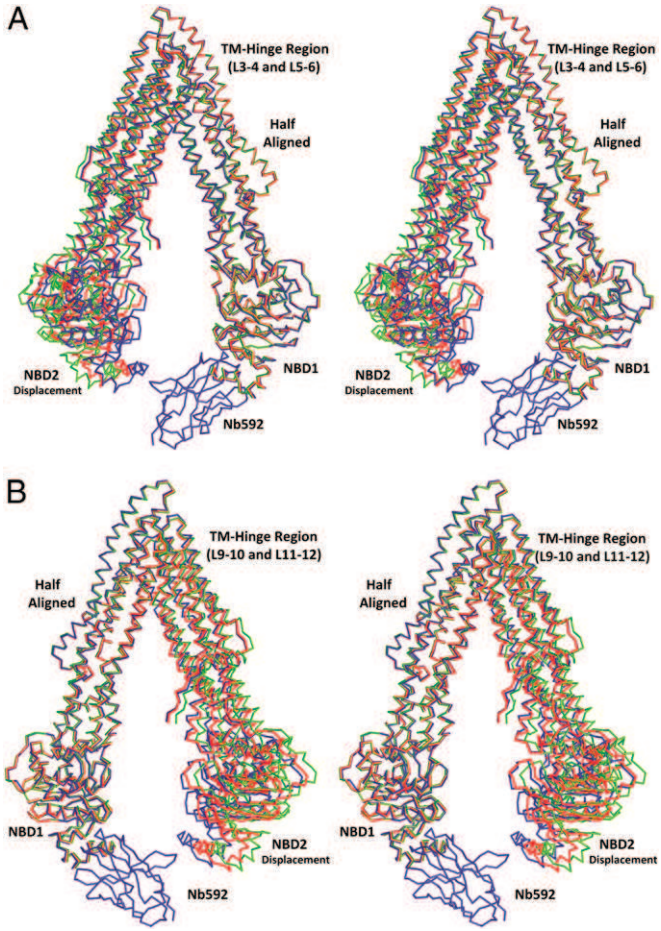


Fig. 4. Conformational changes by P-gp. (*A*) Stereoview of the three P-gp structures described in this study (crystal 1, red; crystal 2, green; P-gp–Nb592 complex [crystal 3], blue) aligned by using residues in TMD1 and NBD1 (residues 33–209, 852–961, and 320–626; designated as "Half Aligned"). The rmsd on Cα atoms for the aligned portion was 0.11 Å between crystal 2 and crystal 1 and 0.28 Å between P-gp–Nb592 complex (crystal 3) and crystal 1. The "TM-hinge regions" for this half of the molecule (L3-4 and L5-6 as described in the text) are marked. The relatively large displacement of the other half of the molecule, including NBD2, is clearly shown. The relative position of the Nb592 is marked in interaction with NBD1. It also makes a smaller contact with NBD2. The P-gp–Nb592 complex is the most closed inward-facing conformation described in this study. (*B*) Same structural alignment as *A* except turned 180° to show the opposite side of the transporter. The "TM-hinge region" comprised of L9-10 and L11-12 is indicated.

corresponding Loop 9–10 (L9-10; residues 849–851) and Loop 10–11 (L10-11; residues 963–968) on the other half of the molecule.

The most striking feature of these P-gp structures is a much larger separation between NBD1 and NBD2, suggesting that the transporter can adopt a much wider inward-facing conformation than previously described. The overall range of distances between residues 626 (C-term of NBD1) to 1271 (C-term of NBD2) sampled by the three inward-facing conformations presented here is ∼29–36 Å, whereas the corresponding distance in the original published structure (9) is only ∼13 Å. Such dynamic conformational flexibility in the inward-facing state has also been observed in biochemical, biophysical, and molecular dynamics simulations experiments on bacterial ABC transporters, such as MsbA (13) and BmrA (16), and recently on mouse P-gp (25). The enlargement of the portals (formed from TM4/TM6 and TM10/TM12) facing the inner-membrane leaflet side may be required for larger substrates like β-amyloid (∼4 kDa in size) (26) to enter and bind inside the substrate-binding pocket.

The wideness of our P-gp structures is comparable to recently published structures of other ABC transporters having similar ABC B–like protein folds (8, 11, 12, 27). Interestingly, the structures of human ABCB10 and *Thermotoga maritima* TM287/288 (10) also have nucleotide analogs bound to separated NBDs, suggesting a state of these transporters just before the formation of the ABC sandwich that is essential for the structural transition to the outward-facing conformation. The degree of NBD separation might also be influenced in part by crystal lattice contacts and may ultimately be constrained in the physiological context by the thickness of the hydrophobic section of the lipid bilayer and the TMDs of the transporter.

The discovery and structural elucidation of unique epitopes on therapeutic targets has tremendous value in the pharmaceutical industry. Antibodies like UlC2 that have been developed against P-gp target epitopes on the TMDs/extracellular surface (28). The structure of P-gp complexed with Nb592 (crystal 3) reveals an epitope located on NBD1 (Fig. 2*B*) that is away from the Walker-A or conserved histidine (H583). The interaction between Nb592 and NBD1 of P-gp appears to be quite specific, as the CDR loops of Nb592 do not bind NBD2, despite a relatively high level of sequence conservation between NBD1 and NBD2 (Fig. 2*D*). The binding location of Nb592 on NBD1 (Fig. 2) suggests that it sterically prevents the NBDs from dimerizing and provides a structural basis to support and explain our biochemical observations that Nb592 (*i*) strongly inhibits ATPase activity, and (*ii*) completely abolishes Vi-induced ADP-trapping, in mouse P-gp (Fig. 3) signifying that NBD dimerization, followed by ATP "occlusion," is essential to the formation of the hydrolysis-competent state during the transport cycle. Future biochemical studies may establish whether Nb592 can bind and inhibit other ABC transporters via a similar mechanism.

Nb592 is a robust protein and will likely bind to its intracellular epitope even when exposed to the reducing glutathiones in the cytoplasm, as our studies suggest (Fig. 3). This result is quite remarkable considering the presence of two highly conserved cysteines, C22 and C96, that form a disulfide bond in the Nb, as observed in our structure. Further validation of these types of inhibitors in whole-cell settings will be necessary to assess their penetration through cell membranes and cross-reactivity with other cellular ATPases. Cell penetration can be improved through innovations such as small peptide/chemical tags (29), whereas specificity may be enhanced through molecular scaffold engineering. Although the IC_{50} of Nb592 (520 nM) is better than the QZ59 inhibitors that cocrystallized in the original P-gp structures (9), it is an order of magnitude less than Tariquidar (43 nM), a P-gp inhibitor that has been used in clinical trials (18, 30). Thus, although Nb592 itself might not directly be used as a clinical inhibitor, it provides a template for the future development of this type of molecular scaffold. Several techniques may also be applied to improve the binding affinity of Nb592 (31).

The X-ray structures of P-gp described here, along with those published (9), will be useful for the molecular modeling of conformational trajectories to understand how the substrate-binding pocket changes during the transport cycle (25). Because human and mouse P-gp share nearly 87% protein sequence identity, our structures present experimentally derived checkpoints useful for docking simulations. The development of these algorithms to more accurately model and predict substrate binding will provide a useful screening tool, complementing in vivo and cell-based studies, for understanding P-gp's role in drug pharmacokinetics. Collectively, these models could facilitate the development of therapeutically important compounds evading P-gp in several clinically relevant contexts, like penetration of drugs through the blood–brain barrier.

Materials and Methods

P-gp Protein Expression and Purification. Gene-optimized mouse P-gp (*ABCB1a*, GenBank JF834158) was expressed as described (9, 32) in 10-L cultures of *Pichia pastoris* grown in a bioreactor (Bioflow 415; New Brunswick Scientific). Protein expression was induced by addition of methanol (3.6 mL/h per liter of culture volume) overnight. Cells were lysed by a single pass through a cell disrupter (TS-Series; Constant Systems) at 40,000 psi. Cell wall and debris were removed by centrifugation (3,500 × *g*, 35 min, 4 °C), and crude plasma membranes were obtained by centrifugation at 35,267 × *g* for 2–3 h at 4 °C. The purification procedure is similar to that described (9) with some modifications. Membranes containing P-gp were resuspended in cold buffer (100 mM NaCl, 15% glycerol, 20 mM Tris at pH 8.0, and Sigma protease inhibitors) and solubilized with a final percentage of ~4.5% (vol/vol) Triton X-100 for 1–2 h at 4 °C. Insoluble material was removed by centrifugation at 38,400 × *g*, 4 °C for 30–60 min, and the supernatant was poured over a metal resin (Ni-NTA Superflow; Qiagen) by FPLC (AKTA; GE Life Sciences). Immobilized P-gp was buffer exchanged into 20 mM Hepes, 20 mM imidazole, 0.04% sodium cholate (Sigma), and 0.0675% β-dodecyl maltoside (DDM). Protein was eluted with buffer containing 200 mM imidazole at pH 7.5. The eluted protein was then diluted 1:10 in buffer containing 20 mM Hepes at pH 8.0, 100 mM NaCl, 0.2 mM tris(2-carboxyethyl)phosphine (TCEP), 0.04% sodium cholate, 0.065% β-DDM, and rebound to a new Ni-NTA column. The column was washed with buffer containing 20 mM imidazole, eluted with 200 mM imidazole, the protein was concentrated (Centricon YM-50 or YM-100; Millipore), spun at 95,000 rpm (TLA120.1 rotor) for 30–60 min at 4 °C, and subjected to gel filtration chromatography (Superdex200 16/60; GE Healthcare) at 4 °C.

Elicitation of Nanobodies Against P-gp. To generate the nanobodies, 2 mg of purified, detergent-solubilized mouse P-gp was injected into a llama (*Llama glama*) over a period of 6 wk to elicit an immune response. The immunization, library construction, and nanobody selection have been performed by following standard procedures according to ref. 33, and a C-terminal His6-tagged nanobody library in pMES4 (GenBank accession no. GQ907248) of 2×10^8 independent clones was established by using the PstI/BstEII site. P-gp–specific phages were recovered by incubating P-gp–coated wells with 100 mM triethylamine at pH 11.0, for 10 min. The P-gp–coated wells were then washed once with Tris·HCl at pH 6.8, and several times with PBS. Finally, freshly grown TG1 cells were added to the wells to recover the noneluted phage. After two rounds of selection, individual colonies were screened for the expression of P-gp–specific nanobodies: Maxisorb 96-well plates were coated overnight at 4 °C with 1 μg/mL P-gp in sodium bicarbonate buffer at pH 8.2. Residual protein-binding sites in the wells were blocked for 2 h at room temperature with 2% milk in PBS. Detection of P-gp–bound nanobodies was performed with a mouse anti-His tag monoclonal (Serotec). Subsequent detection of the mouse anti-tag antibodies was done with an alkaline phosphatase anti-mouse IgG conjugate (Sigma). The absorption at 405 nm was measured 30–60 min after adding the enzyme substrate 4-nitrophenyl phosphate. Plasmids were extracted from the positive clones and transformed in *Escherichia coli* WK6 strains.

Nanobody Expression and Purification. Nb592 protein was produced in the *E. coli* WK6 strain described above. Bacteria were grown in terrific broth to an OD_{600} of 0.7 and then expression was induced by 1 mM IPTG overnight at 28 °C. Bacteria were then pelleted at 7,500 × *g* for 15 min at room temperature. Pellets were resuspended in 15 mL of TES buffer (0.2 M Tris at pH 8.0, 0.5 mM EDTA, and 0.5 M sucrose) and kept under slow agitation for 1 h at 4 °C. Thirty milliliters of fourfold diluted TES buffer was added, and the sample was osmotically lysed under slow agitation for 45 min at 4 °C. Samples were then centrifuged for 30 min at 4 °C and 6,000 × *g*. Supernatant was used for purification on Ni-NTA resin (Qiagen). Binding to the Ni-NTA resin was performed at 4 °C for 1 h. The column was washed with 50 mM phosphate buffer at pH 6.0, 1 M NaCl, then eluted with 50 mM sodium acetate buffer at pH 4.5 and 1 M NaCl. The protein solution was neutralized by using 1 M Tris at pH 7.5. The eluted protein was then subjected to gel filtration (Superdex75 10/300GL; GE Healthcare) at 4 °C.

Formation of P-gp–Nanobody Complex. The P-gp–Nb592 complex was generated by incubating threefold stoichiometric excess nanobody with P-gp purified from the metal resin elution step for 30 min at 4 °C. The complex was then subjected to gel filtration chromatography as described above to remove excess Nb592.

ATPase Inhibition. Purified P-gp was reduced with 1 mM DTT for 30 min on ice, and then excess DTT was removed by passage through 1-mL Sephadex G-50 centrifuge columns equilibrated in 20 mM Hepes at pH 7.4, 10% glycerol, 250 mM NaCl, and 0.1% DDM as described in ref. 34. Protein was activated with 1% (wt/vol) *E. coli* polar lipids (Avanti) for 10 min at room temperature. Lipid-activated P-gp (0.5–2 μg) was preincubated with increasing amounts of Nb592 for 30 min on ice, then ATP hydrolysis assayed with 10 mM MgATP and 100 μM verapamil in a final volume of 50 μL for 15 min at 37 °C. Release of inorganic phosphate was determined by the Malachite green method (34). Negative controls containing P-gp, Nb592, or both were assayed for 1 min at 37 °C and were subtracted as background values. In some experiments, P-gp was substituted for Cys-less P-gp (35), purified in the absence of reducing agents, which gave essentially the same results. The data were fitted by using nonlinear regression on Sigmaplot (v11) with the equation: $Y = d - (a \times x^b/(c^b + x^b))$, where *a* is the verapamil-stimulated activity in the absence of Nb592, *b* is the Hill coefficient, *c* is the concentration for half-maximal inhibition (IC_{50}), *d* is the maximal inhibition, and *x* is the concentration of Nb592.

For vanadate trapping experiments, P-gp was activated with 1 mM DTT and *E. coli* lipids (2:1, wt/wt) for 10 min at room temperature (34). Lipid-activated Pgp (3.3 μg, 3–5 μL) was preincubated with increasing concentrations of Nb592 for 5 min, and reacted with 200 μM 8-azido-[α32P]-ATP, 2 mM Mg^{2+}, and/or 100 μM verapamil and/or 200 μM orthovanadate in a final volume of 100 μL for 15 min at 37 °C. Unbound nucleotide was removed by passage through centrifuge columns, and the eluates UV cross-linked for 7 min on ice as described in ref. 34. Samples were resolved on 10% SDS/PAGE gels stained with Coomassie Brilliant Blue, and the dried gels were exposed to film.

Crystallization, Data Collection, and Structure Determination of P-gp and P-gp-Nb592 Complex. Purified P-gp or P-gp-Nb592 were isolated after gel filtration chromatography at a protein concentration of 1–2 mg/mL and subjected to reductive methylation (36). Freshly made borane and formaldehyde were added to the protein solution at final concentrations of 50 mM and 100 mM, respectively, and incubated in the dark for 2 h at 4 °C with gentle shaking. The reaction was quenched with the addition of ice cold 2.5 mM glycine and incubated for 30 min at 4 °C. Methylated P-gp was then concentrated to 1 mL (Centricon YM-50 or YM-100; Millipore 4) and subsequently diluted with 9 mL of quench buffer (20 mM Tris at pH 7.5, 100 mM NaCl, 0.2 mM TCEP, 0.04% sodium cholate, and 0.065% β-DDM). The concentration/dilution step was repeated two times. In some cases, 2 mM methyl-β cyclodextrin was added to the dilute protein and the mixture was concentrated for crystallization.

Immediately before crystallization, P-gp and the P-gp–Nb592 complex were concentrated to 8–12 mg/mL P-gp–only crystals (crystal form A) described in this work were grown by using the sitting drop method at 4 °C by combining protein and precipitant at 1:1 (volume:volume). Crystal 1 and crystal 2, and all point mutants of Pgp (Table S2), were grown at a protein concentration of 10–12 mg/mL by using 0.1 M Hepes (pH 7–8), 50 mM lithium sulfate, 10 mM EDTA, and 25–29.5% (wt/vol) PEG 600 at 4 °C. These crystals typically appeared after 1–3 d and continued to grow to full size in approximately 2 wk. Crystals of the P-gp–Nb592 complex were grown by using 0.1 M Hepes (pH 7–8) and 22–27% (wt/vol) PEG 600.

X-ray diffraction data were collected on cryo-cooled crystals at the Stanford Synchrotron Radiation Laboratory (BL 11–1), the Advanced Light Source (ALS) (BL5.0.1) and the Advanced Photon Source (23-ID-B, and 23-ID-D). Datasets were processed with HKL2000 (HKL Research, Inc.) and mosflm (37). Experimentally derived protein phases were obtained via the multiple anomalous dispersion technique by using diffraction data collected on the mercury LIII edge and inflection point (Table S1) calculated by the program PHASES. The overall combined phasing power was 1.9, with a figure of merit of 0.732. The experimental electron density map revealed that there was only one molecule in the asymmetric unit. An initial P-gp model was positioned by using PDB ID code 3G5U using the program Coot (38). The structures of crystal 1 and crystal 2 were determined by molecular replacement using the package MolRep as part of the CCP4 suite (37). Similar to the original published structures of P-gp, the N terminus (residues 1–33) was not visualized and no electron density was present for most of the linker region (residues 627–683), which is likely a flexible region connecting the two halves of P-gp.

To further validate these structures of P-gp (crystal 1 and crystal 2), we introduced single-site cysteine mutants throughout the TM regions (Table S2). PCR-based mutagenesis was performed with pairs of complementary mutagenic primers that carry the desired cysteine (codon TGT) by using the *pLIC-Opti-Pgp* expression plasmid (32) as a template, *Pfx50* DNA polymerase (Invitrogen), and the In-Fusion HD cloning kit (Clontech). All plasmids were verified by sequencing. Each mutant was expressed in *P. pastoris* strain

KM71H (9, 32), purified, and crystallized as described above. Crystals of P-gp mutants were soaked with 5 mM ethyl-mercury chloride for 1–2 h, flash cooled, and X-ray data was collected at synchrotrons and by using our in-house X-ray source (Bruker). The identity and position of the corresponding mercury-labeled cysteine residues yielded peaks in the anomalous difference Fourier (Fig. 1*B* and Fig. S2).

For the structure determination of the P-gp–Nb592 complex, the molecular replacement method was used followed by rigid body refinement of the TMD and NBD regions using crystal 1 as the model. The position of Nb592 was determined by using a $2F_o$-F_c difference map using phases derived from only a P-gp model. A homology model for Nb-592 was built from chain A of PDB ID code 1HCV by using Swiss Model (39) and idealized by using the Chiron server (http://troll.med.unc.edu/chiron/login.php). The preliminary nanobody model was manually fit into difference density and refined against the X-ray data by using rigid body refinement.

Crystallographic refinement using native data from all crystals of P-gp was accomplished by using the simulated annealing protocol (mlf target; CNS v1.3) and later using Phenix v1.8.2–1309 (40). A final round of group B-factor refinement and bulk solvent correction produced chemical models for crystals 1–3. The chemical geometry of the refined P-gp and P-gp–Nb592 complex structures were corrected and checked with molprobity (41), yielding models with no violations in the Ramachandran phi-psi plot and good bond angle/bond-length geometry (Table S1) comparable with other structures in this moderate resolution range. All structures of P-gp were validated by using a sigma-A weighted $2F_o$-F_c composite simulated annealing (SA) omit map (iteratively omitting 5% of the model) and multiple F_o-F_c SA difference maps. The F_o-F_c maps were calculated by using CNS v1.3 systematically omitting 11 consecutive residues throughout the model, generously omitting a neighboring sphere size of 4.0 Å, and a map cushion surrounding the omitted region of 2.0 Å. All models of P-gp and P-gp–Nb592 were also validated by using F_o-F_c difference maps Figs. S5–S7. Figures were generated by using PyMOL (14) and Adobe Photoshop 7.0.

ACKNOWLEDGMENTS. We thank Dr. Qinghai Zhang for useful comments on the manuscript, the support staff at beamlines 11-1 (Stanford Synchrotron Radiation Laboratory) and 23-IDD (Advanced Photon Source), and Nele Buys for assistance in the selection of the Nbs. This work was funded by National Institutes of Health (NIH) Grants R01 GM94367 and R01 ES021985 (to P.S., M.V., C.C., C.-W.L., A.B.W., A.C., and G.C.), Innoviris-Brussels (V.G., C.G., E.P., and J.S.), startup funds to A.B.W. from the Scripps Research Institute, cluster 10-ARC1santé and Explora'doc from Rhône-Alpes region (L.M.), the Ligue Contre le Cancer, ANR-09-PIRI-0002-01 and ANR-EMMA-10-049-01 (to P.G.F.), NIH Grant RGM102928, and Cancer Prevention Research Institute of Texas Grant RP101073 (to I.L.U.). C.G. is a Chercheur Qualifié of the Fonds de la Recherche Scientifique.

1. Cascorbi I (2011) P-glycoprotein: Tissue distribution, substrates, and functional consequences of genetic variations. *Handb Exp Pharmacol* (201):261–283.
2. Eckford PD, Sharom FJ (2009) ABC efflux pump-based resistance to chemotherapy drugs. *Chem Rev* 109(7):2989–3011.
3. Schinkel AH, Jonker JW (2003) Mammalian drug efflux transporters of the ATP binding cassette (ABC) family: An overview. *Adv Drug Deliv Rev* 55(1):3–29.
4. Gottesman MM, Ling V (2006) The molecular basis of multidrug resistance in cancer: The early years of P-glycoprotein research. *FEBS Lett* 580(4):998–1009.
5. Marquez B, Van Bambeke F (2011) ABC multidrug transporters: Target for modulation of drug pharmacokinetics and drug-drug interactions. *Curr Drug Targets* 12(5): 600–620.
6. Lee CA, Cook JA, Reyner EL, Smith DA (2010) P-glycoprotein related drug interactions: Clinical importance and a consideration of disease states. *Expert Opin Drug Metab Toxicol* 6(5):603–619.
7. Gutmann DAP, Ward A, Urbatsch IL, Chang G, van Veen HW (2010) Understanding polyspecificity of multidrug ABC transporters: Closing in on the gaps in ABCB1. *Trends Biochem Sci* 35(1):36–42.
8. Ward A, Reyes CL, Yu J, Roth CB, Chang G (2007) Flexibility in the ABC transporter MsbA: Alternating access with a twist. *Proc Natl Acad Sci USA* 104(48):19005–19010.
9. Aller SG, et al. (2009) Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science* 323(5922):1718–1722.
10. Hohl M, Briand C, Grütter MG, Seeger MA (2012) Crystal structure of a heterodimeric ABC transporter in its inward-facing conformation. *Nat Struct Mol Biol* 19(4):395–402.
11. Jin MS, Oldham ML, Zhang Q, Chen J (2012) Crystal structure of the multidrug transporter P-glycoprotein from Caenorhabditis elegans. *Nature* 490(7421):566–569.
12. Shintre CA, et al. (2013) Structures of ABCB10, a human ATP-binding cassette transporter in apo- and nucleotide-bound states. *Proc Natl Acad Sci USA* 110(24): 9710–9715.
13. Zou P, Bortolus M, McHaourab HS (2009) Conformational cycle of the ABC transporter MsbA in liposomes: Detailed analysis using double electron-electron resonance spectroscopy. *J Mol Biol* 393(3):586–597.
14. Hellmich UA, et al. (2012) Probing the ATP hydrolysis cycle of the ABC multidrug transporter LmrA by pulsed EPR spectroscopy. *J Am Chem Soc* 134(13):5857–5862.
15. Doshi R, Woebking B, van Veen HW (2010) Dissection of the conformational cycle of the multidrug/lipidA ABC exporter MsbA. *Proteins* 78(14):2867–2872.
16. Mehmood S, Domene C, Forest E, Jault J-M (2012) Dynamics of a bacterial multidrug ABC transporter in the inward- and outward-facing conformations. *Proc Natl Acad Sci USA* 109(27):10832–10836.
17. Crowley E, McDevitt CA, Callaghan R (2010) Generating inhibitors of P-glycoprotein: Where to, now? *Methods Mol Biol* 596:405–432.
18. Tamaki A, Ierano C, Szakacs G, Robey RW, Bates SE (2011) The controversial role of ABC transporters in clinical oncology. *Essays Biochem* 50(1):209–232.
19. Muyldermans S, et al. (2009) Camelid immunoglobulins and nanobody technology. *Vet Immunol Immunopathol* 128(1-3):178–183.
20. Lam AY, Pardon E, Korotkov KV, Hol WG, Steyaert J (2009) Nanobody-aided structure determination of the EpsI:EpsJ pseudopilin heterodimer from Vibrio vulnificus. *J Struct Biol* 166(1):8–15.
21. Shukla S, Ohnuma S, Ambudkar SV (2011) Improving cancer chemotherapy with modulators of ABC drug transporters. *Curr Drug Targets* 12(5):621–630.
22. Seeger MA, van Veen HW (2009) Molecular basis of multidrug transport by ABC transporters. *Biochim Biophys Acta* 1794(5):725–737.
23. Chen J, Sharma S, Quiocho FA, Davidson AL (2001) Trapping the transition state of an ATP-binding cassette transporter: Evidence for a concerted mechanism of maltose transport. *Proc Natl Acad Sci USA* 98(4):1525–1530.
24. Urbatsch IL, Sankaran B, Bhagat S, Senior AE (1995) Both P-glycoprotein nucleotide-binding sites are catalytically active. *J Biol Chem* 270(45):26956–26961.
25. Wen PC, Verhalen B, Wilkens S, McHaourab H, Tajkhorshid E (2013) On the origin of large flexibility of P-glycoprotein in the inward-facing state. *J Biol Chem* 288(26): 19211–19220.
26. Kuhnke D, et al. (2007) MDR1-P-glycoprotein (ABCB1) mediates transport of Alzheimer's amyloid-beta peptides—implications for the mechanisms of Abeta clearance at the blood-brain barrier. *Brain Pathol* 17(4):347–353.
27. Park YJ, Pardon E, Wu M, Steyaert J, Hol WG (2012) Crystal structure of a heterodimer of editosome interaction proteins in complex with two copies of a cross-reacting nanobody. *Nucleic Acids Res* 40(4):1828–1840.
28. Mechetner EB, et al. (1997) P-glycoprotein function involves conformational transitions detectable by differential immunoreactivity. *Proc Natl Acad Sci USA* 94(24): 12908–12913.
29. Deshayes S, et al. (2010) Structural polymorphism of non-covalent peptide-based delivery systems: Highway to cellular uptake. *Biochim Biophys Acta* 1798(12): 2304–2314.
30. Martin C, et al. (1999) The molecular interaction of the high affinity reversal agent XR9576 with P-glycoprotein. *Br J Pharmacol* 128(2):403–411.
31. De Genst E, et al. (2004) Chemical basis for the affinity maturation of a camel single domain antibody. *J Biol Chem* 279(51):53593–53601.
32. Bai J, et al. (2011) A gene optimization strategy that enhances production of fully functional P-glycoprotein in Pichia pastoris. *PLoS ONE* 6(8):e22577.
33. Conrath KE, et al. (2001) Beta-lactamase inhibitors derived from single-domain antibody fragments elicited in the camelidae. *Antimicrob Agents Chemother* 45(10): 2807–2812.
34. Urbatsch IL, Gimi K, Wilke-Mounts S, Senior AE (2000) Conserved walker A Ser residues in the catalytic sites of P-glycoprotein are critical for catalysis and involved primarily at the transition state step. *J Biol Chem* 275(32):25031–25038.
35. Tombline G, et al. (2006) Expression, purification, and characterization of cysteine-free mouse P-glycoprotein. *Arch Biochem Biophys* 445(1):124–128.
36. Rayment I (1997) Reductive alkylation of lysine residues to alter crystallization properties of proteins. *Methods Enzymol* 276:171–179.
37. Winn MD, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):235–242.
38. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.
39. Schwede T, Kopp J, Guex N, Peitsch MC (2003) SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res* 31(13):3381–3385.
40. Adams PD, et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213–221.
41. Chen VB, et al. (2010) MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66(Pt 1):12–21.

Supporting Information

Ward et al. 10.1073/pnas.1309275110

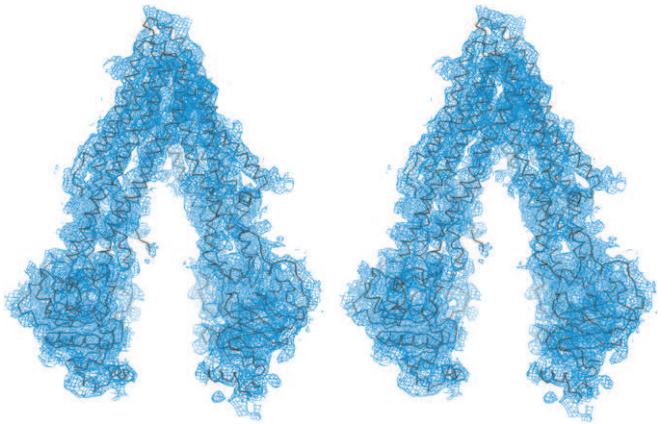


Fig. S1. Experimental electron density map of more open inward-facing conformation of P-gp. Stereoview of 4.5-Å resolution electron density map (blue mesh) contoured at 1σ. The P-gp monomer is presented in gray ribbon. The map was generated by using the program PHASES (1).

1. Furey W, Swaminathan S (1997) PHASES-95: A program package for processing and analyzing diffraction data from macromolecules. *Methods Enzymol* 277:590–620.

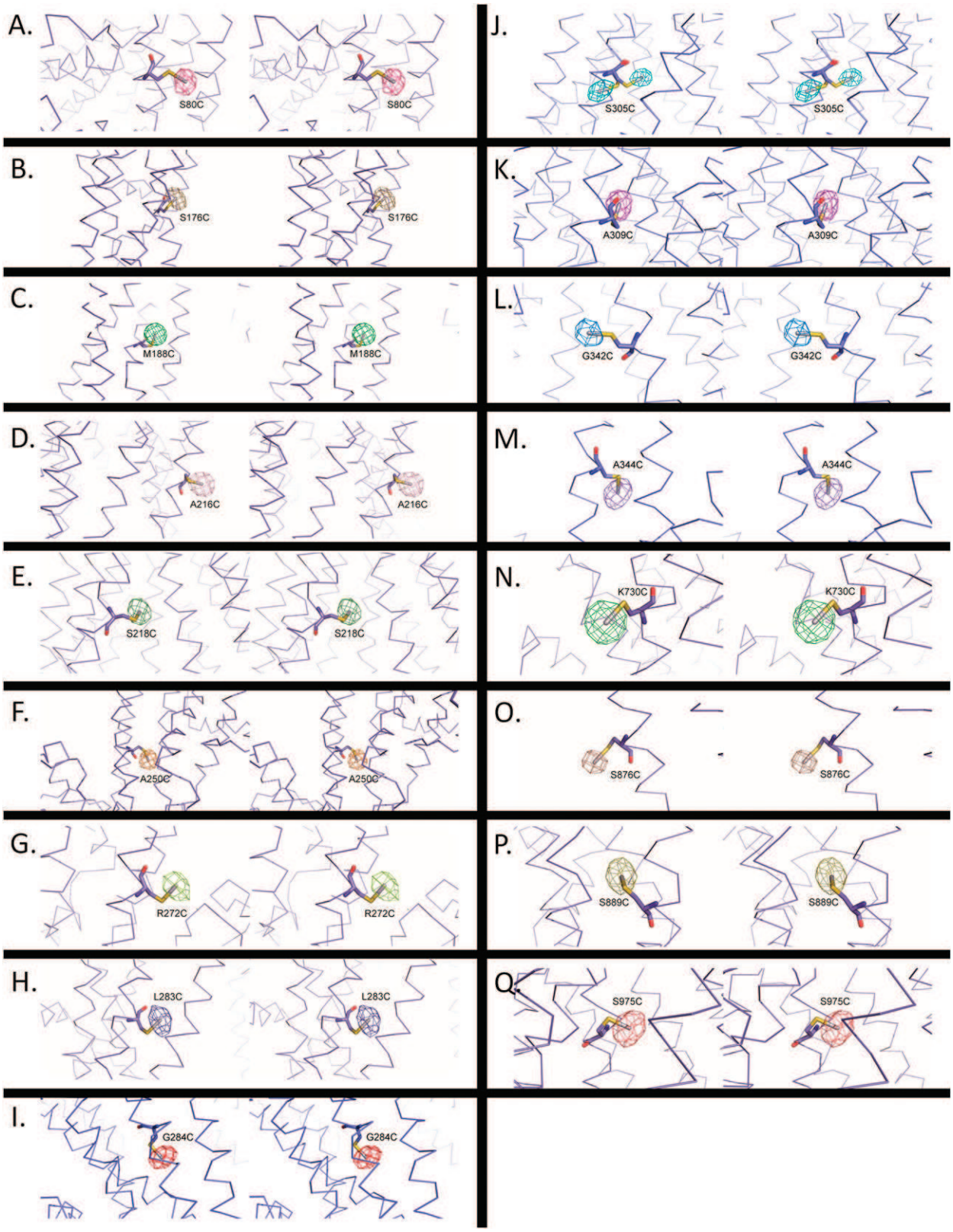


Fig. S2. Closeup views of mercury sites resolved from Hg-soaked crystals of cysteine substitution mutants used for topology verification (Fig. 1*B*). The relative position of the Hg site is shown validating both the position and identity of each residue. Fourier syntheses were generated as described. The mutation positions and sigma cutoffs of the difference maps are as follows: (*A*) S80C, 7σ, (*B*) S176C, 5σ; (*C*) M188C, 7σ; (*D*) A216C, 7σ; (*E*) S218C, 7σ; (*F*) A250C, 5σ; (*G*) R272C, 7σ; (*H*) L283C, 5σ; (*I*) G284C, 4σ; (*J*) S305C, 4σ; (*K*) A309C, 4σ; (*L*) G342C, 4σ; (*M*) A344C, 5σ, (*N*) K730C, 5σ; (*O*) S876C, 6σ; (*P*) S889C, 6σ; (*Q*) S975C, 8σ. Native cysteine sites include: C133, 7σ; C427, 3σ; C713, 7σ; C952, 7σ; C1070, 5σ; C1121, 7σ; C1223; 7σ. For residue 305 located on TM5 of the model, we observed two mercury peaks corresponding to two different rotamers of the mutated cysteine. Both of these positions are superimposed as shown in *J*. Interestingly, one position is on the inside of the molecule, whereas the other is outside, demonstrating that the P-gp may have some degree of local flexibility in structure.

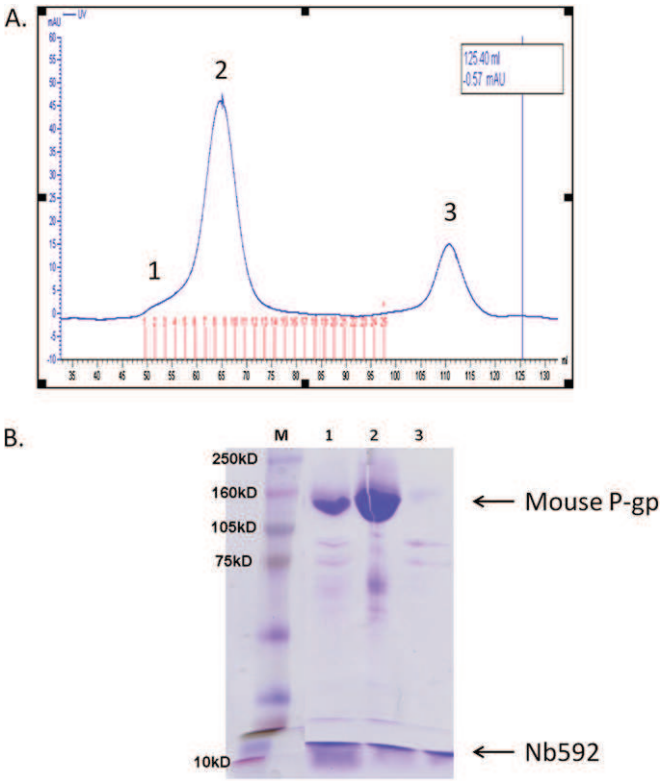


Fig. S3. Size-exclusion chromatography (SEC) plot of P-gp–Nb592 complex. (*A*) Protein sample was injected on a Superdex200 16/60 (GE Healthcare). The SEC column void volume is indicated by (1). The P-gp–Nb592 complex ran as a monodispersed peak eluting at 65 mL (2). The excessive Nb592 sample eluted at 111 mL (3). (*B*) The Coomassie stained SDS/PAGE gel of fractions at selected points shown corresponding to the region marked in the SEC plot.

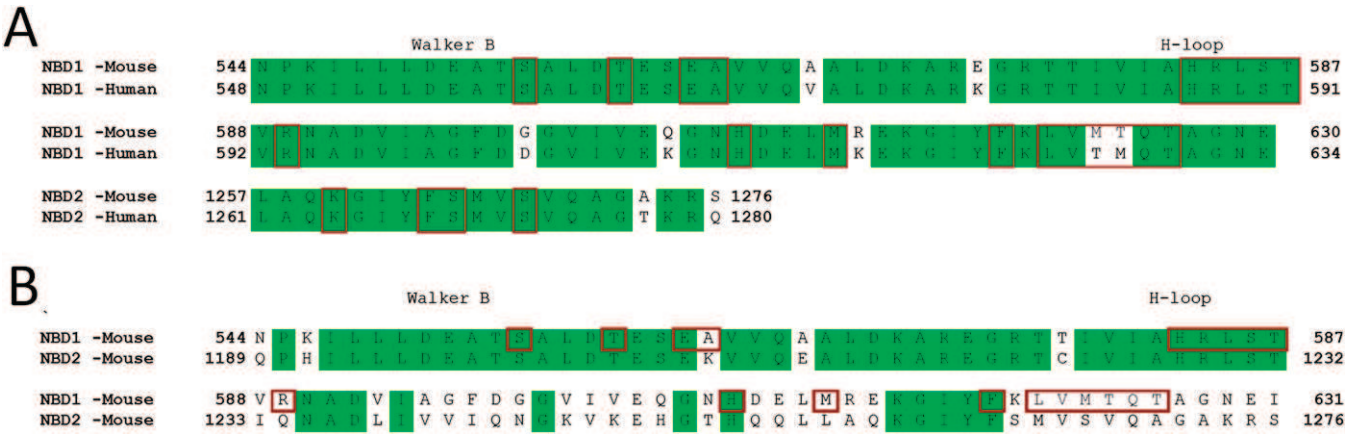


Fig. S4. (*A*) Protein sequence alignment between mouse and human P-gp showing only the regions bound by Nb592 CDR loops. Residues within 5 Å of Nb592 in the mouse P-gp structure, and the corresponding human residues, are boxed in red. Identical residues are indicated in green. (*B*) Protein sequence alignment of mouse NBD1 with mouse NBD2 in the same scheme as in *A*. Nb592 does not bind mouse NBD2.

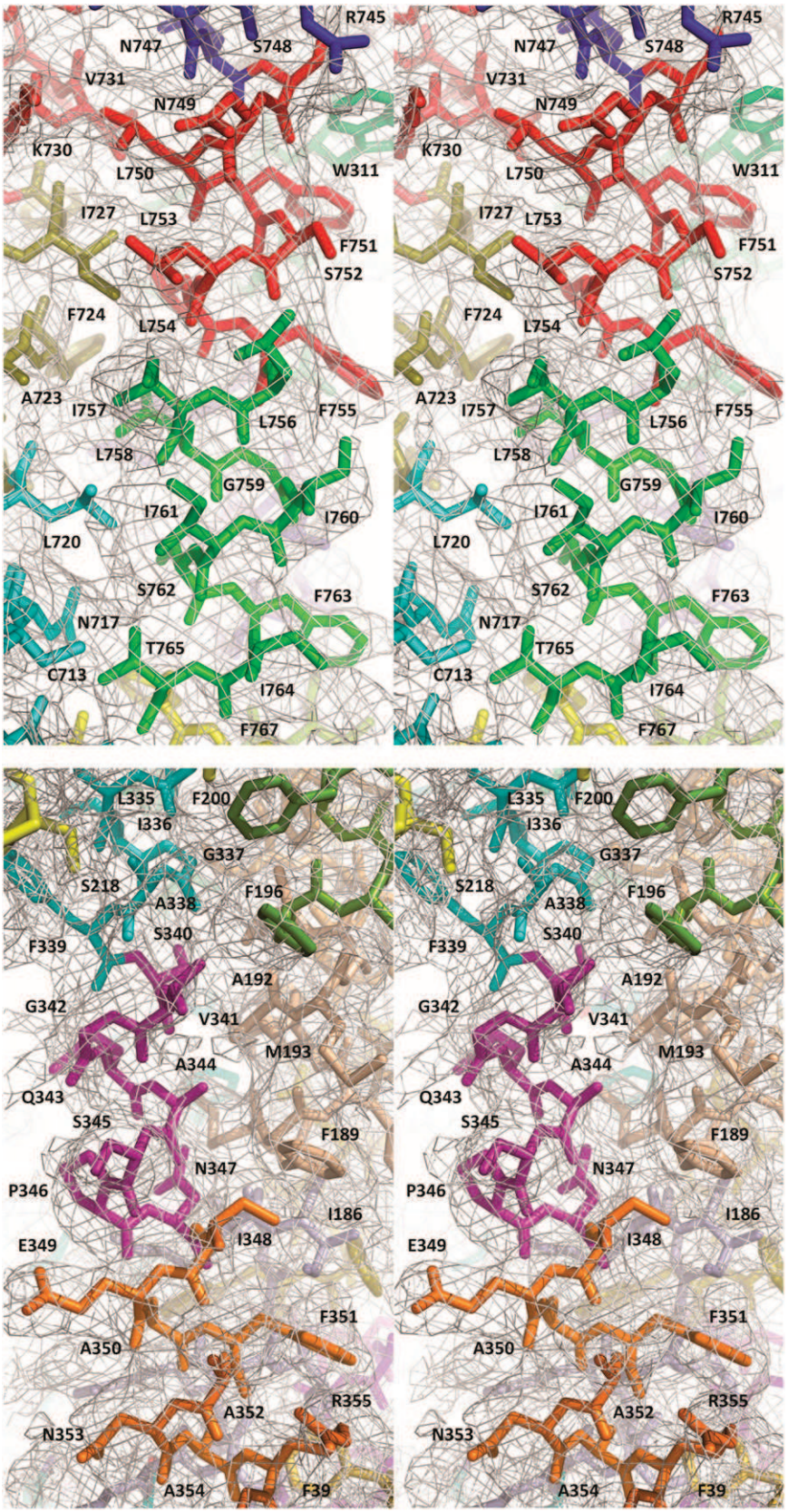


Fig. S5. Stereoviews of F_o-F_c simulated annealing omit map density for crystal 1 with multiple density maps superimposed. The F_o-F_c density maps (gray) were generated as described in the text. Difference density for the entire P-gp molecule was achieved by superimposing individual F_o-F_c maps calculated using sequential segments of nine residues that were iteratively omitted from the model calculated with a neighboring sphere size of 4 Å and a map cushion of 2 Å. The windows of residues omitted from the model are shown in different colors. Maps are contoured at 1.0 σ. The panels are side-by-side stereoviews of superimposed omit maps used for model validation. The panels illustrate example regions of difference maps matching the model. The amino acid residues of the regions shown are indicated.

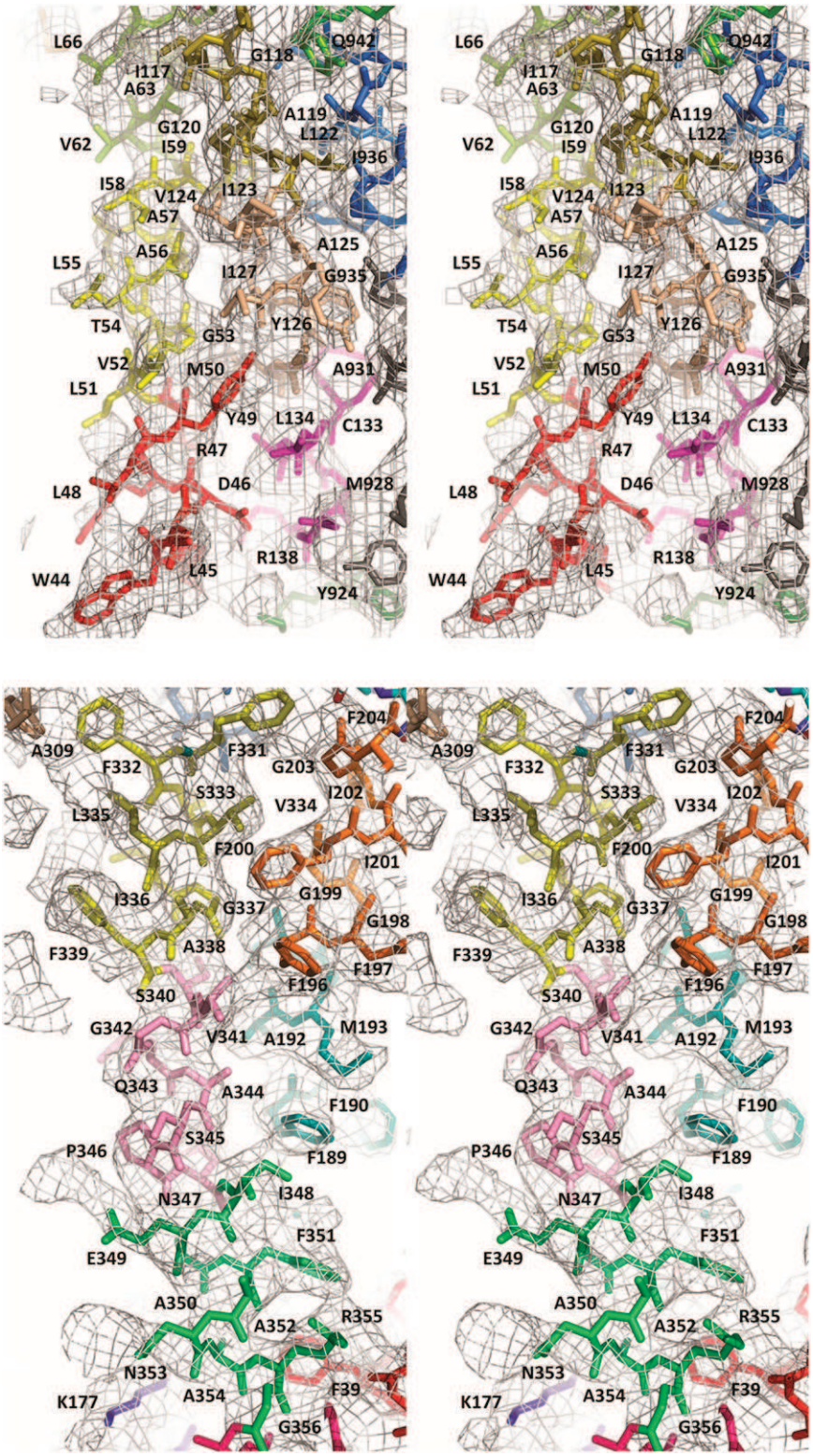


Fig. S6. Stereoviews of F_o-F_c simulated annealing omit map density for crystal 2 with multiple density maps superimposed. The F_o-F_c density maps (gray) were generated in Fig. S5. The windows of residues omitted from the model are shown in different colors. Maps were contoured at 1.0 σ. The panels are side-by-side stereoviews of superimposed omit maps used for model validation. The panels illustrate example regions of difference maps matching the model. The amino acid residues of the regions shown are indicated.

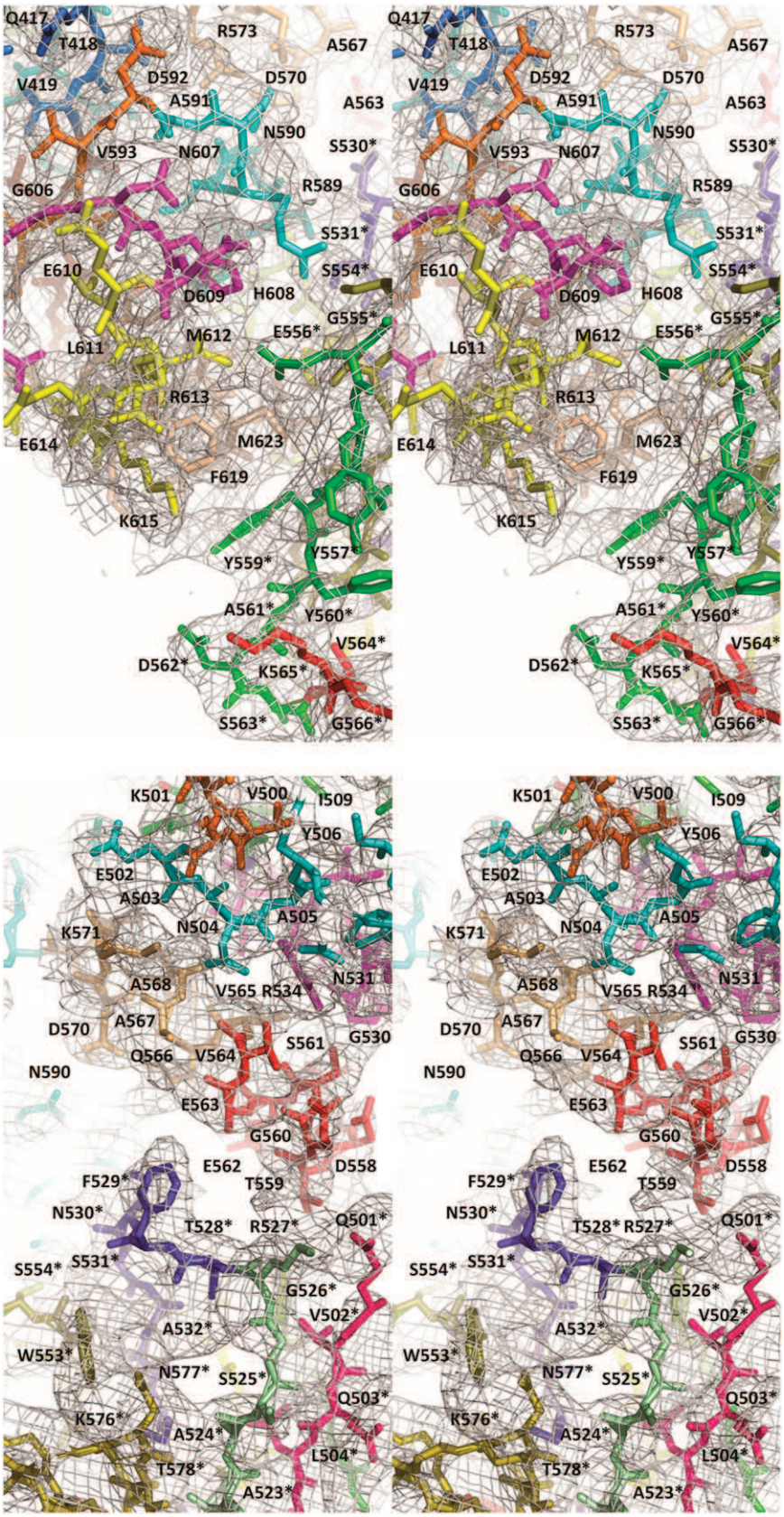


Fig. S7. Stereoviews of F_o-F_c simulated annealing omit map density for the P-gp–Nb592 complex with multiple density maps superimposed. The residues of Nb592 are marked with asterisk (*). The F_o-F_c density maps (gray) were generated as in Fig. S5. The windows of residues omitted from the model are shown in different colors. Maps were contoured at 1.0 σ. The panels are side-by-side stereoviews of superimposed omit maps used for model validation. The panels illustrate example regions of difference maps matching the model. The amino acid residues of the regions shown are indicated.

Table S1. Data collection, phasing, and refinement statistics

	Crystal form A		Crystal form B	Crystal form A	
	Crystal 1	Crystal 2	Crystal 3 (Nb-bound)	Experimental map crystal	
Data collection					
Space group	P212121	P212121	P212121	P212121	P212121
Unit cell	—	—	—	E1	E2
a, *b*, *c*, Å	*a* = 87.40, *b* = 138.65, *c* = 185.13	*a* = 90.65, *b* = 138.29, *c* = 194.72	*a* = 87.10, *b* = 102.47, *c* = 312.10	*a* = 87.92, *b* = 137.37, *c* = 181.44	*a* = 87.92, *b* = 137.41, *c* = 181.83
α, β, γ, °	α, β, γ = 90	α, β, γ = 90	α, β, γ = 90	α, β, γ = 90	α, β, γ = 90
Wavelength, Å	1.00511	0.97945	1.009	1.007	1.009
Resolution, Å	3.8	4	4.1	4.2	4.2
Redundancy	3.4 (3.6)	2.4 (2.3)	5.3 (5.3)	4.9 (4.9)	4.9 (4.9)
R_{sym},* %	0.066 (0.414)	0.038 (0.234)	0.072 (0.350)	0.076 (0.469)	0.076 (0.478)
Completeness, %	96.1 (98.4)	88.7 (91.9)	96.8 (97.9)	98.4 (99.2)	98.4 (99.2)
Model building and refinement					
Resolution, Å	3.8	4	4.1	—	—
Sigma cutoff	2.04	2.02	2.05	—	—
R value, † %	32.47	31.7	32.39	—	—
(Working + Test)				—	—
R value,† %	32.31	31.58	32.28	—	—
(Working Set)				—	—
Free *R* value‡	35.66	33.77	34.45	—	—
Completeness, %	95.29	86.9	95.7	—	—
rmsd					
Bond lengths, Å	0.004	0.004	0.005	—	—
Bond angles, °	0.869	0.909	1.072	—	—
Wilson B, Å^2	98.63	113.94	96.84	—	—
Shrinkage radius, Å	0.9	0.9	0.9	—	—
Solvent radius, Å	1.11	1.11	1.11	—	—
Ramachandran statistics (molprobity)					
Outliers, %	0	0	0	—	—
Allowed, %	8.83	8.91	9.89	—	—
Favored, %	91.17	91.09	90.11	—	—
Rotomer outliers	0	0	0	—	—
Cβ deviations	0	0	0	—	—

Values in parentheses are for the corresponding values in the highest resolution shell.

*$R_{sym} = S|I - \langle I \rangle|S \langle I \rangle$, where I is the measured intensity of each reflection, and $\langle I \rangle$ is the intensity averaged from symmetry equivalents.

†$R = S|F_o - F_c|/S|F_c|$, where F_o and F_c are observed and calculated structure factors, respectively.

‡R_{free} was calculated from a subset of data (5%) omitted from the refinement.

Table S2. Data collection statistics of single site mutations used for topological validation of P-gp model

		Unit cell, Å						
Dataset	Source	*a*	*b*	*c*	Resolution, Å	Multiplicity	R_{sym}*	Completeness
1: S80C	APS23ID-D	88.48	137.9	186.27	4.3	4.3	0.062	99.7
2: S176C	SSRL11-1	87.81	138.37	185.17	4.3	7.2	0.051	94.2
3: M188C	APS23ID-D	89.14	138.86	188.93	4.3	4.5	0.055	97.3
4: A216C	APS23ID-D	88.81	137.66	188.89	4.4	4.2	0.08	99.7
5: S218C	APS23ID-D	89.95	138.76	191.96	4.2	4.3	0.071	99.7
6: A250C	ALS5.0.1	86.04	137.79	184.13	4.7	4	0.078	95.7
7: R272C	SSRL11-1	88.82	140.02	188.99	4.15	3.5	0.105	95.7
8: L283C	SSRL11-1	88.08	138.68	185.9	4.8	3.6	0.076	92.3
9: G284C	SSRL11-1	88.71	137.89	188.31	4.2	4.3	0.062	83.1
10: S305C	ALS5.0.1	90.9	138.8	195.96	4.7	4.4	0.057	93.9
11: A309C	SSRL11-1	86.76	138.01	187.71	6	3.4	0.091	85.4
12: G342C	SSRL11-1	87.31	138	183.71	4.65	3.6	0.074	97.7
13: A344C	SSRL11-1	89.96	138.47	189.95	4.7	3.4	0.057	95.8
14: K730C	SSRL11-1	90.22	138.34	192.73	4.25	4.1	0.057	81.9
15: S876C	APS23ID-D	88.94	138.25	188.61	4.4	4.6	0.069	94.5
16: S889C	SSRL11-1	89.4	137.86	186.61	4.8	2.8	0.09	85.6
17: S975C	SSRL11-1	90.27	138.72	194.05	4.6	3.5	0.053	95.9

*$R_{sym} = S|I - \langle I \rangle| S \langle I \rangle$, where I is the measured intensity of each reflection, and $\langle I \rangle$ is the intensity averaged from symmetry equivalents.

Conclusion

As protein biological function depends on their 3D structure, resolving protein structure is of great interest in medicine and biotechnology. To date few crystal structures of ABC exporters are solved at low resolution (2.85-5.5 Å), substrate-free and in the inward-facing conformation. New structural "snapshots" will help to fill the lacks in the cycle of transport and explain the polyspecific drug recognition.

P-gp is a protein relatively stable, easy to purify and crystallize, with a yield of about 20 mg / fermenter of yeast culture. However, the resolution of these crystals is quite low (3.8-4.5 Å), which prompted us to perform other experiments, including tests of co-crystallization with substrates and inhibitors, supposed to stabilize the more flexible parts, facilitating crystallization and increasing the resolution. This series of tests resulted in crystals with worse diffraction resolution (about 6-7 Å), making it impossible to solve the 3D-structure and localize the compounds.

Related to this, P-gp-Linkerless was supposed to have better resolution that P-gp wild-type protein because the linker is a region in P-gp highly flexible causing disorder in the crystal. Although this was not the case, the P-gp-Linkerless gave resolutions lower (about 8-10 Å). This construction was used to trap P-gp in the outward-facing conformation by preventing either ATP hydrolysis (non-hydrolysable ATP-analogues) or phosphate release (ATP with magnesium and sodium orthovanadate). Despite the low resolution (10 Å), the condition with AMPPNP were in a different space group than WT P-gp, indicating a possible different conformation of P-gp. A further crystallization screening (other PEG, salts, conditions) of P-gp-Linkerless may allow to get a new conformation of P-gp.

To overcome the resolution limitation and improve the accuracy of the mouse P-gp structure the incorporation of heavy atoms (mercury) into the protein allowed to locate the label-atoms unambiguously. Thus, 24 position were validated improving the P-gp model. Three new inward-facing conformation of P-gp were published and may be used *in silico* to understand the transport cycle and to better predict substrates and inhibitor for P-gp. Nevertheless, further x-ray crystallography of P-gp in complex with substrates are needed to fill the lack of experimental information at the molecular level relative to the location of drugs transported.

Up to now, 10 structures of P-gp are known: three from the mouse in 2009, one from *C. elegans*, three from mouse in 2013 and the three 2009 corrected. All in an inward-facing conformation with low resolution (3.4 – 4.4 Å) and differences between them. The impact of these differences has to be analyzed carefully. The global change is that the new mouse structures are similar to the *C. elegans* structure. In the future a structure of P-gp with high resolution (<< 3 Å) will provide an accurate model fitting functional and biochemical data on P-gp.

CHAPTER III: DECIPHERING THE POLYSPECIFICITY OF P-GP

Context

Several investigations allowed to identify different drug-binding sites, among them the H-, R- and P- sites have been the most studies. The H site binds the Hoechst 33342 (and related drugs), the R site binds anthracyclins such as daunorubicin and the P site binds the likes of prazosin and progesterone. Up to date, only two enantiomeric inhibitors, QZ59-SSS and QZ59-RRR, were co-crystallized with mouse P-gp [166], raising the question of their distinct or shared location with drug-substrates binding sites. Thus, the aim of this study was to use the structural information brought with QZ59 enantiomers to try to locate the P-, H- and R- drug-substrate binding sites.

Transport experiments were designed with whole cells, for ensuring a fully functional environment for the pump. Briefly, drug efflux was measured by flux cytometry on NIH3T3 cells stably transfected with an expression plasmid either empty or encoding the human P-gp. Cells were incubated either with or without substrate (concentration range 0 at 6 µM) and inhibitors (concentration range 0 at 4 µM). After incubation and washing, the intracellular drug amount was estimated. In collaboration with Dr. Emilie Henin and Pr. M. Tod (Faculté de Médecine Lyon-sud Charles Mérieux, Lyon), each data sets were fitted with either classical enzymology or built inhibition models. They were evaluated by using the Goodness-of-fit (GOF) and the Akaike's Information Criterion, AIC (as detailed in Methods and summarized in the table of supporting information of publication III).

The conditions presented here were set up to compare drugs accumulation depending on the presence or absence of QZ59s inhibitors, leading to apparent value parameters as indicated always labelling them “app”. This is classically done when one cannot establish true values of kinetic parameters (K_m, K_i, …). Determining true kinetic parameters, and especially the initial rate of drug efflux solely mediated by P-gp, is much more complicated and was (hopefully) not necessary to establish the inhibition mechanism of QZ59s inhibitors.

The case of bodipy-prazosin

The bodipy-prazosin (BP) was used as an analog of prazosin [223], for exploring the effects of QZ59-RRR or QZ59-SSS on drug binding to the P-site. The contribution of P-gp to the BP efflux is deduced by subtracting the accumulation of BP in control NIH3T3 cells (maximal drug accumulation, triangles in Figure 70A) from the P-gp–expressing NIH3T3 cell (minimal drug accumulation, squares in Figure 70A). The P-gp-mediated efflux of BP displayed a complex pattern, being maximal up to 1 µM and then decreasing above (squares in Figure 70A). This behavior is typical of an inhibition caused by an excess of substrate. Data regression with the corresponding equation (e.q. 2.6 in Table 1 of publication III: $V_m S/(K_m+S\,(1+S/K_{SI}))$) gave a $V_{m,app}$ of 60 ± 9 pmolmg^{-1}s^{-1}, a $K_{m,app}$ of 1.1 ± 0.2 µM and a substrate inhibition constant $K_{si,app}$ of 0.6 ± 0.1 µM.

In order to be sure that the effect of BP was an inhibition by the substrate, the behavior of prazosin towards its bodipy analog was checked expecting to be competitive. Surprisingly, such

effect was not observed, the efflux of bodipy-prazosin added at 1 µM being unchanged up to 30 µM prazosin (Figure 70B). This suggests that bodipy-prazosin does not behave like prazosin, and the fluorescent analog was not further used in this study.

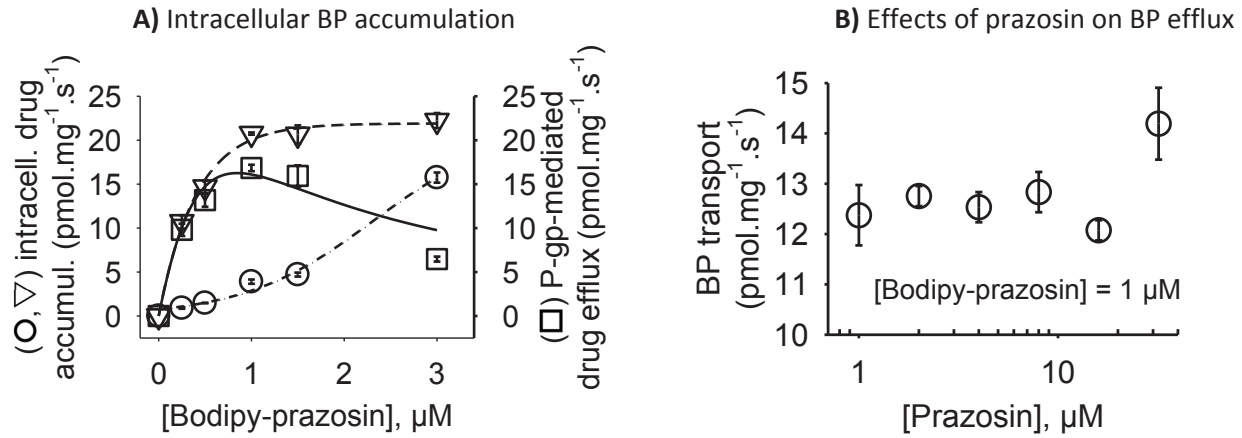


Figure 70. Intracellular bodipy-prazosin accumulation and P-gp-mediated transport.

A) BP accumulations either in NIH3T3 (triangles) and NIH3T3-P-gp (circles). BP transport mediated by P-gp is the difference between both accumulations (squares).

B) BP transport rate as a function of prazosin (circles).

The case of Hoechst 33342 and daunorubicin

Same kind of experiments were carried out toward the Hoechst 33342 and daunorubicin. Results revealed distinct inhibitory mechanisms of the RRR or SSS QZ59 enantiomers toward drug-substrates binding to the H and R sites. The QZ59-RRR displayed a mixed (with a marked noncompetitive tendency) inhibition of Hoechst 33342 transport and a non-competitive inhibition of daunorubicin transport (Figure 2 in the publication III). Thus, QZ59-RRR inhibited drug efflux mainly through an inhibitory site distinct to the R- and H- transport sites. However, at higher concentrations (≥ 5 µM), it tends to share the H- site and enhanced the daunorubicin-mediated self-inhibition. The QZ59-SSS was found to compete with both substrates (Figure 3 in the publication III) with a tendency to activate the Hoechst 33342 efflux (at high QZ59-SSS and Hoechst 33342 concentrations). As the RRR enantiomer, QZ59-SSS also enhanced the self-inhibitory action of daunorubicin efflux. Thus the QZ59-SSS shares both sites with a positional substrate overlap much more marked for the H-site than for the R-site ($K_{I,app}$ values of 0.15 and 0.3 µM towards Hoechst 33342 and daunorubicin transport respectively).

At first thought, these results are conflicting because QZ59 enantiomers are located in a common groove in the outer leaflet of the membrane domain [166] where the R- and H-drugs-binding sites are distinct [227]. However, the contradiction is only apparent. Looking closer to the kinetic constants, QZ59-RRR displayed a moderate competitive tendency toward Hoechst 33342 ($K_{i1,app}$ = 5 µM > $K_{i2,app}$ = 1.6 µM) while it acted strictly non-competitively toward daunorubicin.

Docking experiments suggest that the Hoechst 33342 covered both QZR59-SSS locations while the daunorubicin shared that of the most embedded QZ59-SSS (Figure 71).

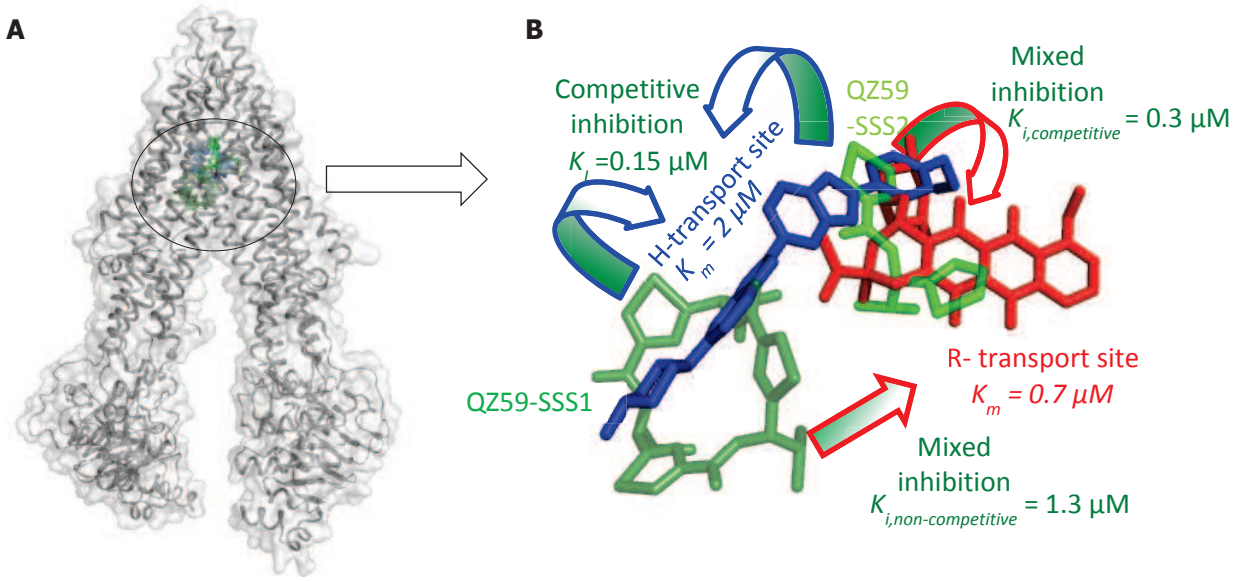


Figure 71. Molecular localization of Hoechst 33342 and daunorubicin in P-gp
A) 3D structure of P-gp (pdb code: 3G61) with the two enantiomers of the cyclic hexapeptide QZ59 inhibitor. The QZ59 SSS is show in green sticks (upper) and dark green (lower), the QZ59- is show in sticks light blue. **B)** Location of sites. The colors for the cyclic hexapeptide QZ59 inhibitor is the same as in A. Hoechst 33342 (blue) and daunorubicin (red) are shown in the best-rated molecule location from independent docking simulations. The arrows show the effect of inhibition (with their kinetic constants) each QZ59 Site transport Hoechst 33342 and daunorubicin.

The conformational trajectory from the 2009 P-gp structures (PDB codes: 3G61, 3G5U) up to the mouse 2013 structures (4KSB, 4KSC), shows that the R and H drug-binding site locations that we propose exists at a given time, a situation which seems reasonable in a context of drug translocation (Video S1 in supporting information of publication III).

These results are published in FEBS journal and were presented in the following conferences:

- French-Belgian annual ABC meeting 2013 edition (Octobre 2013, Lyon, France). Communication: Deciphering the polyspecificity of human multidrug resistance P-glycoprotein. Martinez L.

- Gordon conferences, Multi-Drug Efflux Systems (Mars 2013, Ventura, CA). Poster presentation: Binding sites for Hoechst 33342 and QZ59-SSS overlap in human ABCB1. Martinez, L.; Arnaud, O. ; Henin, E. ; Tao, H. ; Chaptal, V. ; Tod, M. ; Di Pietro, A. ; Zhang, Q.; Chang, G. ; and Falson, P.

- 4th FEBS Special Meetingon ABC Proteins, FEBS « Federation of European Biochemical Societies » (March 2012, Innsbruck, Austria). Poster presentation: P-glycoprotein inhibition mechanism of QZ59. Martinez L., Arnaud O., Henin E., Tao H., Chaptal V., Andrieu T., Dussurgey S., Tod M., Di Pietro A., Zhang Q., Chang G., and Falson P.

Publication III

the FEBS Journal

Understanding polyspecificity within the substrate-binding cavity of the human multidrug resistance P-glycoprotein

Lorena Martinez[1],†, Ophélie Arnaud[1],†, Emilie Henin[2],†, Houchao Tao[3], Vincent Chaptal[1], Rupak Doshi[4], Thibault Andrieu[5], Sébastien Dussurgey[5], Michel Tod[2], Attilio Di Pietro[1], Qinghai Zhang[3], Geoffrey Chang[4] and Pierre Falson[1]

1 Drug Resistance Mechanism and Modulation group, Ligue 2013 certified, Molecular and Structural Basis of Infectious Systems, Mixed Research Unit between the National Centre for Scientific Research and Lyon I University n°5086, Institute of Biology and Chemistry of Proteins, France
2 Ciblage Thérapeutique en Oncologie, Mixed Research Team n°3738, Faculté de Médecine Lyon-sud Charles Mérieux, OULLINS Cedex, France
3 Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA, USA
4 Skaggs School of Pharmacy and Pharmaceutical Sciences, Department of Pharmacology, School of Medicine University of California, San Diego, La Jolla, CA, USA
5 Unité Mixe de Service n°3444, Tour Inserm Cervi, Lyon cedex, France

Keywords
ABC transporters; cancer; drug efflux; drug resistance; drug-binding sites; P-glycoprotein

Correspondence
Pierre Falson, Drug Resistance Mechanism and Modulation group, Molecular and Structural Basis of Infectious Systems, Mixed Research Unit between the National Centre for Scientific Research and Lyon I University n°5086, Institute of Biology and Chemistry of Proteins, 7, passage du Vercors, 69367 Lyon, France
Fax: +33 4 7272 2604
Tel: + 33 6 4658 0266
E-mail: pierre.falson@ibcp.fr

†These authors contributed equally to this work

(Received 10 September 2013, revised 6 November 2013, accepted 7 November 2013)

doi:10.1111/febs.12613

Human P-glycoprotein (P-gp) controls drugs bioavailability by pumping structurally unrelated drugs out of cells. The X-ray structure of the mouse P-gp ortholog has been solved, with two *SSS* enantiomers or one *RRR* enantiomer of the selenohexapeptide inhibitor QZ59, found within the putative drug-binding pocket (Aller SG, Yu J, Ward A, Weng Y, Chittaboina S, Zhuo R, Harrell PM, Trinh YT, Zhang Q, Urbatsch IL *et al.* (2009). Science 323, 1718–1722). This offered the first opportunity to localize the well-known H and R drug-binding sites with respect to the QZ59 inhibition mechanisms of Hoechst 33342 and daunorubicin transports, characterized here *in cellulo*. We found that QZ59-*SSS* competes efficiently with both substrates, with $K_{I,app}$ values of 0.15 and 0.3 μM, which are 13 and 2 times lower, respectively, than the corresponding $K_{m,app}$ values. In contrast, QZ59-*RRR* non-competitively inhibited daunorubicin transport with moderate efficacy ($K_{I,app}$ = 1.9 μM); it also displayed a mixed-type inhibition of the Hoechst 33342 transport, resulting from a main non-competitive tendency ($K_{i2,app}$ = 1.6 μM) and a limited competitive tendency ($K_{i1,app}$ = 5 μM). These results suggest a positional overlap of QZ59 and drugs binding sites: full for the *SSS* enantiomer and partial for the *RRR* enantiomer. Crystal structure analysis suggests that the H site overlaps both QZ59-*SSS* locations while the R site overlaps the most embedded location.

Introduction

The human P-glycoprotein (P-gp) is an integral membrane protein that actively pumps endo/exogenous compounds out of cells using energy derived from ATP binding and hydrolysis [1]. The expression of

Abbreviations
ABC, ATP-binding cassette; AIC, AKAIKE Index Criterion; AICc, corrected AIC; GOF, Goodness-Of-Fit; MTS-rhodamine, methanethiosulfonate-rhodamine; P-gp, Pleiotropic glycoprotein; TM, Transmembrane.

P-gp is up-regulated in many cancer cells, where it reduces the intra-cellular concentrations of many chemotherapeutic drugs, thereby conferring multidrug resistance (MDR) [2]. P-gp is one of the best-known pumps in this context, in addition to multidrug resistance protein 1 (MRP1/ABCC1) and breast cancer resistance protein (BCRP/ABCG2); all these proteins belong to the same ATP-binding cassette (ABC) transporter superfamily [3]. The involvement of such pumps in chemoresistance requires elucidation of the mechanisms of multidrug export and a targeted inhibition. P-gp transports a broad spectrum of molecules sharing a marked hydrophobicity but structurally divergent [4,5]. This polyspecificity is partly explained by the existence of at least three predicted drug-binding sites, termed H-, R-, and P- site. The H site binds Hoechst 33342 and quercetin, the R site preferentially binds rhodamine 123 and anthracyclins such as daunorubicin and doxorubicin, while the P site binds compounds such as prazosin and progesterone [6,7]. A fourth site has also been suggested, which may bind non-transported modulators like GF120918 or nicardipine [8].

Of the various drug-binding sites of P-gp, the H and R sites have been more rigorously investigated. The use of thiol-reactive substrates in competition with drugs, coupled with cysteine-scanning mutagenesis, has contributed to the identification of several amino acid residues from transmembrane domain that are close to or participate in the drug binding sites (reviewed in [9]). Notably, I340 (in transmembrane domain (TM) 6), A841 (in TM9), and L975/V981/V982 (in TM12) map to the R site [10]. The location of the H site is more controversial as fluorescence resonance energy transfer investigations using Hoechst 33342 suggested a location in the cytoplasmic/inner leaflet side of the transporter [11], but use of a pharmacophore for the same dye suggested a location within the outer leaflet [12].

The recent mouse P-gp (*Abcb*1a gene) X-ray structures reveal a large, internal drug-binding cavity enclosed between the two 'halves' of the 'full' transporter [13–15]. The pump was co-crystallized with the *RRR* and *SSS* enantiomers of the cyclic selenohexapeptide, QZ59 [13]. One molecule of QZ59-*RRR*, and two of QZ59-*SSS* were identified in their respective co-crystal structures. As is not uncommon for membrane proteins, these 3D-structures were obtained at a rather modest resolution of 3.8–4.35 Å, but, due to the three selenium atoms they included, QZ59s may be unambiguously located within the outer leaflet interface of the drug-binding cavity. QZ59s have been found to inhibit mouse P-gp-mediated efflux of drugs such as colchicine and calcein, and the verapamil-stimulation component of mouse P-gp's ATP hydrolysis activity [13]. Until now, no other compound, as either inhibitor or substrate drug, has been co-crystallized with P-gp, so the QZ59 enantiomers are unique tools for localizing the drug-binding sites more precisely, at least inside the present co-crystallized pump conformation. A detailed characterization of their inhibition mechanisms (competitive, non-competitive, mixed, etc.) would indicate whether they share, none, one or even several drug transport sites, which ones, and to what extent. Such information is crucial to understanding the transport mechanism mediated by ABC pumps; it is also critical for guiding the design of compounds that act at the molecular level with the highest efficiency.

We addressed this question in the present study. Our results revealed distinct inhibitory mechanisms of Hoechst 33342 and daunorubicin effluxes by each QZ59 enantiomer. As expected, determination of such mechanisms suggests localization of these substrate-binding sites within the transporter's binding cavity.

Results

In order to test the effects of QZ59 inhibitors on the transport by P-gp of substrates that binding to the H- and R-sites, as detailed in the Experimental procedures, we used whole cells to provide a native environment for the pump in terms of membrane lipid composition, substrate partitioning and bioavailability, protein trafficking, and membrane potential [16]. The concentrations and kinetic parameters estimated under these conditions are only 'apparent', but, from the point of view of the present study, are fully comparable.

We measured the transport of daunorubicin and Hoechst 33342 mediated by P-gp (Fig. 1), from which we estimated maximal rates, V_m, of 41 ± 3 and 34 ± 3 pmol·s^{-1}·mg^{-1} (of protein cell content) and Michaelis constants, K_m of 0.7 ± 0.2 and 2 ± 0.4 μM for daunorubicin and Hoechst 33342, respectively, using Eqn 2.1 (Table 1). These values are in the same range as those previously reported for daunorubicin [16] and Hoechst 33342 [17] showing that the system used here allows characterization of the QZ59 inhibition towards these substrates. For further experiments, the data sets were subjected to a detailed enzymatic analysis (Table S1) using non-linear regressions, and modeled using the equations shown in Table 1. We then evaluated the models by statistical tools such as Goodness-Of-Fit (GOF) and the corrected Akaike Information Criterion (AIC_c). As classically performed, the choice of the best model is guided by the largest decrease in AICc (ΔAIC_c) when using the Michaelis-Menten equation as reference (Eqn 2.1).

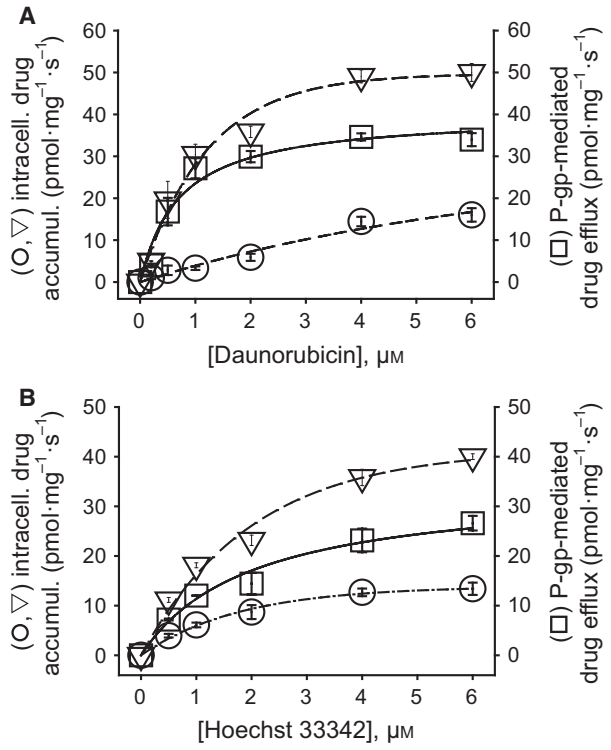


Fig. 1. Intracellular accumulation and P-gp-mediated transport of daunorubicin and Hoechst 33342 in control and P-gp-expressing NIH3T3 cells. Daunorubicin (A) and Hoechst 33342 (B) were added to NIH3T3 cells (open triangles) and NIH3T3/P-gp cells (open circles) and quantified as described in Experimental procedures. The net drug or dye accumulations corresponding to the difference in accumulation between the cell types are indicated by open squares. Data in triplicates were fitted as described in Experimental procedures using Eqn 1.1 (triangles and circles) and Eqn 2.1 (squares).

Mechanism of drug efflux inhibition by QZ59-*RRR*

We first measured the QZ59-*RRR*-influenced inhibition of Hoechst 33342 efflux (to assess the H site) and daunorubicin efflux (to assess the R site) by human P-gp. Transport data in the absence or presence of increasing concentrations of the drug-substrate and QZ59-*RRR* were collected using the cell line NIH3T3/P-gp which over-express human P-gp, using non-expressing NIH3T3 cells as a control (Fig. 2A–F).

The mixed inhibition model (Eqn 6.1) best fitted the dataset for inhibition of Hoechst 33342 export (Fig. 2A–C), leading to the largest ΔAICc (174.8, Fig. 2B). A detailed analysis of model fitting and scoring is shown in Fig. S1A. This mixed inhibition has a strong non-competitive contribution, as suggested by the inhibition constants $K_{i1,app}$ (for P-gp alone) of 5 ± 2 μM and $K_{i2,app}$ (for the P-gp/Hoechst 33342 complex) of 1.6 ± 0.2 μM (Fig. 2C and Table S1). The non-competitive inhibition model (Eqn 4.1) confirmed this tendency, giving a comparable ΔAICc (171.7, Fig. 2B). As shown in Fig. 2C, this non-competitive mixed inhibition suggests that, at the $K_{i2,app}$ concentration (1.6 μM) QZ59-*RRR* binds to an inhibitory site, distinct from the H site, while at high concentrations, i.e. in the range of $K_{i1,app}$ (≥5 μM), QZ59-*RRR* shares the H site.

The inhibition of P-gp-mediated daunorubicin efflux by QZ59-*RRR* is shown in Fig. 2D–F and Fig. S1B. The initial kinetic analysis suggested a non-competitive model (Table S1). Additionally, we observed inhibition mediated by daunorubicin itself of its own efflux, which is clearly visible at high daunorubicin and QZ59-*RRR* concentrations (Fig. 2D). Regression analyses performed on these data with Eqns 4.6 and 4.7, which include these effects, gave the same ΔAICc scores (191.1) for the competitive model (Eqn 4.6) and the non-competitive model (Eqn 4.7) (Fig. 2E). However, these models did not correctly fit the data obtained at high daunorubicin and QZ59-*RRR* concentrations (Fig. 2E and Fig. S1B). These data suggest that QZ59-*RRR* may enhance daunorubicin-mediated self-inhibition. Indeed, introducing this effect into Eqns 4.8 and 4.9 led to the largest ΔAIC_c, 225.2 and 227.4, for models including competitive and non-competitive inhibition by daunorubicin (Fig. 2E). Although these models are fairly close in terms of their fit (Fig. S1B), the latter model (Eqn 4.9), with the significantly larger ΔAIC_c, is more likely to be correct. Thus, the QZ59-*RRR*-mediated inhibition of daunorubicin transport is non-competitive ($K_{I,app} = 1.9 \pm 0.4$ μM), and coupled with daunorubicin-mediated self-inhibition enhanced by QZ59-*RRR* ($K_{SI,app} = 7 \pm 3$ μM).

Taken together, our data suggest that QZ59-*RRR* inhibits drug efflux mainly through an inhibitory site, distinct to the R- and H- transport sites. At high concentrations however, QZ59-*RRR* partially shares or overlaps with the H-site, and also enhances the self-inhibitory action of R-site drug-substrates.

Mechanism of drug efflux inhibition by QZ59-*SSS*

The mouse P-gp/QZ59 co-crystal structures reveal distinct binding sites for the *RRR* and *SSS* enantiomers of QZ59, with the latter being present in two different locations. We tested whether this distinction has a bearing on the mechanism of drug efflux inhibition.

The inhibition of P-gp-mediated Hoechst 33342 efflux by QZ59-*SSS* is shown in Fig. 3A–C. The data at different concentrations of QZ59-*SSS* revealed a sigmoidal distribution, suggesting cooperative behavior

Table 1. Enzyme inhibition models used in this study.

Eqn	Models	Velocity equations
2.1	No inhibition (NI)	$V_m*S/(K_m + S)$
3.1	QZ59 competitive inhibition (CI)	$V_m*S/(K_m*(1 + I/K_I) + S)$
4.1	QZ59 non-competitive inhibition (NCI)	$V_m/(1 + I/K_I)*S/(K_m + S)$
5.1	QZ59 uncompetitive inhibition (UI)	$V_m/(1 + I/K_I)*S/(K_m/(1 + I/K_I) + S)$
6.1	QZ59 mixed inhibition (MI)	$V_m/(1 + I/K_{i2})*S/(K_m*(1 + I/K_{i1})/(1 + I/K_{i2}) + S)$
2.2	Cooperativity (coop.) + NI	$V_m*S^h/(K_m{}^h + S^h)$
3.2	Coop. + QZ59 CI	$V_m*S^h/(K_m{}^h*(1 + I/K_I) + S^h)$
4.2	Coop. + QZ59 NCI	$V_m/(1 + I/K_I)*S^h/(K_m{}^h + S^h)$
5.2	Coop. + QZ59 UI	$V_m/(1 + I/K_I)*S^h/(K_m{}^h/(1 + I/K_I) + S^h)$
6.2	Coop. + QZ59 MI	$V_m/(1 + I/K_{i2})*S^h/(K_m{}^h*(1 + I/K_{i1})/(1 + I/K_{i2}) + S^h)$
3.3	QZ59 CI and activation (act.)	$V_m*(1 + I/K_A)*S/(K_m*(1 + I/K_I)/(1 + I/K_A) + S)$
4.3	QZ59 NCI and act.	$V_m*(1 + I/K_A)/(1 + I/K_I) *S/(K_m/(1 + I/K_A) + S)$
3.4	Coop. + QZ59 CI and act.	$V_m*(1 + I/K_A)*S^h/(K_m{}^h*(1 + I/K_I)/(1 + I/K_A) + S^h)$
3.5	Coop. + QZ59 CI and act. + QZ59 effect on coop.	$V_m*(1 + I/K_A)*S^{h*(1+I/K_A)}/(K_m{}^h{}^{*(1+I/K_A)}*(1 + I/K_I)/(1 + I/K_A) + S^{h*(1+I/K_A)})$
2.6	substrate (sub.) CI	$V_m*S/(K_m + S*(1 + S/K_{SI}))$
3.6	Sub. CI + QZ59 CI	$V_m*S/(K_m*(1 + I/K_I) + S*(1 + S/K_{SI}))$
4.6	Sub. CI + QZ59 NCI	$V_m/(1 + I/K_I)*S/(K_m + S*(1 + S/K_{SI}))$
5.6	Sub. CI + QZ59 UI	$V_m/(1 + I/K_I)*S/(K_m/(1 + I/K_I) + S*(1 + S/K_{SI}))$
6.6	Sub. CI + QZ59 MI	$V_m/(1 + I/K_{i2})*S/(K_m*(1 + I/K_{i1})/(1 + I/K_{i2}) + S*(1 + S/K_{SI}))$
2.7	Sub. NCI	$V_m/(1 + S/K_{SI})*S/(K_m + S)$
3.7	Sub. NCI + QZ59 CI	$V_m/(1 + S/K_{SI})*S/(K_m*(1 + I/K_I) + S)$
4.7	Sub. NCI + QZ59 NCI	$V_m/(1 + S/K_{SI})/(1 + I/K_I)*S/(K_m + S)$
5.7	Sub. NCI + QZ59 UI	$V_m/(1 + S/K_{SI})/(1 + I/K_I)*S/(K_m/(1 + I/K_I) + S)$
6.7	Sub. NCI + QZ59 MI	$V_m/(1 + S/K_{SI})/(1 + I/K_{i2})*S/(K_m*(1 + I/K_{i1})/(1 + I/K_{i2}) + S)$
2.8	Sub. CI + QZ59 effect on sub. CI	$V_m*S/(K_m+S*(1 + S/K_{SI}*I/K_I))$
3.8	Sub. CI + QZ59 CI + QZ59 effect on sub. CI	$V_m*S/(K_m*(1 + I/K_I)+S*(1 + S/K_{SI}*I/K_I))$
4.8	Sub. CI + QZ59 NCI + QZ59 effect on sub. CI	$V_m/(1 + I/K_I)*S/(K_m+S*(1 + S/K_{SI}*I/K_I))$
5.8	Sub. CI + QZ59 UI + QZ59 effect on sub. CI	$V_m/(1 + I/K_I)*S/(K_m/(1 + I/K_I)+S*(1 + S/K_{SI}*I/K_I))$
6.8	Sub. CI + QZ59 MI + QZ59 effect on sub. CI	$V_m/(1 + I/K_{i2})*S/(K_m*(1 + I/K_{i1})/(1 + I/K_{i2})+S*(1 + S/K_{SI}*I/K_{i2}))$
2.9	Sub. NCI + QZ59 effect on sub. NCI	$V_m/(1 + S/K_{SI}*I/K_I)*S/(K_m + S)$
3.9	Sub. NCI + QZ59 CI + QZ59 effect on sub. NCI	$V_m/(1 + S/K_{SI}*I/K_I)*S/(K_m*(1 + I/K_I) + S)$
4.9	Sub. NCI + QZ59 NCI + QZ59 effect on sub. NCI	$V_m/(1 + S/K_{SI}*I/K_I)/(1 + I/K_I)*S/(K_m + S)$
5.9	Sub. NCI + QZ59 UI + QZ59 effect on sub. NCI	$V_m/(1 + S/K_{SI}*I/K_I)/(1 + I/K_I)*S/(K_m/(1 + I/K_I) + S)$
6.9	Sub. NCI + QZ59 MI + QZ59 effect on sub. NCI	$V_m/(1 + S/K_{SI}*I/K_{i2})/(1 + I/K_{i2})*S/(K_m*(1 + I/K_{i1})/(1 + I/K_{i2}) + S)$

Equations were used as detailed in the Experimental procedures. S = substrate concentration (μM); I = QZ59 concentration (μM); V_m = maximal substrate efflux rate (fmol transported drug·cell^{-1}·h^{-1}; K_m = Michaelis constant (μM); K_I = inhibition constant (μM); K_{i1} = inhibition constant for the $E + I \rightarrow E{\cdot}I$ partial reaction (μM); K_{i2} = inhibition constant for the $E{\cdot}S + I \rightarrow E{\cdot}S{\cdot}I$ partial reaction (μM); h = Hill number, K_A = activation constant (μM); K_{SI} = substrate inhibition constant (μM). Eqn, equation; NI, no inhibition; CI, competitive inhibition; NCI, non-competitive inhibition; UI, uncompetitive inhibition; MI, mixed inhibition, coop., cooperativity; act., activation; sub, substrate.

of the pump triggered by QZ59-*SSS* binding. Moreover, the results obtained at higher inhibitor concentrations revealed an activation effect by QZ59-*SSS* on the efflux of Hoechst 33342. We performed simulations accordingly using Eqns 2.1–3.5, exploring the effects of cooperativity, inhibition and activation, independently or together. A detailed analysis is shown in Fig. S2A. Among these models, the competitive inhibition models best fitted the data set (Eqn 3.1–3.5, Fig. 3B). Furthermore, the models that included activation and cooperativity together (Eqns 3.4 and 3.5) led to the largest ΔAIC_c (95.4 and 105.7), in contrast to the models that included them separately (Eqns. 3.2 and 3.3: values of 47.9 and 79.9). The best model (Eqn 3.5; GOF shown in Fig. S2A) suggests that QZ59-*SSS* competes with Hoechst 33342 transport (Fig. 3C), with very good efficiency as indicated by the low value of the inhibition constant, $K_{I,app}$ (0.15 ± 0.04 μM), 13 times lower than the $K_{m,app}$ (2 μM). At high QZ59-*SSS* and Hoechst 33342 concentrations, QZ59-*SSS* tends to activate the Hoechst 33342 efflux, in a cooperative manner. The estimated activation constant $K_{A,app}$ of 4.5 μM suggests that this effect is mediated either through binding to additional low-affinity sites and/or with a poor efficacy.

The effects of QZ59-*SSS* on P-gp-mediated daunorubicin efflux are shown in Fig. 3D–F. A detailed analysis is shown in Fig. S2B. The distribution of the

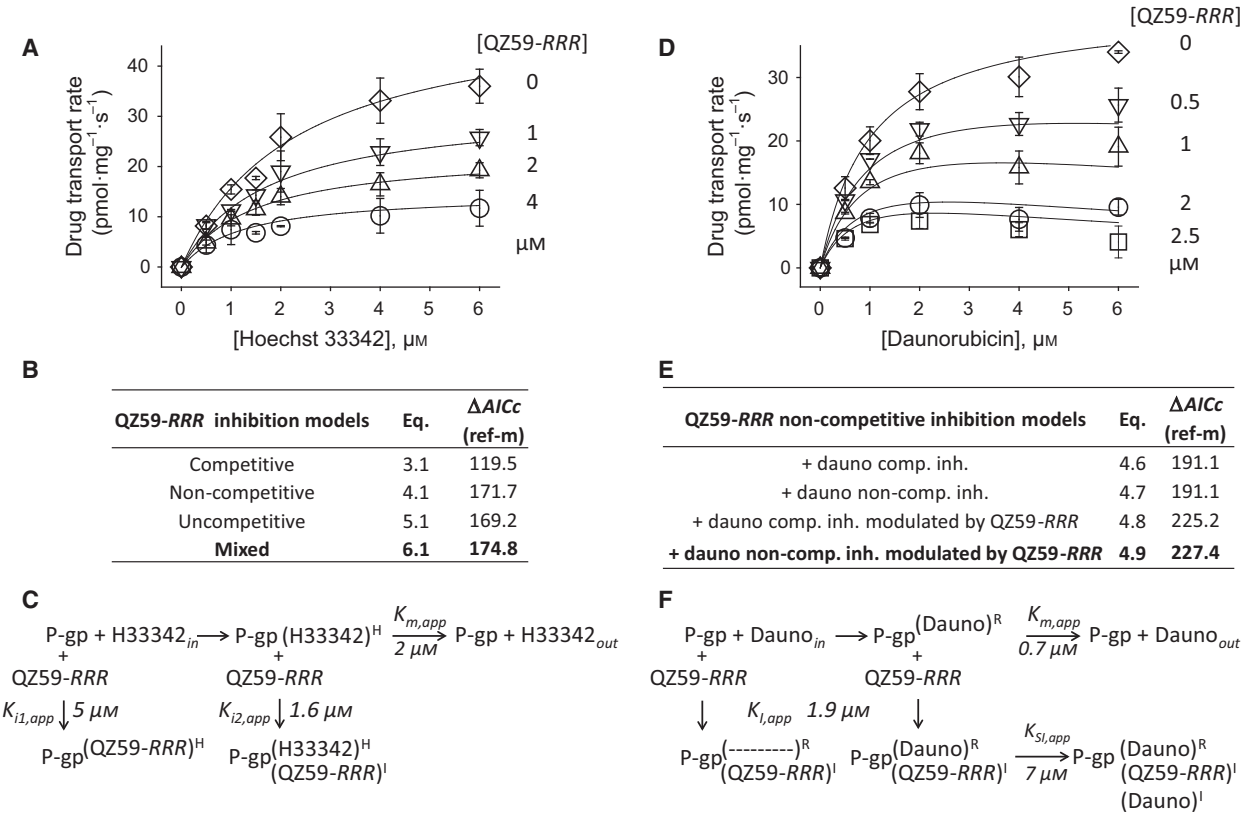


Fig. 2. QZ59-*RRR* effect on P-gp-mediated drug transport. A, D. Plots of Hoechst 33342 (A) and daunorubicin (D) transport rates as a function of drug and QZ59-*RRR* concentrations. The experiments were performed in triplicate, generating 84 (A) and 90 (D) measurements. respectively. Traces correspond to the best fit obtained using Eqn 6.1 (A) or Eqn 4.9 (D) (Table 1). (B, E) ΔAIC_c scores of the tested models in respect of the reference model, Eqn 2.1. (C, F) Reaction schemes of substrates transport and QZ59-*RRR* effects showing the estimated constants. H33342, Hoechst 33342; Dauno, daunorubicin; R and H, R and H transport sites; I, inhibition site. K_{i1}, K_{i2}, K_I, K_{SI} and K_m correspond to the inhibition and Michaelis constants of the reactions, as described in the text.

values suggests inhibition by daunorubicin at high substrate and QZ59-*SSS* concentrations as observed with QZ59-*RRR*, but to a lower extent. Competitive and mixed inhibition models (Eqns 3.9 and 6.9, Fig. 3E) gave the largest ΔAIC_c values (199 and 199.5, respectively). The mixed inhibition model had a marginally better fit, as shown in Fig. 3D and confirmed by the corresponding GOF values (Fig. S2B). As drawn in Fig. 3F, the mixed inhibition model suggests that QZ59-*SSS* competes with daunorubicin at the R transport site. This effect occurs with an inhibition constant $K_{i1,app}$ of 0.35 ± 0.08 μM, which is approximately two times lower than the K_m of daunorubicin transport. According to this model, QZ59-*SSS* also binds to the P-gp/daunorubicin complex on an inhibitory site, with an inhibition constant $K_{i2,app}$ of 1.3 ± 0.9 μM. $K_{i2,app}$ is approximately four times higher than $K_{i1,app}$, indicating a marked competitive inhibition tendency. This is confirmed by the close ΔAIC_c obtained with the competitive model (Eqn 3.9), leading to a $K_{I,app}$ of 0.28 ± 0.05 μM. At high concentrations, daunorubicin binds to an inhibitory site, as described above for the effect of QZ59-*RRR* on daunorubicin efflux. This effect is limited since the inhibition constant $K_{SI,app}$ is estimated to be 16 μM (Fig. 3F).

Taken together, these data suggest that QZ59-*SSS* competes with drug substrates, which are historically known to bind the H and R sites, with a positional substrate overlap that is much more marked for the H site than for the R site. Additionally, at higher concentrations this inhibitor tends to activate the efflux of drugs that bind the H site.

Discussion

Due to its significant role in imparting multidrug resistance properties to cancer cells, P-gp is a validated therapeutic target. Inhibition of its drug export function is critical in order to improve the efficacy of many clinically important chemotherapeutic drugs. Thus,

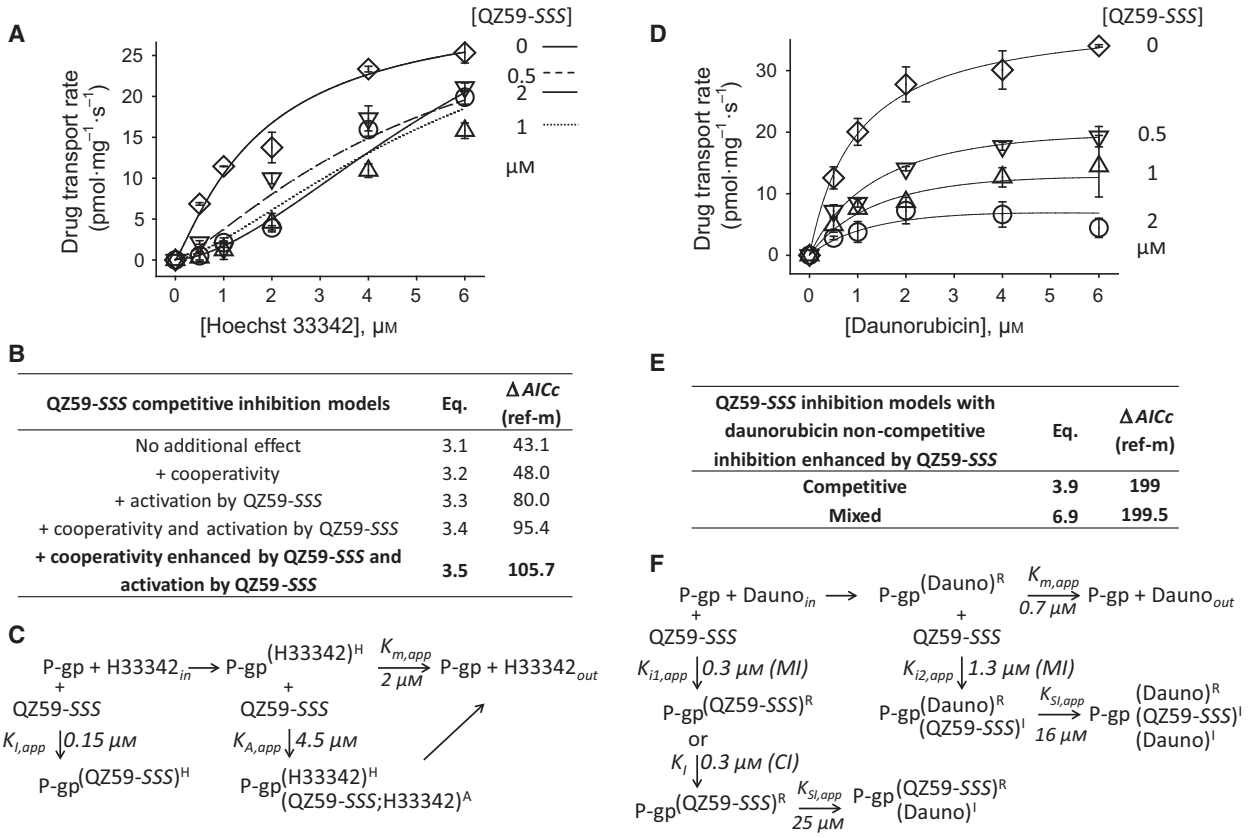


Fig. 3. QZ59-*SSS* effects on P-gp-mediated drug transport. (A, D) Plots of Hoechst 33342 (A) and daunorubicin (D) transport rates as a function of drug and QZ59-*SSS* concentrations. The experiments were performed in triplicate generating 72 measurements in both cases. Traces correspond to the best fits obtained with Eqn 3.5 (A) or Eqn 6.9 (D) (Table 1). For clarity, traces in (A) are shown in solid lines for [QZ59-*SSS*] = 0 and 2 μM, and as dashed and dotted lines for [QZ59-*SSS*] = 0.5 and 1 μM, respectively. (B, E) ΔAIC_c scores of the tested models in respect of the reference model, Eqn 2.1. (C, F) Reaction schemes of substrate transport and QZ59-*SSS* effects showing the estimated constants. H33342, Hoechst 33342; Dauno, daunorubicin; QZ59S, QZ59-*SSS*; R and H, R and H transport sites; I, inhibition sites; A, activation site. K_{i1}, K_{i2}, K_I, K_{SI}, K_A and K_m correspond to the inhibition, activation and Michaelis constants of the reactions, as described in the text.

understanding the inhibitory mechanisms pertinent to this pump will help to develop more potent and selective inhibitors. The system used here allowed estimation of the apparent transport and inhibition constants; we found them to range between 0.1 and 25 μM, values that correlate well with K_D recently estimated for compounds that interact with P-gp [18,19].

The QZ59-*RRR* and -*SSS* enantiomers are two inhibitors that have been co-crystallized with mouse P-gp [13,15], each in one of the 6 inward-facing conformations resolved to date [14]. The mouse P-gp is 87% identical in protein sequence to human P-gp, suggesting that the inhibitory mechanism of QZ59 compounds towards the mouse and human P-gp are similar. We found that in contrast to the *RRR* enantiomer, the *SSS* enantiomer competes with drugs that bind to canonical H and R sites. This result is intriguing because QZ59-*RRR* and -*SSS* are distributed in a common groove in the outer leaflet of the membrane domain (Fig. 4A) while R- and H-drugs-binding sites are distinct from each other, as previously reported [7].

This contradiction is only superficial. Indeed, the $K_{I,app}$ values for QZ59-*SSS* competitive inhibition toward Hoechst 33342 and daunorubicin (0.15 and 0.3 μM respectively) suggest an overlap between drugs and QZ59-*SSS*, as the corresponding $K_{m,app}$ are 13- and twofold higher. In addition, QZ59-*RRR* shows a moderate competitive tendency toward Hoechst 33342 but acts strictly non-competitively toward daunorubicin. This suggests that the H site broadly overlaps with the QZ59-binding groove, contrarily to the R site, which partially shares it. Docking simulations performed with Hoechst 33342 in the mouse P-gp X-ray structure indeed gave the best *in silico*

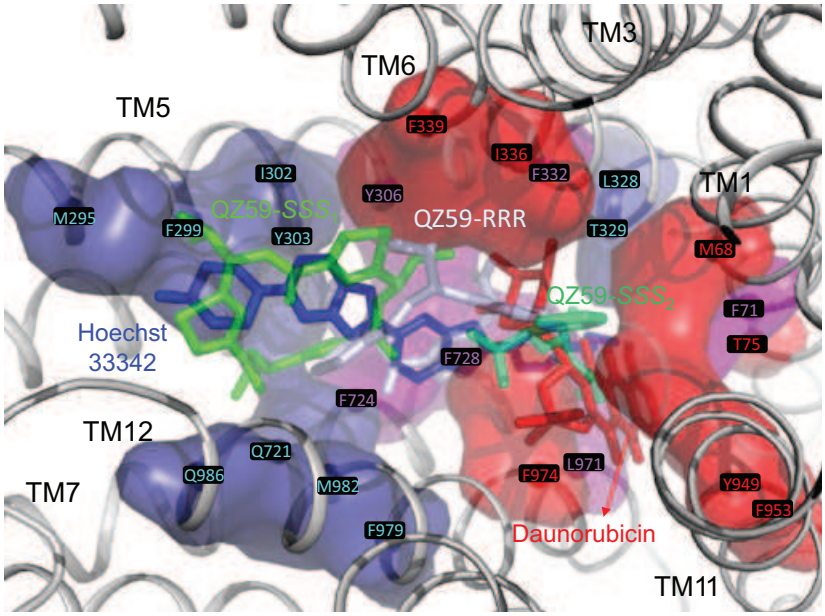


Fig. 4. *In silico* localization of the H and R drug-binding sites in the mouse P-gp. QZ59 binding region of the mouse P-gp X-ray structure (PDB code 4LSG) is shown in inset. QZ59-*SSS*$_1$, QZ59-*SSS*$_2$ and QZ59-*RRR* are shown in green and pale blue sticks, as resolved in their X-ray structures (PDB codes 3G61/4M2T and 3G60/4M2S, respectively). Hoechst 33342 (blue) and daunorubicin (red) correspond each to the best-rated molecule location from independent docking simulations as detailed in Fig. S3A,B (–11.6 and –13.4 kcal/mol). Residues forming the putative H and R drug-binding sites are shown in surface and labeled; they are colored in blue for the H site, in raspberry for the R site and purple for those being common to both sites. The figure was drawn using Pymol 1.6.

affinities for locations along the groove (Fig. S3A). Notably, as defined, the H site is bordered by residues F299, Y303 and Y306 (Fig. 4A) which have been proposed to be part of it [12]. Another series of docking simulations done with daunorubicin gave the best *in silico* affinities for locations mainly overlapping the most embedded QZ59-*SSS*, (a second binding region was also found close to the most accessible QZ59-*SSS*) (Fig. S3B). Similar results were obtained by docking the thiol-reactive analog of rhodamine, methanethiosulfonate (MTS) rhodamine (Fig. S3C) used previously by Loo and Clarke for mapping cysteine-mutagenized residues protected from modification by rhodamine B [10]. As defined, the R site overlaps the most embedded QZ59-*SSS* location and almost not that of QZ59-*RRR* (Fig. 4A), an observation fully consistent with the QZ59 enantiomers inhibition patterns. When considering the residues forming each H and R pocket (Fig. 4B,C), about half are distinct, in agreement with the competition between each drug at concentrations above 2 μM as previously reported [7]. A recent study on elacridar and tariquidar-based inhibitors of P-gp reached the same conclusions [20].

The location overlap between Hoechst 33342 and QZ59-*RRR* suggested in Fig. 4A reflects well the mixed non-competitive nature of the inhibition pattern resulting from a poor competitive component ($K_{i1,app}$ 5 μM) and a non-competitive one [$K_{i2,app}$ 1.6 μM, close to the $K_{m,app}$ (2 μM)]. This modest range of inhibition constants suggests that QZ59-*RRR* may shift in the presence of Hoechst 33342, similarly to the situation observed with proflavin and ethidium bound to the multidrug-binding transcription repressor QacR [21]. Combining the inward-facing conformations resolved to date [13–15] (Movie S1) shows that the R and H drug-binding site locations proposed here (Fig. 4) exist transiently, a situation that seems reasonable in a context of drug translocation. Experimentally determined structures of P-gp with bound drug substrates and inhibitors, together with computational simulations, will assist in obtaining a better understanding of this fascinating membrane protein pump.

Experimental procedures

Reagents

Daunorubicin and Hoechst 33342 were obtained from Sigma Aldrich (Saint Quentin-Fallavier, France). The QZ59-*RRR* and QZ59-*SSS* cyclic hexapeptides were synthesized as described by Tao *et al.* [22]. The compounds

were dissolved in 100% dimethylsulfoxide at 20 mM stock concentration and stored at −20 °C.

Cell culture

The NIH3T3 parental cell line and NIH3T3/P-gp drug-resistant cell line transfected with human MDR1/A-G185 [23], were from American Type Culture Collection (Manassas, VA) and used as described previously [24]. Cells were grown in cell growth medium containing Dulbecco's modified Eagle's medium (PAA Laboratories/GE Healthcare Europe GmbH, Velizy-Villacoublay, France), 10% fetal bovine serum (PAA laboratories) and 1% penicillin/streptomycin (PAA Laboratories). The NIH3T3/P-gp growth medium was also supplemented with 60 ng·mL^{-1} colchicine. Cells were incubated at 37 °C in humidified 5% CO_2.

Drug transport

We followed the protocol described by Spoelstra *et al.* [16], who fully characterized P-gp-mediated daunorubicin efflux. We used steady-state conditions similar to those described by Litman *et al.* [25], incubating control cells and P-gp-expressing cells with up to 6 μM daunorubicin (rather than 5 μM) for 1 hour in the culture medium. The same protocol was used for Hoechst 33342. About 50 000 NIH3T3 or NIH3T3/P-gp cells were incubated as above for 24 h in cell growth medium, and then 1 h in the same medium with or without 0–6 μM daunorubicin or Hoechst 33342 and with or without 0–4 μM QZ59s. After incubation, cells were washed with NaCl/Pi (PBS, PAA laboratories), trypsinised, and stored on ice for ~ 30 min before quantification of their intracellular drug amount by flow cytometry. We used flow cytometry to confirm that virtually no cell lysis occurred during the overall treatment as cell fragments and dead cells corresponded to less than 0.1%. Flow cytometry was performed with a FACS Calibur cytometer or FACS LSR II (Becton Dickinson Biosciences, Franklin Lakes, New Jersey, USA). Daunorubicin and Hoechst 33342 were excited using 488 and 355 nm lasers; the corresponding emissions were recorded with a 530/30 band pass filter and 450/50 band pass filter, respectively. All experiments were standardized with mid-range FL1 fluorescence beads (Becton Dickinson Biosciences). Data were collected with CELLQUEST Pro 4.0 software or FACSDIVA 6.1.3 software (Becton Dickinson Biosciences) and then exported to FlowJo (TreeStar, Ashland, Oregon, USA) for analysis. To facilitate the kinetic analysis, arbitrary fluorescence values were converted to apparent intracellular drug concentrations (nmol·mg^{-1}) by quantifying spectrophotometrically the intracellular drug amounts (Fig. S4). This led to estimates that were close to those reported previously [16]. The drug concentration range was then optimized, showing that a maximal accumulation was reached close to 6 μM, with or without QZ59 inhibitors (Fig. S5) [16]. Although these drug/dye partly are well known to accumulate into the nucleus, this effect had a limited impact since we only considered the difference in accumulation between control NIH3T3 and NIH3T3-P-gp cells incubated in comparable conditions [25]. Intracellular drug accumulation was fitted using equations 1.1 or 1.2:

$$y = a(1 - e^{-bx}) \tag{1.1}$$

$$y = a/(1 + e^{-(x-x_0)/b}) \tag{1.2}$$

Regression analyses

Equations 2.1–7.1 that were used or built to fit P-gp-mediated drug efflux are shown in Table 1. Regressions were performed using the *nls* function in the statistical software R v2.14.1. Scripts are available upon request. Model validation and selection were achieved by evaluating each model by goodness of fit (GOF) diagnostics (Lowess function in R) comparing predictions to observations, and then calculating the Akaike's Information Criterion (AIC) score (Eqn 8, AIC function in R):

$$AIC = -2^{*}log(likelihood) + (k \times n_{\mathrm{par}}) \tag{8}$$

In which, n_{par} represents the number of parameters in the fitted model and k the penalty per parameter to be used (default $k = 2$). The log(likelihood) term is a measure of the fit between predicted and observed values (the lower the deviance value, the better the model); the $(k \times n_{\mathrm{par}})$ term is a penalty for over-fitting by increasing the number of parameters [26]. As the data points/parameters ratio was within 14–40 in this study, we used the corrected AIC score, AIC_c (Eqn 9), which gives a more accurate score by taking into account the number of points and parameters:

$$AIC_c = AIC + 2k(K+1)/(N-K-1) \tag{9}$$

Where N is the number of data points and K is the number of parameters +1. The simplest model with the smallest loss of information due to the prediction gives the smallest AIC_c score [27]. Models were compared by calculating the difference of AIC_c between a reference model and the tested one (m), using the Michaelis and Menten equation (Eqn 2.1) as reference, without inhibition (Table 1). The best model gives the largest ΔAIC_c:

$$\triangle AIC_{c(\mathrm{ref-m})} = AIC_{c(\mathrm{ref})} - AIC_{c(\mathrm{m})} \tag{10}$$

Akaike's weights were calculated for models with close AIC_c scores, typically giving a difference of ~3 or lower, using Eqn (11):

$$Probability = 100 \times e^{-0.5\triangle AIC_{c(\mathrm{m-ref})}}/(1+e^{-0.5\triangle AIC_{c(\mathrm{m-ref})}} \tag{11}$$

A value $\Delta AIC_{c(m\text{-}ref)}$ of 0 will give a probability of 0.5, indicating that the reference and the model tested have the same probability of existence. A $\Delta AIC_{c(m\text{-}ref)}$ of 2, which corresponds to the default value of k in Eqn 8, will give a probability of 0.27, indicating that the model tested is 73% more accurate than the reference model. In the same way, a $\Delta AIC_{c(m\text{-}ref)}$ of 6 will give a probability of 0.05, indicating that the model is 95% more accurate than the reference, etc [28].

Docking

Simulations were performed with AutoDock Vina 1.1.2 [29] and AutoDockTools 4 [30] for preparing ligands (QZ59-*RRR*, QZ59-*SSS*, Hoechst 33342, daunorubicin and MTS-rhodamine) and receptor (mouse apo-Pgp, PDB file 4LSG).The docking box, designed to cover the main part of the membrane region, had a size of $x = 40$, $y = 38$ and $z = 36$ (using a spacing of 1 Å), and was centered on $x = 20$, $y = 52$ and $z = -2$. Docking were performed by allowing flexibility to leu64, met67, met68, phe71, met295, ile302, phe299, tyr303, tyr306, phe332, ile336, phe339, gln721, phe724, phe728, met945, tyr949, phe953, leu971, val973, phe974, ile977, val978 and met982 (55 rotatable bonds). The 'exhaustivity' option was set to the default value of 8. These settings were tested by docking QZ59-*SSS* and QZ59-*RRR* on the mouse P-gp X-ray structure 4LSG (Fig. S6A-B). Scripts and 3D datasets are available upon request.

Acknowledgements

The authors thank Dr. Hélène Cortay (Drug Resistance Mechanism and Modulation group, Molecular and Structural Basis of Infectious Systems, Mixed Research Unit between the National Centre for Scientific Research and Lyon I University n°5086) for useful comments on the manuscript. The authors declare no conflict of interest. This work was funded by NIH to RD and GC (RO1 GM94367). LM was supported by the Rhône-Alpes region (Cluster 10-ARC1 santé, Explora'doc). EH was funded by Fondation Synergie Lyon Cancer and the Ligue Nationale contre le Cancer. PF, VC, LM and ADP were supported by the *Ligue Contre le Cancer*, PF was supported by ANR-EMMA-10-049-01 and ANR-13-BSVS-0001-01, and Lyon Science Transfert (AAP 2009 L616).

References

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

2 Gottesman MM, Fojo T & Bates SE (2002) Multidrug resistance in cancer: role of ATP-dependent transporters. *Nat Rev Cancer* **2**, 48–58.

3 Dean M, Hamon Y & Chimini G (2001) The human ATP-binding cassette (ABC) transporter superfamily. *J Lipid Res* **42**, 1007–1017.

4 Szakacs G, Paterson JK, Ludwig JA, Booth-Genthe C & Gottesman MM (2006) Targeting multidrug resistance in cancer. *Nat Rev Drug Discov* **5**, 219–234.

5 Eckford PDW & Sharom FJ (2009) ABC efflux pump-based resistance to chemotherapy drugs. *Chem Rev* **109**, 2989–3011.

6 Shapiro AB, Kelly Fox, Ping Lam & Ling V (1999) Stimulation of P-glycoprotein-mediated drug transport by prazosin and progesterone. *Eur J Biochem* **259**, 841–850.

7 Shapiro AB & Victor Ling (1997) Positively cooperative Sites for drug transport by P-glycoprotein with distinct drug specificities. *Eur J Biochem* **250**, 130–137.

8 Martin C, Berridge G, Higgins CF, Mistry P, Charlton P & Callaghan R (2000) Communication between multiple drug binding sites on P-glycoprotein. *Mol Pharmacol* **58**, 624–632.

9 Loo TW & Clarke DM (2008) Mutational analysis of ABC proteins. *Arch Biochem Biophys* **476**, 51–64.

10 Loo TW & Clarke DM (2002) Location of the Rhodamine-binding site in the human multidrug resistance P-glycoprotein. *J Biol Chem* **277**, 44332–44338.

11 Qu Q & Sharom FJ (2002) Proximity of bound Hoechst 33342 to the ATPase catalytic sites places the drug binding site of P-glycoprotein within the cytoplasmic membrane leaflet. *Biochemistry* **41**, 4744–4752.

12 Pajeva IK, Globisch C & Wiese M (2004) Structure–function relationships of multidrug resistance P-glycoprotein. *J Med Chem* **47**, 2523–2533.

13 Aller SG, Yu J, Ward A, Weng Y, Chittaboina S, Zhuo R, Harrell PM, Trinh YT, Zhang Q, Urbatsch IL *et al.* (2009) Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science* **323**, 1718–1722.

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

15 Li J, Jaimes KF & Aller SG (2013) Refined structures of mouse P-glycoprotein. *Protein Sci* **23**, 34–46.

16 Spoelstra EC, Westerhoff HV, Dekker H & Lankelma J (1992) Kinetics of daunorubicin transport by P-glycoprotein of intact cancer cells. *Eur J Biochem* **207**, 567–579.

17 Wang E-J, Casciano CN, Clement RP & Johnson WW (2000) Two transport binding sites of P-glycoprotein are unequal yet contingent: initial rate kinetic analysis by ATP hydrolysis demonstrates intersite dependence. *Biochim Biophys Acta* **1481**, 63–74.
18 Melchior DL, Sharom FJ, Evers R, Wright GE, Chu JWK, Wright SE, Chu X & Yabut J (2012) Determining P-glycoprotein–drug interactions: evaluation of reconstituted P-glycoprotein in a liposomal system and LLC-MDR1 polarized cell monolayers. *J Pharmacol Toxicol Methods* **65**, 64–74.
19 Marcoux J, Wang SC, Politis A, Reading E, Ma J, Biggin PC, Zhou M, Tao H, Zhang Q, Chang G *et al.* (2013) Mass spectrometry reveals synergistic effects of nucleotides, lipids, and drugs binding to a multidrug resistance efflux pump. *Proc Natl Acad Sci U S A* **110**, 9704–9709.
20 Pajeva IK, Sterz K, Christlieb M, Steggemann K, Marighetti F & Wiese M (2013) Interactions of the multidrug resistance modulators Tariquidar and Elacridar and their analogues with P-glycoprotein. *Chem Med Chem* **8**, 1701–1713.
21 Schumacher MA, Miller MC & Brennan RG (2004) Structural mechanism of the simultaneous binding of two drugs to a multidrug-binding protein. *EMBO J* **23**, 2923–2930.
22 Tao H, Weng Y, Zhuo R, Chang G, Urbatsch IL & Zhang Q (2011) Design and synthesis of Selenazole-containing peptides for cocrystallization with P-glycoprotein. *Chem Bio Chem* **12**, 868–873.
23 Cardarelli CO, Aksentijevich I, Pastan I & Gottesman MM (1995) Differential effects of P-glycoprotein inhibitiors on NIH3T3 cells transfected with wild-type (G185) or mutant (V185) multidrug transporters. *Cancer Res* **55**, 1086–1091.
24 Arnaud O, Koubeissi A, Ettouati L, Terreux R, Alamé G, Grenot C, Dumontet C, Di Pietro A, Paris J & Falson P (2010) Potent and fully noncompetitive peptidomimetic inhibitor of multidrug resistance P-glycoprotein. *J Med Chem* **53**, 6720–6729.
25 Litman T, Skovsgaard T & Stein WD (2003) Pumping of drugs by P-glycoprotein: A two-step process? *J Pharmacol Exp Ther* **307**, 846–853.
26 Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Autom Control* **19**, 716–723.
27 Burnham KP & Anderson DR (2004) Multimodel inference: understanding AIC and BIC in model selection. *Sociol Methods Res* **33**, 261–304.
28 Motulsky HJ & Christopoulos A (2003) Fitting Models to Biological Data Using Linear and Nonlinear Regression. A Practical Guide to Curve Fitting. Graphpad Software Inc., San Diego.
29 Trott O & Olson AJ (2010) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **31**, 455–461.
30 Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS & Olson AJ (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem* **30**, 2785–2791.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site:

Figure S1. (A) Detailed process for fitting the inhibition by QZ59-*RRR* of Hoechst 33342 efflux. (B) Fitting of the inhibition by QZ59-*RRR* of daunorubicin efflux.

Figure S2. (A) Fitting of the inhibition by QZ59-*SSS* of Hoechst 33342 efflux. (B) Fitting of the inhibition by QZ59-*SSS* of daunorubicin efflux.

Figure S3. (A) Hoechst 33342 docking on the mouse P-gp. (B) Daunorubicin docking on the mouse P-gp. (C) MTS rhodamine docking on the mouse P-gp.

Figure S4. Intracellular drugs quantitations.

Figure S5. Intracellular drug-substrate saturation.

Figure S6 (A) QZ59-*SSS* docking on the mouse P-gp. (B) QZ59-*RRR* docking on the mouse P-gp.

Table S1. Detailed results of the regression analyses.

Movie S1. Conformational changes of the mouse P-gp visualized by morphing.

Supporting information for:

UNDERSTANDING POLYSPECIFICITY WITHIN THE SUBSTRATE-BINDING CAVITY OF THE HUMAN MULTIDRUG RESISTANCE P-GLYCOPROTEIN

Lorena Martinez, Ophélie Arnaud, Emilie Henin, Houchao Tao, Vincent Chaptal, Rupak Doshi, Thibaud Andrieu, Sébastien Dussurgey, Michel Tod, Attilio di Pietro, Qinghai Zhang, Geoffrey Chang and Pierre Falson

Table S1. Detailed results of the regression analyses. As detailed in Methods, the Table shows the inhibition model tested, the number of points used for the test, the resulting AIC and corrected AIC (AIC_c) which takes into account the number of parameters in the considered equation and the constant K, the difference in AICc (ΔAICc) to the reference model (no inhibition for the overall data set; the larger the difference the better the model), the Akaike's weights, and the deduced kinetic parameters.

Figure S1A. Detailed process for fitting the inhibition by QZ59-*RRR* of Hoechst 33342 efflux. Plots of Hoechst 33342 transport velocities in respect of Hoechst 33342 concentrations and with 0, 1, 2 and 4 μM QZ59-RRR are shown with diamonds, blue triangles-down, green triangles-up and red circles, respectively. Fits carried out with each enzymatic model (Eqns 2.1–6.1, detailed in Tables 1 and S1) are shown in the 1st column. As detailed in the Experimental procedures, fits were evaluated by the Goodness-of-fit (GOF) shown in the 2nd column in which the magenta line corresponds to the lowess function. The AICc shown in the 5th column was also used to evaluate models and scoring them (ΔAICc, 6th column), in respect of a reference corresponding to the simplest Michaelis & Menten, Eqn 2.1. Each AICc takes into account the number of experimental points N (84, 3rd column) and the number of parameters used in each model (*npar*, 4th column).

As shown, each inhibition model gives an AICc ranging from 423.9 to 368.6, lower than the reference model which does not includes any inhibition (543.4). Among these models, the mixed inhibition model gives the lowest AIC_c (368.6; the lower the AIC the better the model), and the largest ΔAIC_c (174.8). These scores are significantly better than those calculated for the non-competitive inhibition model (AIC_c of 371.7 and ΔAIC_c of 171.7). The difference in AIC_c between both models is of 371.7-368.6 = 3.1 making that the non-competitive inhibition model is 4.7 times less probable than the mixed inhibition model (see the Akaike's weights in Table S1). Such a rationale provides the classification indicated in the 7th column.

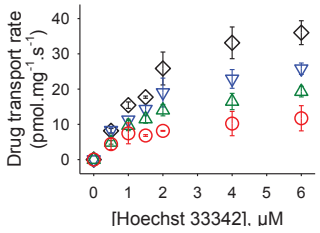
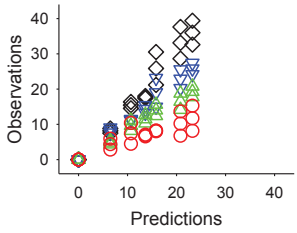
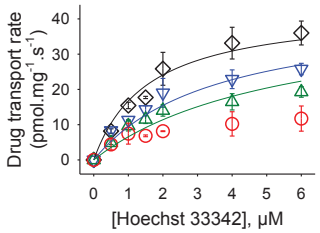
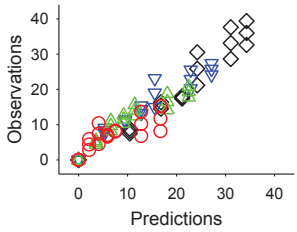
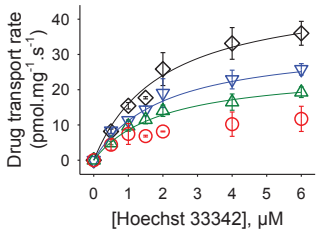
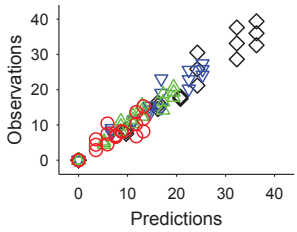
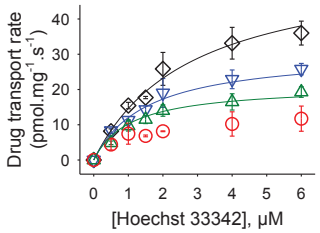
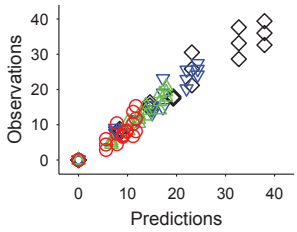
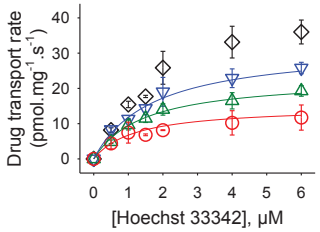
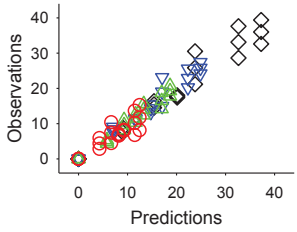
Model	GOF	N	*npar*	AICc	ΔAICc (ref-mod)	rank
No inhibition (eq. 2.1, reference)						
		84	2 K_m V_m	543.4	0	-
Competitive inhibition (eq. 3.1)						
		84	3 K_m V_m K_I	423.9	119.5	4
Non-competitive inhibition (eq. 4.1)						
		84	3 K_m V_m K_I	371.7	171.7	2
Uncompetitive inhibition (eq. 5.1)						
		84	3 K_m V_m K_I	374.2	169.2	3
Mixed inhibition (eq. 6.1)						
		84	4 K_m V_m K_{i1} K_{i2}	368.6	174.8	1

Figure S1B. Fitting of the inhibition by QZ59-*RRR* of daunorubicin efflux. Plots of daunorubicin transport velocities in respect of daunorubicin concentrations and with 0, 0.5, 1, 2 and 2.5 µM QZ59-*RRR* are shown with diamonds, blue triangles-down, green triangles-up, red circles and magenta squares, respectively. Only fits carried out with enzymatic models shown in Fig. 2E are detailed (Eqns 4.6–4.9) added of the standard non-competitive inhibition model, Eqn. 4.1. Other details are in Table S1. As in Fig. S1A, fits are shown in the 1st column, GOF in the 2nd column in which the orange line corresponds to the lowess function. AIC_c and ΔAIC_c, are in the 5th and 6th columns, still calculated in respect of the simplest Michaelis & Menten, Eqn 2.1 (not shown, see Table S1). The number of experimental points N was of 90 (3rd column) and the number of parameters, *npar*, used in each model is displayed in the 4th column.

As detailed in Table S1, although the non-competitive model for QZ59-*RRR* inhibition of daunorubicin efflux (Eqn 4.1, AIC_c of 423.2) was not the best starting model when compared to the uncompetitive or mixed inhibition models (Eqns 5.1 and 6.1; AIC_c of 415.7 and 416.5), it allowed to reach the lowest AIC_c when adding an effect of inhibition by substrate enhanced by QZ59-*RRR* (Eqns 4.8 or 4.9). Models based on Eqns 4.1, 5.1 and 6.1 did not fit correctly data at high daunorubicin and QZ59-*RRR* concentrations (shown in the figure with Eqn 4.1). As the pattern of these data is typical of a competitive/non-competitive inhibition by substrate at high concentrations, we tested these effects in the corresponding equations (Eqns 4.6 and 4.7). The resulting fits were not better, with AIC_cs even higher (424.4). Since this substrate inhibition was more pronounced at high QZ59-*RRR* concentration (squares), we evaluated the possibility that QZ59-*RRR* enhances such inhibition, leading to models with Eqns 4.8 and 4.9. This allowed to get a much better fit and a significant decrease of the AIC_c (390.3 and 388.2). As Eqn 4.9 gave the lowest AIC_c score (388.2), this suggests that daunorubicin at high concentration (K_{SI} of 7 µM) reaches another location than the R site which hampers its own efflux, in the presence of high concentrations of QZ59-*RRR*.

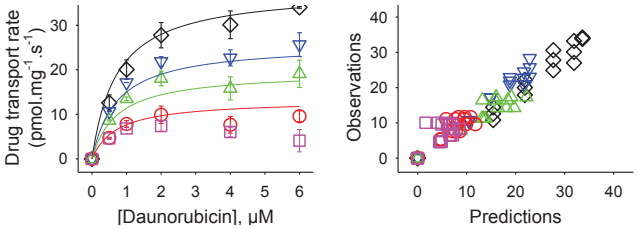
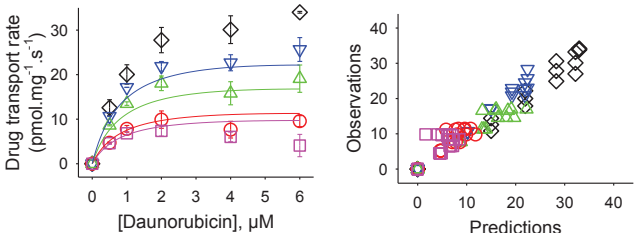
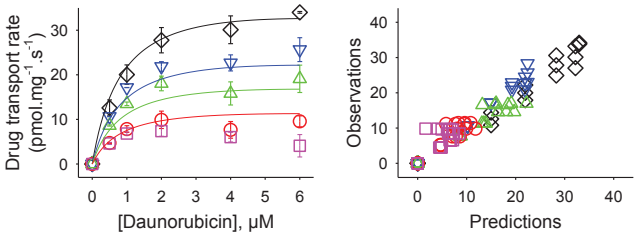
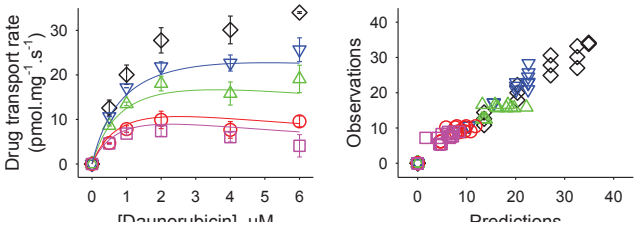
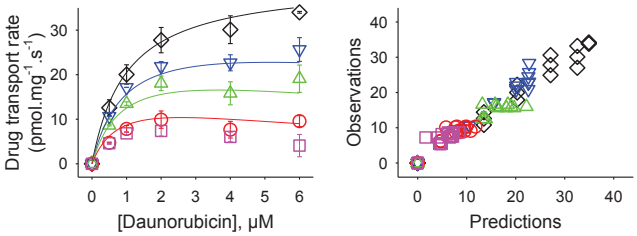
Model	GOF	N	*npar*	AICc	ΔAICc (ref-mod)	rank
Non-competitive inhibition by QZ59-RRR (eq. 4.1)						
		90	3 K_m V_m K_I	423.2	192.3	3
… + competitive inhibition by daunorubicin (eq. 4.6)						
		90	4 K_m V_m K_I K_{SI}	424.4	191.1	4
… + non-competitive inhibition by daunorubicin (eq. 4.7)						
		90	4 K_m V_m K_I K_{SI}	424.4	191.1	4
… + competitive inhibition by daunorubicin enhanced by QZ59-RRR (eq. 4.8)						
		90	4 K_m V_m K_I K_{SI}	390.3	225.2	2
… + non-competitive inhibition by daunorubicin enhanced by QZ59-RRR (eq. 4.9)						
		90	4 K_m V_m K_I K_{SI}	388.2	227.3	1

Figure S2A. Fitting of the inhibition by QZ59-*SSS* of Hoechst 33342 efflux. Plots of Hoechst 33342 transport velocities in respect of Hoechst 33342 concentrations and with 0, 0.5, 1, and 2 µM QZ59-*SSS* are shown with diamonds, blue triangles-down, green triangles-up, and red circles, respectively. Only fits carried out with enzymatic models displayed in Fig. 3B are detailed (Eqns 3.1–3.5). Other details are in Table S1. As in Fig. S1A, fits are shown in the 1st column, GOF in the 2nd column in which the magenta line corresponds to the lowess function. AIC_c and ΔAIC_c, are in the 5th and 6th columns, still calculated in respect of the simplest Michaelis & Menten, Eqn 2.1 (not shown, see Table S1). The number of experimental points N was of 72 (3rd column) and the number of parameters, *npar*, used in each model is shown in the 4th column.

As detailed in Table S1, the competitive model for QZ59-*SSS* inhibition of Hoechst 33342 efflux (Eqn 3.1) gave the lowest AIC_c (352.5) among the standard models of inhibition. However, such a model only fitted data at 0 and 0.5 µM QZ59-*SSS*. As the data distribution of Hoechst 33342 efflux rates suggested a cooperativity together with an activation by QZ59-*SSS*, such effects were evaluated in models with Eqn 3.2 (cooperativity alone), Eqn 3.3 (activation alone) and Eqn 3.4 (both effects). The latter gave the best decrease of AIC_c (300.2 compared to 347.6 and 315.6), however data at high Hoechst 33342 and QZ59-*SSS* concentrations were still poorly fitted. A significant improvement was obtained when evaluating an enhancement by QZ59-*SSS* of the cooperativity (Eqn 3.5, AIC_c of 289.9). Other models did not get a better fit.

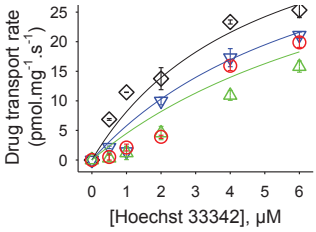
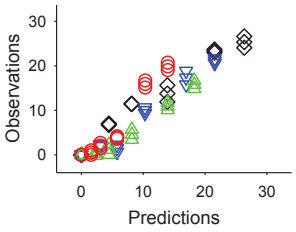
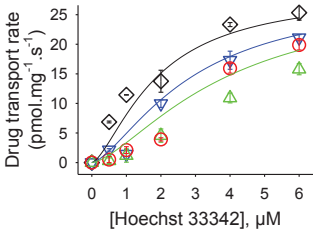
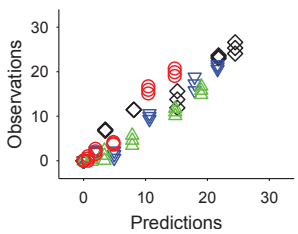
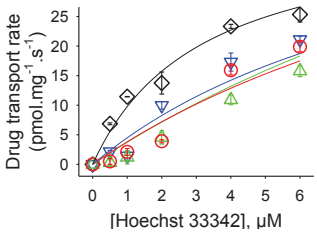
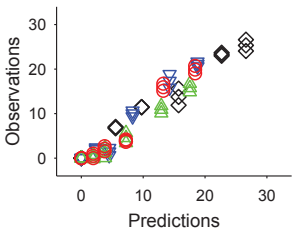
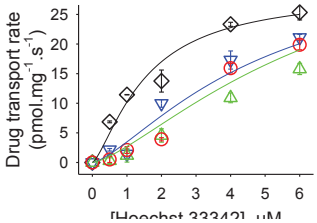
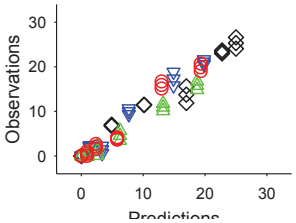
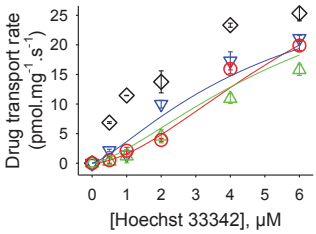
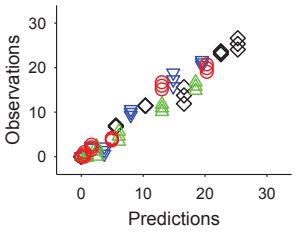
Model	GOF	N	*npar*	AICc	ΔAICc (ref-mod)	rank
Competitive inhibition by QZ59-SSS (eq. 3.1)						
		72	3 K_m V_m K_I	352.5	43.1	5
… + allostery (eq. 3.2)						
		72	4 K_m V_m K_I h	347.6	48.0	4
… + activation by QZ59-SSS (eq. 3.3)						
		72	4 K_m V_m K_I K_A	315.6	80.0	3
… + allostery and activation by QZ59-SSS (eq. 3.4)						
		72	5 K_m V_m K_I K_A h	300.2	95.4	2
… + allostery enhanced by QZ59-SSS and activation by QZ59-SSS (eq. 3.5)						
		72	5 K_m V_m K_I K_A h	289.9	105.7	1

Figure S2B. Fitting of the inhibition by QZ59-*SSS* of daunorubicin efflux. Plots of daunorubicin transport velocities in respect of daunorubicin concentrations and with 0, 0.5, 1 and 2 µM QZ59-SSS are shown with diamonds, blue triangles-down, green triangles-up and red circles, respectively. Fits carried out with enzymatic models for competitive, non-competitive, mixed inhibition and those shown in Fig. 3E are detailed (Eqns 3.1, 4.1, 6.1, 3.9 and 6.9). Other details are in Table S1. As in Fig. S1A, fits are shown in the 1st column, GOF in the 2nd column in which the magenta line corresponds to the lowess function. AIC_c and ΔAIC_c, are in the 5th and 6th columns, still calculated in respect of the simplest Michaelis & Menten, Eqn 2.1 (not shown, see Table S1). The number of experimental points N was of 72 (3rd column) and the number of parameters, *npar*, used in each model is shown in the 4th column.

As detailed in Table S1, the competitive, non-competitive, uncompetitive and mixed inhibition models could fit the dataset but not data at high substrate and inhibitor concentrations (daunorubicin = 6 µM, QZ59-*SSS* = 2 µM). Such data suggested an inhibition by substrate as already concluded with QZ59-*RRR*, and thus the corresponding models were evaluated. As shown, taking into account this inhibition by substrate enhanced by QZ59-*SSS*, both the competitive and mixed inhibition models (Eqns 3.9 and 6.9) produced a better fit together with a lower AIC_c.

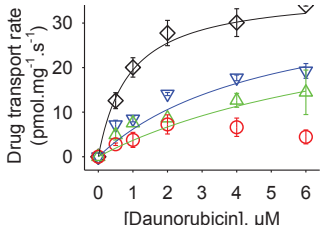
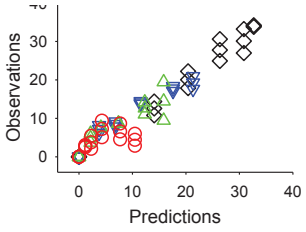
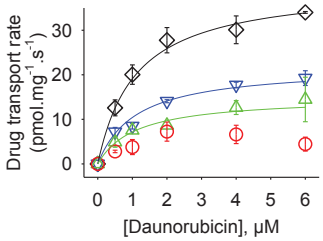
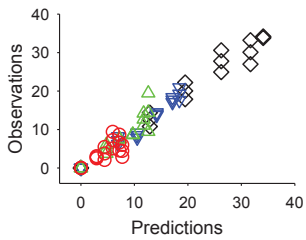
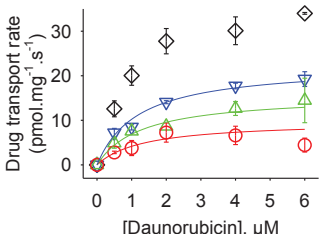
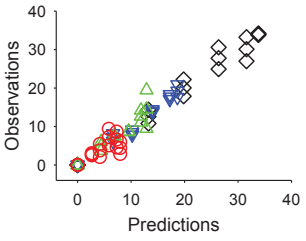
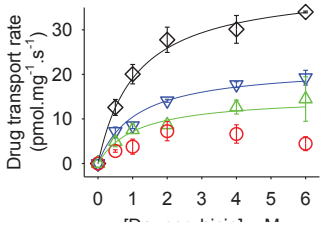
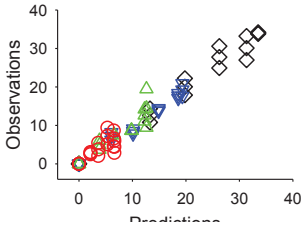
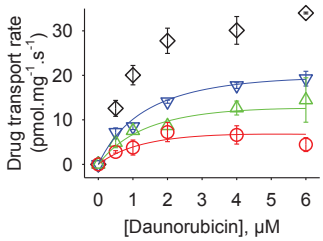
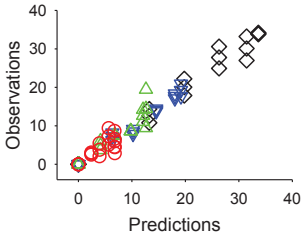
Model	GOF	N	*npar*	AICc	ΔAICc (ref-mod)	rank
Competitive inhibition by QZ59-SSS (eq. 3.1)						
		72	3 K_m V_m K_I	347.2	149.0	5
Non-competitive inhibition by QZ59-SSS (eq. 4.1)						
		72	3 K_m V_m K_I	298.1	198.1	3
Mixed inhibition by QZ59-SSS (eq. 6.1)						
		72	3 K_m V_m K_I	299.3	197.0	4
Comp. inh. by QZ59-SSS + non-comp. inh. by daunorubicin enhanced by QZ59-SSS (eq. 3.9)						
		72	4 K_m V_m K_I K_{SI}	297.2	199.1	2
Mixed inh. by QZ59-SSS + non-comp. inh. by daunorubicin enhanced by QZ59-SSS (eq. 6.9)						
		72	5 K_m V_m K_{I1} K_{I2} K_{SI}	296.8	199.5	1

Figure S3A. Hoechst 33342 docking on the mouse P-gp X-ray structure. Docking experiments were performed with the PDB file 4LSG, as described in the Experimental procedures. P-gp is shown as a surface, focused on the QZ59s binding region. Crystal locations of QZ59-*SSS* are shown in green sticks, the most buried being partially resolved; crystal location of QZ59-*RRR* is shown in pale blue sticks. The 15 poses of the docked Hoechst 33342 are shown in lines and CPK. The 3 main locations are shown in sticks, colored from the highest to the lowest affinities in magenta, orange and yellow. The calculated affinities are listed on the right. The picture was prepared with Pymol 1.6.

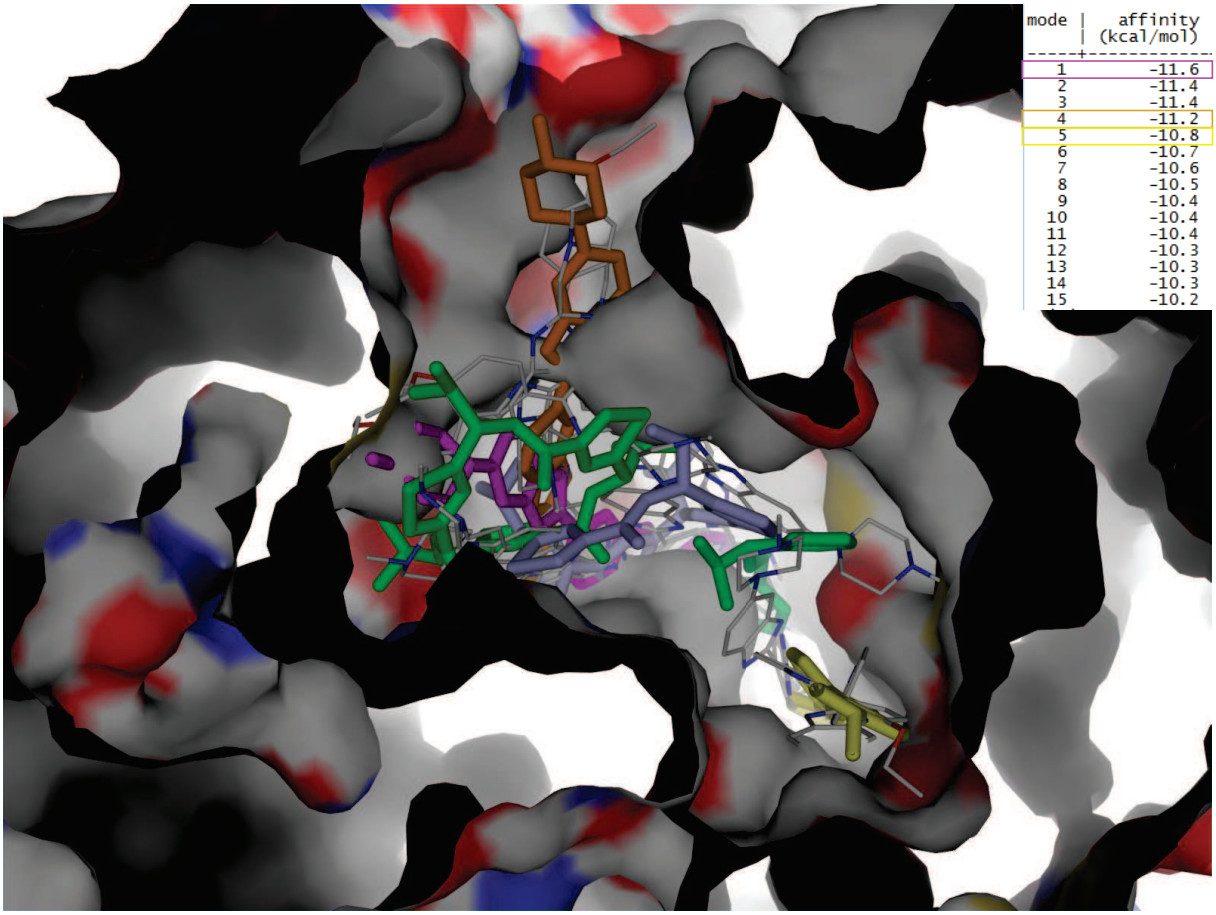


Figure S3B. Daunorubicin docking on the mouse P-gp X-ray structure. Docking experiments were performed as in Fig. S3A. The 14 poses of the docked daunorubicin are shown in lines and CPK. The two main locations are shown in sticks, colored in magenta for the highest affinity and orange for the second best affinity. The calculated affinities are listed on the right.

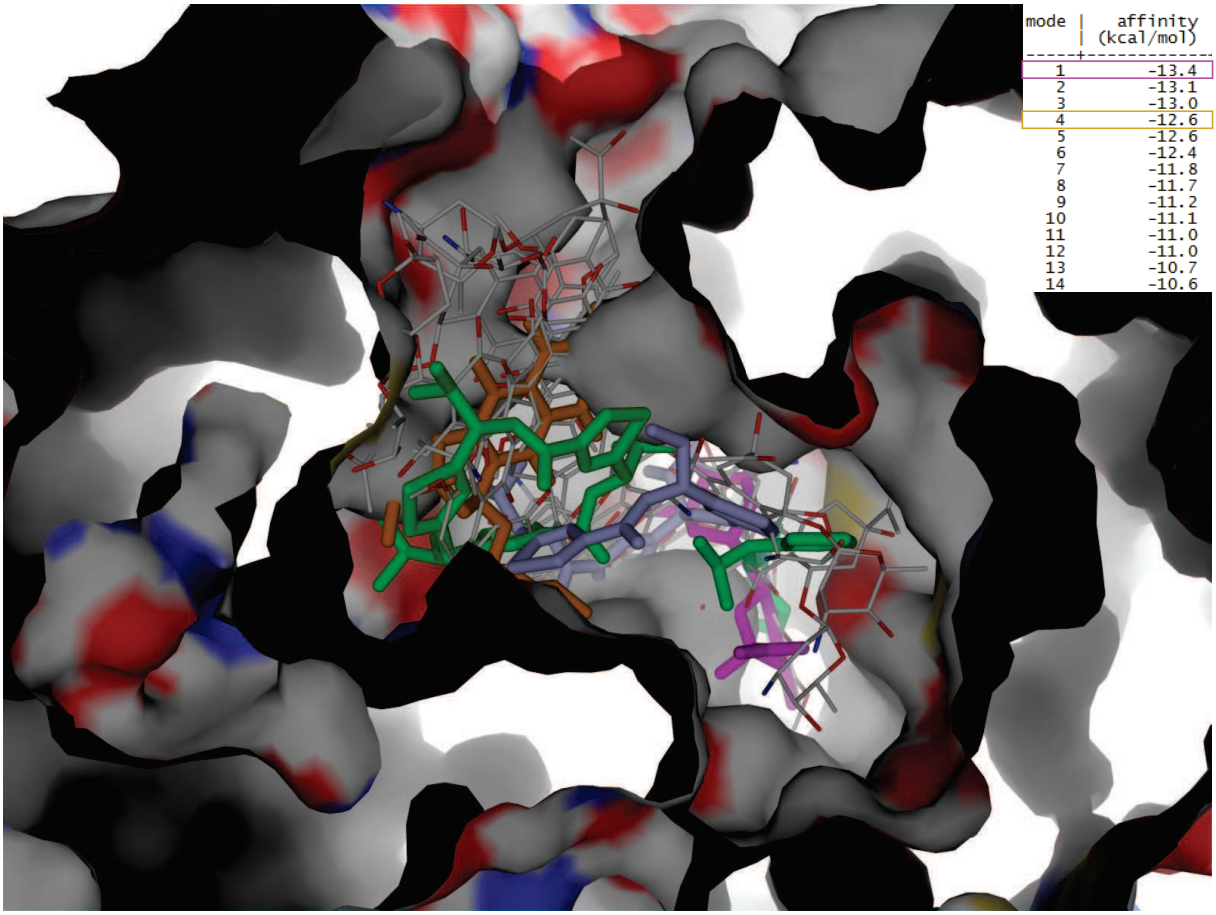


Figure S3C. MTS rhodamine docking on the mouse P-gp X-ray structure. Docking experiments were performed as in Fig. S3A. The 15 poses of the docked MTS rhodamine are shown in lines and CPK. The 3 main locations are shown in sticks, colored from the highest to the lowest affinities in magenta, orange and yellow. The calculated affinities are listed on the right.

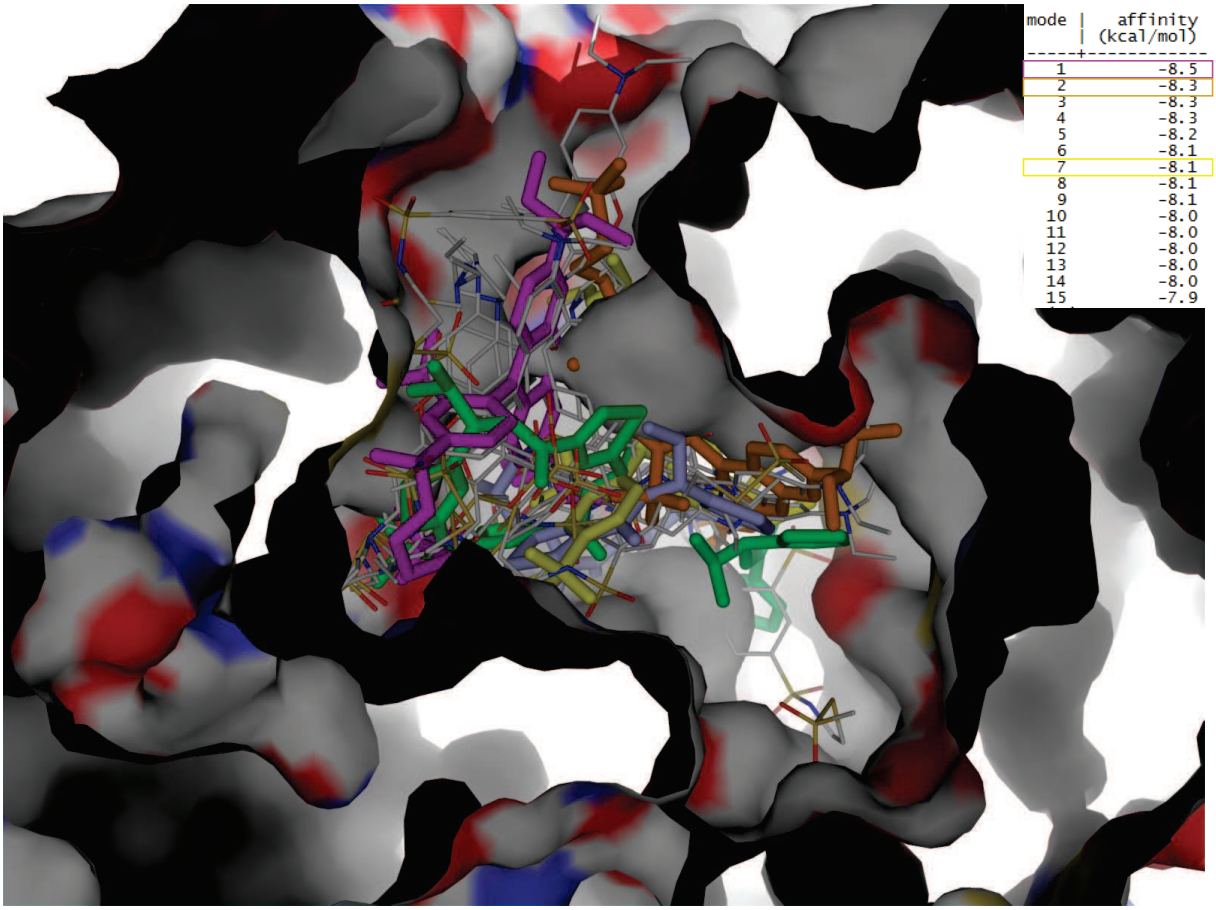


Figure S4. Intracellular drugs quantitations. About 50,000 NIH3T3 cells grown for 24 h in cell growth medium, 37 °C, 5% CO_2, were incubated for 1 h at 37 °C in 200 µL of the same medium with 6 µM of Hoechst 33342 or daunorubicin, corresponding to the maximal fluorescence values of 69,250 and 600 obtained by flow cytometry. Cells were thereafter washed with NaCl/Pi buffer, trypsinised and collected by centrifugation at 5,000 rpm for 5 min. Cells were lysed by suspending the pellets in 200 µL NaCl/Pi buffer added of 1% SDS followed by a 30-min incubation. The non-extracted material was then discarded by a centrifugation at 10,000 rpm, and drugs remaining in the soluble fraction were then quantified by spectrophotometry using standard curves displayed in the Figure. The latter were done from 0 to 3-6 µM of each compound in 200 µL NaCl/Pi buffer by integrating the pic area between 338-400 and 375-600 nm for Hoechst 33342 and daunorubicin, respectively. The intracellular amount of drugs was then deduced as detailed in the Table, per cell or per mg. In this case the protein amount was estimated with the BinChinconinic Acid assay, extracting proteins from 10^7 cells with 1% SDS [4]. A concentration of 4 mg/ml was found, corresponding to 1.2 mg/10^7 cells or 120 pg/cell.

	Hoechst 33342	**Daunorubicin**
Intracellular drug fluorescence at 6 µM	**69,250**	**600**
Pic area	**2.4**	**5.2**
nmol/50,000 cells	**0.6**	**1**
fmol/cell	**12**	**20**
nmol.mg^{-1}	**100**	**170**

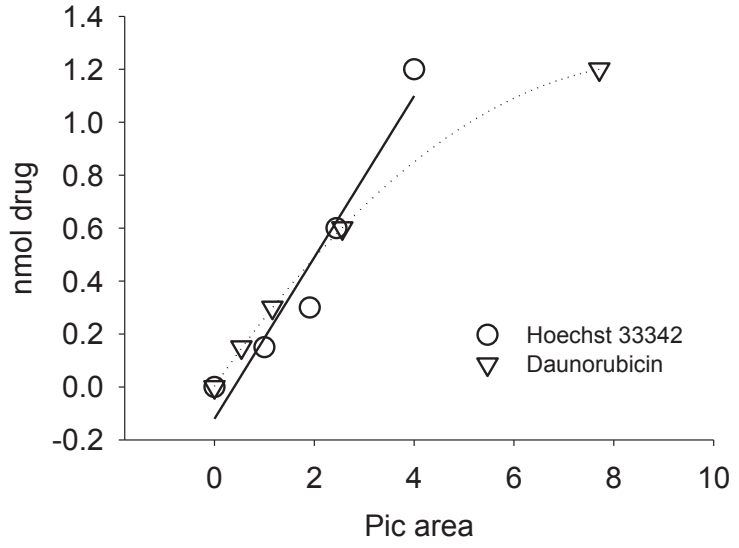


Figure S5. Intracellular drug-substrate saturation. The concentration range for each drug was evaluated by incubating 50,000 NIH3T3 cells with 0-6 µM drugs for 1 h at 37 °C in cell growth medium, and with increasing QZ59s concentrations. After incubation, cells were washed with NaCl/Pi buffer, trypsinised and stored at 4 °C before intracellular drug fluorescence quantitation by flow cytometry. Fluorescence values were then converted into nmol·mg^{-1} as described in the Experimental procedures.

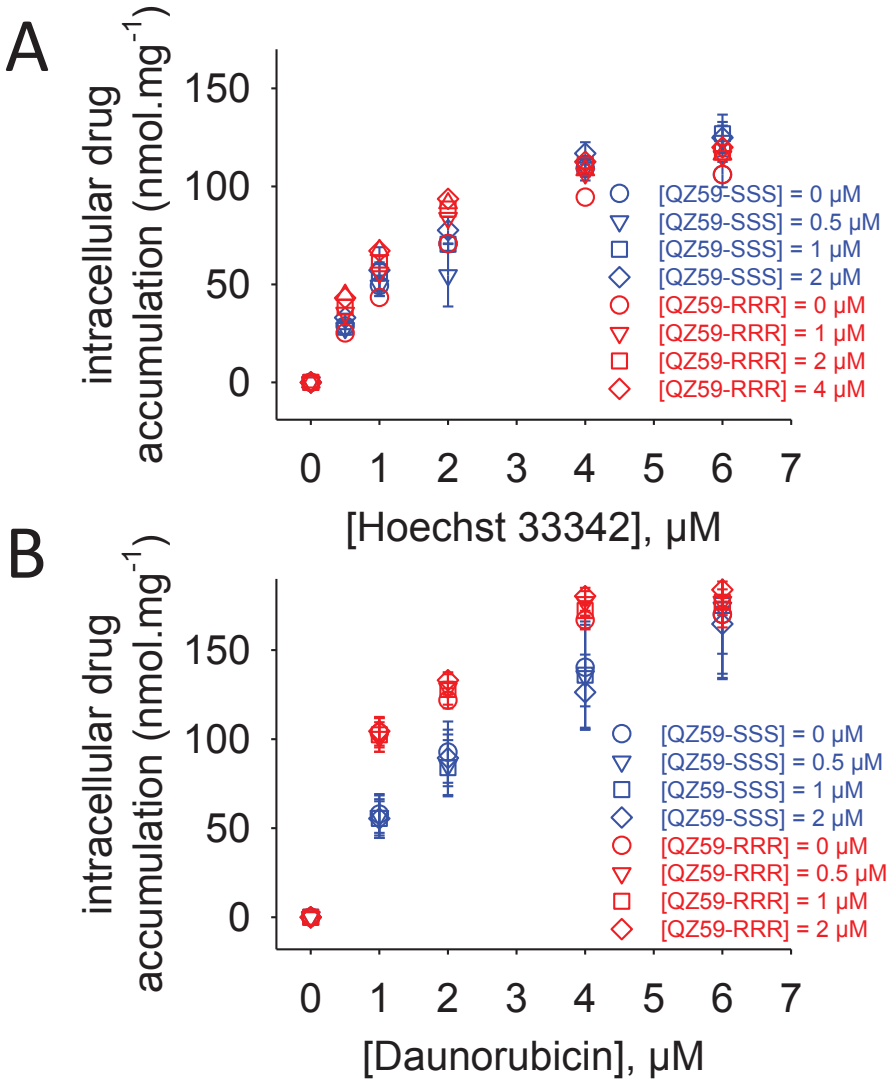


Figure S6A. QZ59-*SSS* docking on the mouse P-gp X-ray structure. Docking experiments were performed as in Fig. S3A. The 15 poses of the docked QZ59-*SSS* are shown in lines and CPK. The 4 main locations are shown in sticks, colored from the highest to the lowest affinities in magenta, orange, yellow and firebricks. The calculated affinities are listed on the right.

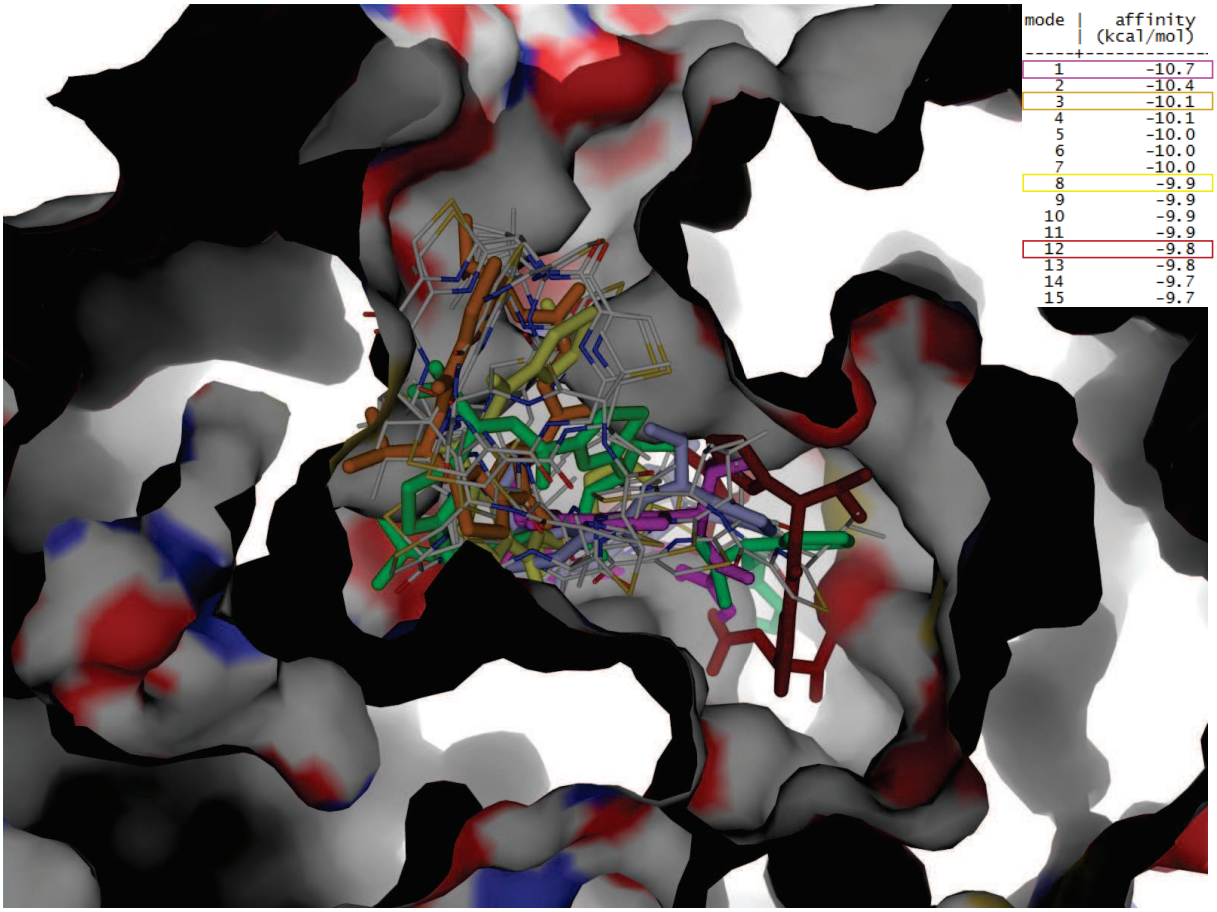
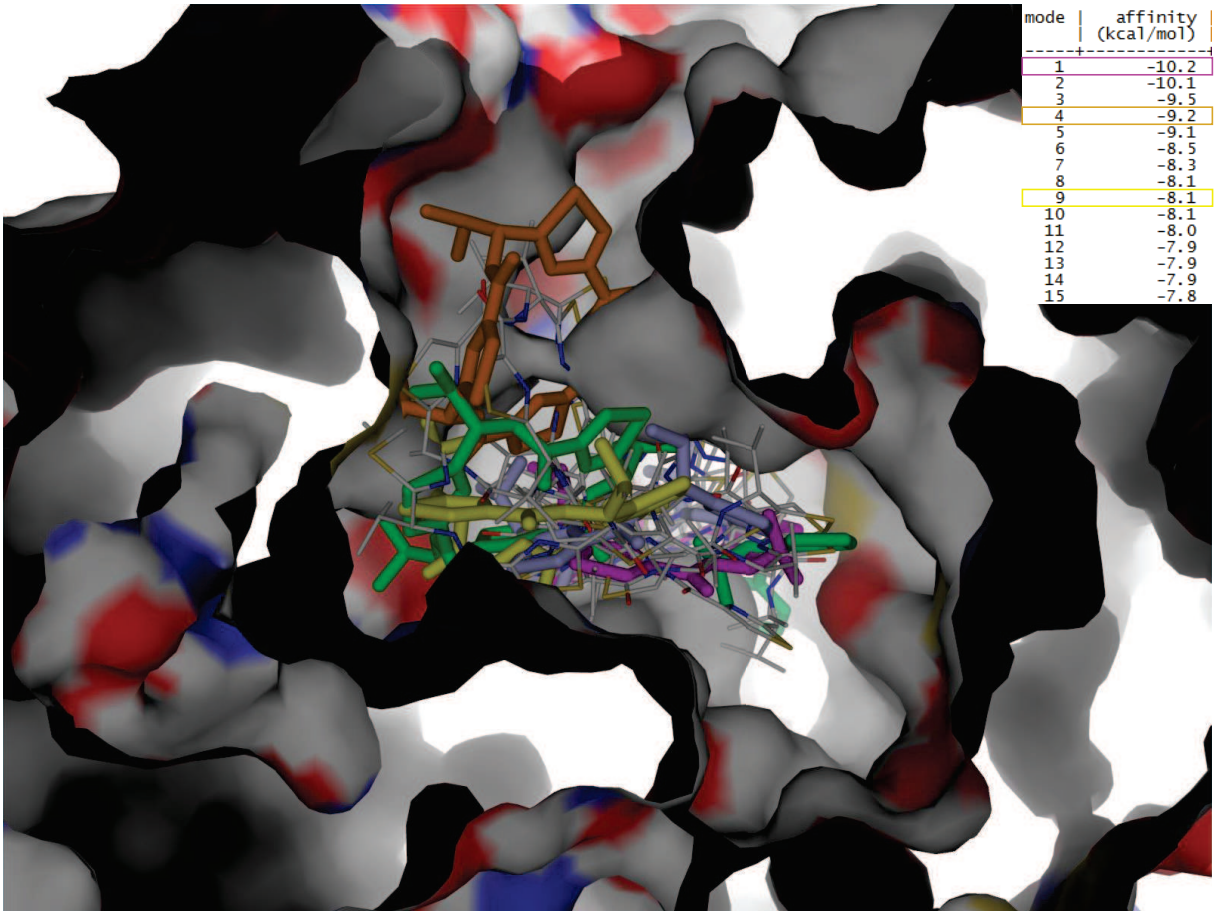


Figure S6B. QZ59-*RRR* docking on the mouse P-gp x-ray structure. Docking experiments were performed as in Fig. S3A. The 15 poses of the docked QZ59-*RRR* are shown in lines and CPK. The 3 main locations are shown in sticks, colored from the highest to the lowest affinities in magenta, orange and yellow. The calculated affinities are listed on the right.



Movie S1. Conformational changes of the mouse P-gp visualized by morphing. Inward-facing mouse P-gp x-ray structures PDB codes 4KSC, 4KSB and 4LSG were aligned on their N-ter transmembrane domain moiety and subjected to morphing with Pymol 1.6. The movie displays a general view, then focused on the drug-binding pocket, in ribbon and then in surface modes, colored in rainbow, from blue for the N-ter to red for the C-ter. Hoechst 33342 and daunorubicin are shown as docked in their best locations of highest affinity. The second part of the movie displays a surface view of the drug-binding pocket, showing the conformational changes of residues forming the H-site (cyan), the R-site (red) and those common to both sites (violet).

References

1. Loo, T. W. & Clarke, D. M. (2002) Location of the Rhodamine-binding Site in the Human Multidrug Resistance P-glycoprotein, *Journal of Biological Chemistry.* **277**, 44332-44338.

2. Ward, A. B., Szewczyk, P., Grimard, V., Lee, C. W., Martinez, L., Doshi, R., Caya, A., Villaluz, M., Pardon, E., Cregger, C., Swartz, D. J., Falson, P. G., Urbatsch, I. L., Govaerts, C., Steyaert, J. & Chang, G. (2013) Structures of P-glycoprotein reveal its conformational flexibility and an epitope on the nucleotide-binding domain, *Proc Natl Acad Sci U S A.* **110**, 13386-91.

3. Aller, S. G., Yu, J., Ward, A., Weng, Y., Chittaboina, S., Zhuo, R., Harrell, P. M., Trinh, Y. T., Zhang, Q., Urbatsch, I. L. & Chang, G. (2009) Structure of P-Glycoprotein Reveals a Molecular Basis for Poly-Specific Drug Binding, *Science.* **323**, 1718-1722.

4. Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J. & Klenk, D. C. (1985) Measurement of protein using bicinchoninic acid, *Anal Biochem.* **150**, 76-85.

5. Li, J., Jaimes, K. F. & Aller, S. G. (2013) Refined structures of mouse P-glycoprotein, *Protein Sci.*

Conclusion

One main mechanism of drug resistance involves multi-drug efflux pumps which expel drugs out of the cells. How these proteins pump out hundreds of toxins while displaying marked specificities is mysterious. Here we describe a molecular basis of this mechanism of poly-specificity, having localized the H- and R-sites of the human P-gp. This two main sites by which drugs efflux occurs, were exemplified with Hoechst 33342 and daunorubicin.

From a technical point of view, we achieved this study by bringing together skills of biologists, crystallographers and mathematicians which allowed us to build and validate new mechanistic models. This leads to important findings, of broad interest for people interested in enzyme mechanisms and more precisely drug efflux, bridging cellular effects to structural data. Beyond, we are convinced that the approach and tools developed here will be useful to characterize other ABC transporters (e.g., ABCG2) allowing the evaluation of the impact of such topologies on drug-transport.

From docking studies with the mouse P-gp structure (PDB code: 4LSG), it was possible to identify the lining residues for each site. The H and R sites have about half of specific residues and half of commons ones. R site specific residues are mainly located in the N-terminal halve, M68 and T75 in TM1, I336 and F339 in TM6, Y949 and F953 in TM11 and F974 in TM12. Residues specific for the H site are mostly located on the C-terminal halve, M295, F299, I302 and Y303 in TM5, L328 and T329 in TM6, Q721 in TM7, F979 in TM8, and M982 and Q986 in TM12. Common residues are mostly located close to the most buried QZ59-SSS location: F71 in TM1, Y306 in TM5, F332 in TM6, F724 and F728 in TM7 and L971 in TM12. Among the residues proposed here to participate to the R site, I336, F339 and L971 were previously mapped by Loo and Clarke, using the rhodamine analog, rhodamine-MTS [233] the other mapped residues, I977 and V978, although not included in the present selection remain quite close. Concerning the H site, F299, Y303 and Y306 were also previously proposed to be part of the H site [235].

Another observation is that these sites exist in P-gp in a given time, as shown in the Video S1 (supporting information of publication III), which put all the known P-gp structures solved until today and the R and H drug-binding site locations that we propose.

Two papers were published recently making allusion at the putative H and R drug-binding site locations in the (mouse or human) P-gp. Both of them used Docking approaches. The first used AutoDock VINA [287], while the second used MOE and the GOLD suite [288]. Despite using the same docking program (Autodock VINA), our results contradict those found by Ferreira *et al.*, while they are close to those from Pajeva *et al.*, (Figure 72). However it should be noted that we bring here direct information by using the QZ59 inhibitors (for which the location is known through their co-crystalization with the pump) in competition with Hoechst 33342 or daunorubicin.

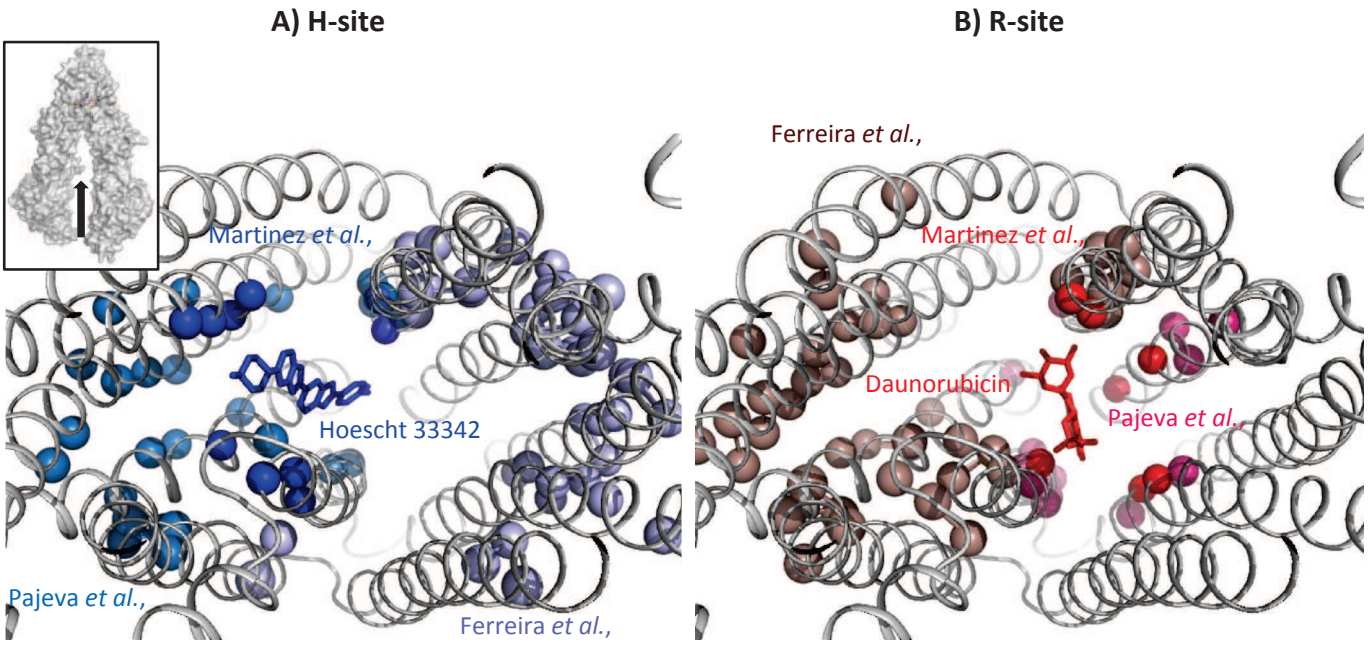


Figure 72. Mapped residues of H- and R- binding sites in the mouse P-gp (PDB code: 4LSG).
The H site is displayed in left **(A)** and the R site in the right **(B)**.
The Hoescht 33342 is colored in blue and the daunorubicin in red. Residues proposed by Martinez *et al.*, are colored in blue (H-site) and red (R-site). Residues proposed by Ferreira *et al.*, are colored in lightblue (H-site) and darksalmon (R-site). Residues proposed by Pajeva *et al.*, are colored in skyblue (H-site) and warmpink (R-site).

CONCLUSION AND FUTURE DIRECTIONS

The P-glycoprotein (P-gp or ABCB1), multidrug resistance-associated protein (MRP1 or ABCC1) and breast cancer resistance protein (BCRP or ABCG2) are membrane proteins ubiquitously produced by cells. Their main function is to reduce the concentration of cytotoxic compounds (such as anticancer, antiviral, antifungal and antibacterial drugs) by using the energy from ATP binding and hydrolysis. These proteins cover a critical field in drug disposition and drug resistance to chemotherapeutic treatments, for which solutions remain to be found.

During these three years of PhD, I tried to make a contribution to basic and applied understanding of the functional mechanisms of P-glycoprotein. I focused in one strategy to fight against chemoresistance due to the activity of P-gp (chapter I). I also was interested to find key aspects of the mechanism of P-glycoprotein by trying to bring new structural "snapshots" of P-gp (chapter II) and by locating two drug-binding sites into P-gp (chapter III).

One strategy to overcome P-gp activity: dual-function molecules

For many years the main strategy to overcome P-gp-mediated resistance was the development of an effective, specific and not toxic inhibitor to block its transport function. However, since P-gp is expressed in non-target tissues to protect them, its inhibition plus the anticancer drug may lead to increased toxicity. This was the scenario for most clinical trials, which were ended up due to the intolerable side effects. Thus, other approaches (design molecules as no substrate, multifunctional drugs, the "Achilles' heel" or collateral sensitivity, regulation of P-gp transcription or translation) have been developed. A combination of them with delivery systems (e.g., nanoparticles, liposomes, micelles) may enhance the efficacy of chemotherapy avoiding the toxicity.

Related to this, dual agents to overcome HIV resistance due to P-gp efflux have been described [279, 280]. In our laboratory, dipeptide analogs derivated from the reversin 121 (CT1300 to CT1361 molecules) were previously reported as potent P-gp inhibitors [257]. Since their chemical structure shares certain similarities to current protease inhibitors, we hypothesized that these compounds might have an inhibitory action against HIV-1 protease. To evaluate our hypothesis, *in-silico* docking experiments of these compounds were carried out into the HIV-1 protease by using Sybyl software. According to their score, the CT1347 was ranked in the best binding affinity group for HIV-1 protease, while the CT1353 was ranked in a group of poor efficacy. As scoring functions are generally poorly efficient to discriminate active from inactive compounds (false positive or negative compounds), we analyzed a second time by AuPosSOM (Automatic analysis of Poses using SOM). This program uses all the ligand poses between the molecule and the target to classify the molecules into the same interatomic contacts cluster. Surprisingly, the CT1347 was clustered in the poor efficient compounds group, while the CT1353 in the one that could inhibit the activity of HIV-1 protease. However, cell based and FRET assay showed that these compounds cannot inhibit the HIV-1 protease activity. This lack of correlation between *in silico* and *in vitro* results is often described in the literature. The use of others docking programs (e.g., AutoDock-vina) could help to

establish a more reliable *in silico* protocol to screen other potent compounds in the laboratory library.

In parallel, a similar strategy was implemented. New HIV-1 protease inhibitors (compounds numbered 36 to 43, publication I) developed by Dr. J. Hasserodt (ENS-Lyon), were tested as inhibitors or substrates of P-gp and ABCG2. To evaluate their capacity to inhibit these transporters, mitoxantrone accumulation in either controls (NIH3T3 or HEK293 pcDNA3) or expressing P-gp or ABCG2 cells lines (NIH3T3-P-gp or HEK293 pcDNA3 ABCG2) was measured by flow cytometry. Results did not show any prevention of the mitoxantrone efflux by P-gp or ABCG2, suggesting that these compounds probably are not inhibitors of such transporters. To determine if these compounds are translocated out of the cells by P-gp aor ABCG2, their toxicities in the previous cells estimated by the MTT assay was used. The expression of P-gp or ABCG2 did not change the effective concentrations (EC_{50}) of these compounds, suggesting that they are not transported by these pumps. Thus, these compounds are not dual agents but they are attractive candidates because they are not recognized by P-gp or ABCG2 transporters. Further refinement of these molecules to improve their inhibition potential into the nanomolar range is needed. To this end, X-ray structures of HIV-1 protease in complex with the inhibitor may furnish the molecular insights to refine them. Also, structures of P-gp or ABCG2 in complex with the inhibitor may lead to develop safe and effective dual agents acting as inhibitors of HIV replication and resistance mediated by P-gp or ABCG2.

X-ray structures of P-gp

Up to date, several structures of P-gp were solved at resolution range of 3.4-4.4 Å. All of them in an inward-facing configuration (nucleotide binding domains are distant and transmembrane segment are facing toward the cytoplasmic side) and only two in complex with an inhibitor (QZ59 seleno-enantiomers). However, structures of P-gp in a pre- and post-hydrolytic conformations together with substrate bound are still missing to elucidate its mechanism.

Co-crystallization of P-gp with its substrates (daunorubicin, hoechst33342, mitoxantrone and rhodamine123), its inhibitors (CT1347, CT1357, CT1336, CT1364…) and HIV-1 protease inhibitors (dauronavir, amprenavir, indinavir, and the last protease inhibitors: compounds numbered 36 to 43) were realized in collaboration with Professor Geoffrey Chang (UCSD, San Diego California, USA). Despite successful crystal formation, the resolution of the trials (> 6 Å) was not enough to solved the structure and localize the compounds. This is a common scenario in membrane protein crystallization, due to the difficulty to find optimum conditions yielding good diffraction.

In all P-gp structures the linker region (residues 627-683) which connect the first nucleotide binding domain (NBD1) with the transmembrane segment seven (TM7), was not resolved, suggesting that these residues are highly flexible. We hypothesized that a construction of P-gp without its linker (named P-gp-Linkerless) might have a better resolution. This construction was also used to trap P-gp in a pre- and post-hydrolytic conformation by preventing either ATP hydrolysis

(non-hydrolysable ATP-analogues, such as AMPPNP) or phosphate release (sodium orthovanadate Na_3VO_4, traps an ADP•Pi•Vi complex). Crystals in most conditions (P-gp- Linkerless alone, in presence of either AMPPNP or ATP + Na_3VO_4) were obtained. Although they didn't diffract well (about 8-10 Å) the experiments are encouraging because the condition with AMPPNP has a different space group (Monoclinic, P2) compared to the P-gp- Linkerless alone (Orthorhombic, $P2_12_12_1$). This indicates that the P-gp-Linkerless conformation is different when in complex with AMPPNP. A further crystallization screening (other PEG, salts, conditions) of P-gp-Linkerless with AMPPNP may allow to get a new conformation of P-gp.

In order to validate the mouse P-gp structure, mercury was used to label selected residues throughout the transmembrane helices and to localize them unambiguously. The position of 24 residues confirm the registration of amino acids in the experimentally electron density map of mouse P-gp. This series of crystallography experiments have led us to solved, by molecular replacement using the validated model, three new inward-facing conformations all wider than those from 2009 (publication II). One of the structures was in complex with a camel antibody (Nb592), bound to NBD1. This Nb is a good inhibitor of the ATP hydrolysis of P-gp.

Interestingly, arginine-scanning mutagenesis by Loo and Clarke showed that homology models of human P-gp based on these crystal structures of mouse P-gp are consistent with orientation of transmembrane segment five and nine (TM5 and TM9) [289]. While the same orientation in the originals mouse P-gp models were incompatible.

Recently, structures of mouse P-gp from 2009 were corrected [169] and according to the authors these new mouse structures have discrepancies in other regions not validated by mercury-labelled cysteines (coupling helix 1, TM8, TM12 and the N-terminal elbow helix).

It's seem that as happened before, these structures start to trigger a debate in the scientific community. Further high resolution structure of P-gp could give enough detail to build an accurate model fitting functional and biochemical data on P-gp.

Localization of two drug binding sites in P-gp

Many biochemical experiments argue that P-gp has a single large pocket containing discrete sub-sites to accommodate different types of substrates. Among these, the R- and H-sites, respectively binding anthracyclins and Hoechst 33342 related compounds, have been more rigorously investigated. But, the direct localization of these drug-binding sites are still unknown, due to the fact that no structure with co-crystallized substrates is available.

The mouse ortholog was crystallized in complex with RRR and SSS enantiomers of the cyclic selenohexapeptide QZ59. They are inhibitors of the calcein and colchicine P-gp-mediated transports [166]. Since their positions are unambiguously located into P-gp thanks to the three selenium atoms, we hypothesized that they may share one or several drug transport sites. To evaluate our hypothesis inhibition models (competitive, non-competitive, mixed…) of these inhibitors with Hoechst 33342

(representing the H-site) and daunorubicin (representing the R-site) were carried out. Results show distinct mechanisms of inhibition by the two enantiomers. The RRR enantiomer was found mainly non-competitive of Hoechst 33342 and daunorubicin transport. Thus, QZ59-RRR inhibited drug efflux mostly through an inhibitory site distinct to the R- and H- transport sites. While the SSS enantiomer was found to compete effectively with both substrates, indicating that QZ59-SSS shares both sites. On the other hand, at higher concentrations of inhibitor and substrate, other effects such as an improvement of daunorubicin-mediated self-inhibition by both enantiomers and an activation of Hoechst 33342 efflux by the SSS enantiomer were found. In order to provide a clearer picture about the location of these two sites, these finding together with docking experiments were carried out. Thus the R-site is surrounded by residues M67, M68, I70, M74, Y113, F728, M945, Y949, A950, I977, V978. While the H-site by L300, Y303, Y306, L335, F339, F924, F928.

It seems also that these sites exist in P-gp at a given time, as shown in the Video S1 (supporting information of publication III), which displays the H- and R- sites in the conformational change from the 2009 P-gp structures (PDB codes: 3G61, 3G5U) up to the mouse 2013 structures (4KSB, 4KSC). This seems reasonable in a context of drug translocation.

To conclude

Research in multidrug efflux pumps field, in the next years, is expected to reach a stage where the mechanism by which these proteins translocate a wide variety of molecules out of the cell would be well understood. To accomplish this, new biochemical and biophysical data combined with higher resolution structures of ABC exporters trapped in different conformations, either with or without substrate, are needed. At this stage, it will be interesting to witness how this wealth of knowledge will be translated to circumvent multidrug resistance in the clinic.

RÉSUMÉ FRANÇAIS

La capacité des cellules cancéreuses, des bactéries ou des virus à survivre en présence d'une large gamme de molécules cytotoxiques représente un problème de santé publique qui devient de plus en plus important dans le traitement de diverses maladies (cancer, infections virales ou bactériennes). Cette résistance est parfois trouvée pour des agents auxquels les cellules n'ont jamais été exposées : c'est ce que l'on appelle une résistance « intrinsèque ». D'autres développent une résistance à un agent auquel ils s'étaient révélés initialement sensibles : c'est ce que l'on appelle la résistance « acquise », et même à d'autres composés auxquelles elles n'ont jamais été exposées et n'ayant aucune homologie structurale entre elles : phénotype appelé MDR pour « Multi-Drug Resistance ».

Parmi les différents mécanismes de résistance, les transporteurs ABC tels que la P-glycoprotéine (P-gp), les protéines de résistance multi-drogues (MRPs), la protéine de résistance multi- drogues du cancer du sein (ABCG2), ont été désignées comme acteurs potentiels. En effet, la fonction principale de ces protéines membranaires est d'évacuer hors de la cellule une grande variété de molécules possèdent de structures et de compositions chimiques très différentes. De par leur expression au niveau de différentes barrières physiologiques (intestins, foie, barrière hémato-encéphalique, placenta,…), elles diminuent la concentration du médicament dans l'organisme. De plus, ces transporteurs sont exprimés dans les cellules cibles (cancer, bactérie et lymphocytes) conduisant à des concentrations intracellulaires inefficaces de l'agent cytotoxique. Ainsi, l'étude fonctionnelle de ces transporteurs présente un intérêt double, tant du point de vue fondamental que du point de vue des applications cliniques.

Durant les trois années de ma thèse, j'ai particulièrement étudié P-gp et essayé d'apporter une contribution à la compréhension de sa fonction autour de trois axes, qui ont fait l'objet de publications scientifiques et sont présentées à la fin de chaque partie. Je me suis intéressée à une stratégie de lutte contre la chimiorésistance due à l'activité de la P-gp, en développant une nouvelle classe d'inhibiteurs de la protéase du VIH-1 qui ne sont ni transportés par la P-gp ni par ABCG2 (chapitre I). Un aspect clé pour comprendre le mécanisme de la P-gp est d'identifier des nouvelles structures dans des conformations différentes, par exemples en pré- et post-hydrolyse de l'ATP et en complexe avec des substrats. Ainsi, nous avons déterminé trois nouvelles structures de la P-gp. Une d'entre-elles illustre l'interaction entre la P-gp et un nano-anticorps au niveau de la partie C-terminale du premier domaine de fixation des nucléotides, qui fige la P-gp dans une nouvelle conformation ouverte vers l'intérieur (chapitre II). Pour finir, nous avons développé une stratégie pour localiser en détail deux sites de fixation de P-gp en caractérisant les modes d'inhibition de deux inhibiteurs précédemment co-cristallisés avec la pompe (chapitre III).

Chapitre 1. Molécules à double fonction, une stratégie pour surmonter la résistance liée à l'activité de la P-gp

Un obstacle majeur à l'éradication totale du virus HIV est la présence de réservoirs du virus dans des endroits très peu accessibles. Dans le cerveau, par exemple, l'échec des nombreux agents anti-VIH, tels que les inhibiteurs de protéase, est dû à leur pénétration limitée dans la barrière hématoencéphalique [240]. Cette limitation est en partie attribuable à la présence de transporteurs ABC, notamment P-gp, ABCG2 et MRP1, qui effluent ces agents antirétroviraux. Actuellement tous les inhibiteurs de la protéase du VIH utilisés dans la thérapie contre le SIDA, HAART (« anti-retroviral therapy ») sont soit de substrats de transport soit de inhibiteurs de la P-gp (le ritonavir, l'indinavir, le saquinavir et le nelfinavir) [16, 17]. Nous développons une approche innovante qui consiste à synthétiser des molécules ayant un double effet : inhiber la protéase et inhiber la P-gp sans être transportée par la P-gp. De cette façon l'anti-protéase ne sera pas évacuée par la P-gp et pourra s'accumuler dans des zones rendu normalement inaccessibles lors que la P-gp est active.

Approche I : des inhibiteurs de la P-gp pourraient-ils aussi inhiber l'activité de la protéase du VIH-1?

La structure chimique de certains inhibiteurs de la P-gp développés au laboratoire [257] est proche de celles des inhibiteurs de la protéase du VIH-1. Ces composés sont dérivés de la réversine 121, un inhibiteur puissant de la P-gp. Ils sont spécifiques de la P-gp avec un niveau élevé d'inhibition (sub-micromolaire) et ne sont pas transportés par la P-gp (non compétitifs pour les sites de transport) [257]. Afin de déterminer si ces inhibiteurs de la P-gp pouvaient se fixer sur la protéase (et potentiellement l'inhiber), des expériences de « docking » ont été menées avec le logiciel « Sybyl molecular-modeling » (SYBYL 1.3. Tripos Inc).

Dans notre cas, le récepteur utilisé correspond à la protéase du VIH-1 (code 1MUI dans la PDB « protein data bank »). Les ligands étudiés correspondent à :

- neuf inhibiteurs décrits pour la protéase (contrôles positifs) : amprenavir (APV, code PDB : 3NU3), saquinavir (SQV, code PDB : 3DIX), daurunavir (DRV, code PDB : 3PWM), tripanavir (TPV, code PDB : 2O4P), atazanavir (ATV, code PDB : 3EKY), nelfinavir (NFV, code PDB : 3EKX), ritonavir (RTV, code PDB : 2B60), indinavir (IDV, code PDB : 2B7Z) et lopinavir (LPV, code PDB : 1MUI) ;

- sept inhibiteurs de la P-gp (possibles contrôles négatifs) : QZ59-RRR, QZ59-SSS, zosuquidar, GF120918, verapamil, tariquidar et progestérone ;

- 28 dérivés de la réversine 121 (molécules d'intérêt): molécules CT1300 à CT1361.

Comme prévu, les inhibiteurs de protéase sont dans le groupe de meilleur affinité (bleu et vert dans le Tableau 1), et la plupart des témoins négatifs dans le groupe d'affinité modeste ou mauvaise (orange et rouge dans le Tableau 1). D'après les scores obtenus (Tableau 1), de nombreux

inhibiteurs de la P-gp dérivés de la réversine 121, (CT) présenteraient une bonne affinité pour la protéase (score > 7,5).

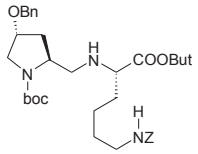
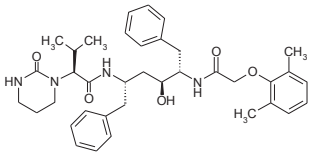
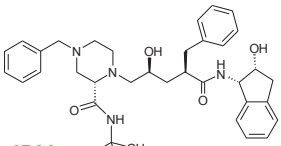
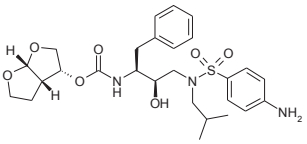
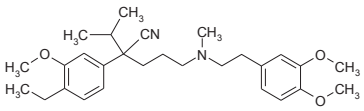
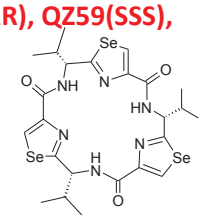
Score >10	Score 7.5-10
LPV, CT1316, CT1361, CT1326, CT1327, CT1347, CT1338, CT1344, CT1336, CT1342, CT1301, CT1337, CT1302, NFV, CT1346, CT1345, SQV, CT1339	**CT1357, TPV, CT1343, CT1329, CT1345, IDV, CT1348, DRV, RTV, CT1340, CT1354, APV, CT1341, CT1300, Zosuquidar, GF120918, CT1356, ATV, CT1355**
CT1347   **LPV**	 **IDV**  **DRV**
Score 5-7.5	**Score 0-5**
CT1333, Verapamil, Tariquidar, CT1351, CT1353, Progesterone	**QZ59(RRR), QZ59(SSS),**
Verapamil 	 **QZ59-RRR**

Tableau 1 : Résultats des expériences de liaisons de ligands *in silico*.
Les meilleures scores d'affinité pour la protéase du VIH-1 sont en bleu (scores > 10), les moins bons sont en rouge (scores entre 0 et 5).

Étant donné que la fonction score n'est pas très performante pour discriminer les composés actifs des inactifs, une seconde analyse (AuPosSOMS [283]) a été effectuée pour classer les molécules en fonction de la similitude de contact entre le récepteur et le ligand. Cette analyse (Figure 1) montre que quelques molécules CT (CT1300, CT1357, CT1327, CT1355, CT1340 et CT1345) pourraient inhiber l'activité de la protéase VIH-1 (entourés en bleu dans la Figure 1).

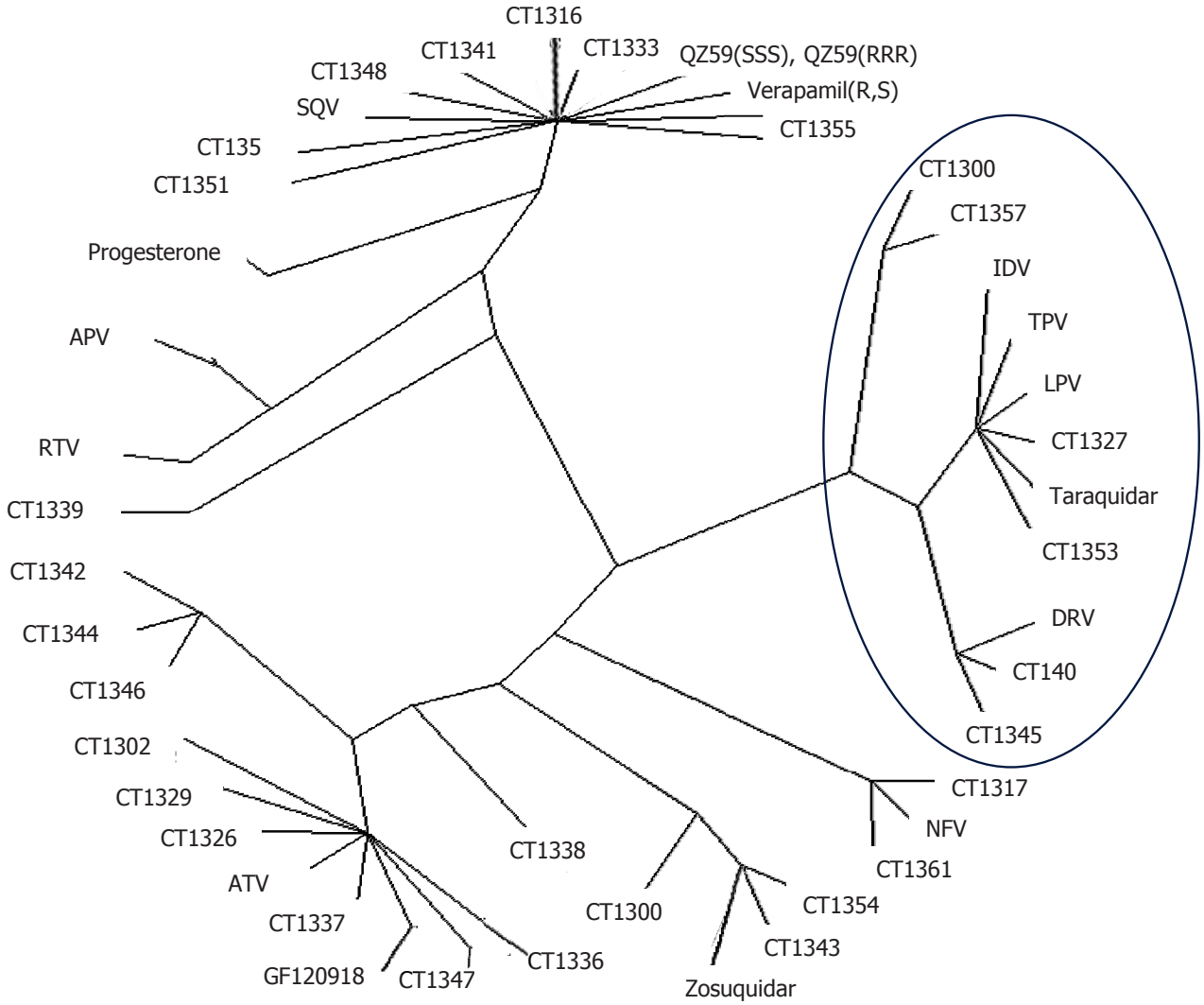


Figure 1 : Regroupement par "Automatic analysis of Poses using SOM" (AuPosSOMS) des molécules potentiellement inhibitrices de la protéase VIH-1. Les plus efficaces sont entourées en bleu

L'efficacité prédite pour les CT diffère selon la méthode d'analyse. Nous avons donc testé *in cellulo* deux molécules (CT1347 et CT1353) dont les prédictions se contredisent en fonction de la méthode d'analyse.

Afin de tester la capacité de ces composés à inhiber la protéase, un test cellulaire a été mis en place en collaboration avec le Dr. R. Wolkowicz (UCSD, USA) [254]. Pour cela, l'activité de la protéase est mesurée dans son milieu naturel, les cellules lymphocytaires. La protéase du VIH-1 est produite, dans ces cellules, en fusion entre les deux domaines de la protéine Gal4 : «N-terminal DNA-binding domain» (DBD) et «C-terminal transactivation domain» (TAD). En présence de doxycycline et d'un inhibiteur très peu actif, la protéase protéolyse la fusion. Ainsi, elle détache le domaine DBD du domaine TAD de Gal4. Or, dans ce système, l'intégrité de Gal4 est indispensable à l'expression de la «Green Fluorescente Protein» (GFP). En conséquence, les cellules GAL4/PR en présence de doxycycline et d'un inhibiteur actif, expriment la GFP car la protéase est inhibée et ne

peut pas s'auto-extraire. Ceci a été observé pour les inhibiteurs de référence, l'indinavir et le saquinavir (respectivement au debout cercle rouge et vert pâles dans la Figure 2). À l'inverse, les cellules GAL4/PR en présence de doxycycline et d'un inhibiteur peu actif, n'ont pas d'expression de la GFP. C'était le cas des composés CT1347 et CT1353 (respectivement au debout cercle orange et vert dans la Figure 2), ce qui suggère qu'ils ne sont pas capables d'inhiber la protéase dans cet état.

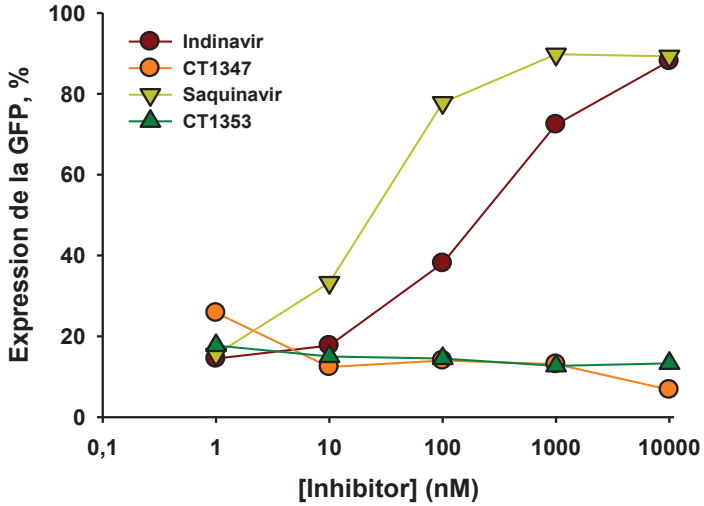


Figure 2 : Inhibition de la protéase HIV-1 par le CT1347 et CT1353. L'indinavir et le saquinavir ont été utilisés comme contrôles positifs.

Cette absence d'inhibition est peut-être due à leur manque d'efficacité; elle peut être aussi due au fait que ces composés sont relativement hydrophobes et s'accumulent dans la membrane plasmique. Ainsi ils ne peuvent pas atteindre la protéase du VIH-1, qui est présente dans le cytoplasme. Ainsi, nous avons testé l'activité inhibitrice de ces composés *in vitro* par « fluorescence resonance energy transfer » (FRET). Cette méthode utilise un substrat marqué avec un donneur de fluorescence et un accepteur non fluorescent. Lorsque cette molécule est entière, l'accepteur est suffisamment proche pour atténuer l'émission de fluorescence du donneur. À l'inverse, lorsque cette molécule est clivée par la protéase du VIH, l'effet d'atténuation est perdu (Figure 3A). Ainsi, la protéase du VIH en présence du substrat fluorogène génère la plus forte fluorescence. L'addition d'un inhibiteur efficace diminue la fluorescence. Comme prévu, les inhibiteurs de protéase (pepstatine, indinavir et saquinavir) bloquent le clivage du substrat par la protéase du VIH (diminution de fluorescence sur la Figure 3B). En présence de nos composés, la fluorescence n'a pas changé, ce qui suggère que ces composés n'inhibent pas la protéase aux concentrations testées (Figure 3B).

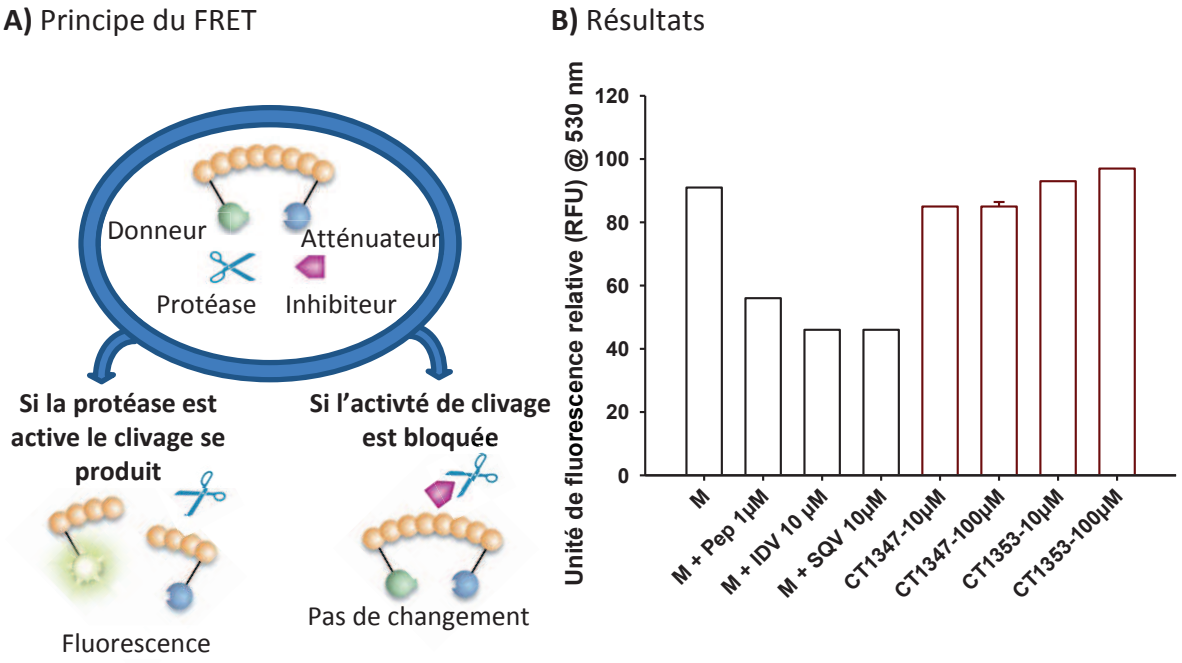


Figure 3 : Inhibition de composés par dosage FRET.
Excitation/émission = 490 nm/530 nm, en utilisant un spectrofluorimètre pour microplaques (Tecan Infinite M1000). M: mélange réactionnel (tampon + VIH-1 protéase + substrat); Pep: pepstatine; IDV: indinavir; SQV: saquinavir

Les tests réalisés i*n cellulo* et *in vitro* ont donc montré que les composés CT1347 et CT1353 ne sont pas capables d'inhiber l'activité de la protéase du VIH-1. Ce manque de corrélation entre les résultats *in silico* et *in vitro* est souvent décrit dans la littérature. L'utilisation d'autres programmes de tests de liaison de composés *in silico* (par exemple, AutoDock-vina) pourraient aider à établir une meilleure relation entre prédictions et observations pour cribler d'autres composés.

Approche II : les inhibiteurs de la protéase du VIH-1 pourraient-ils aussi inhiber la P-gp ou ABCG2 ?

En parallèle, nous avons profité de la mise en œuvre de ces techniques pour tester de nouveaux inhibiteurs de protéase du VIH-1 (collaboration avec le Dr J. Hasserodt, ENS-Lyon) en tant qu'agents inhibiteurs ou substrats de la P-gp ou ABCG2. Ces molécules (numérotées de 36 à 43) comportent un motif hétérocyclique uréique [286] présentant l'avantage de ne pas se baser sur un motif peptidique susceptible à générer des résistances. Elles inhibent la protéase par des concentrations de l'ordre de 30 µM (Figure 4).

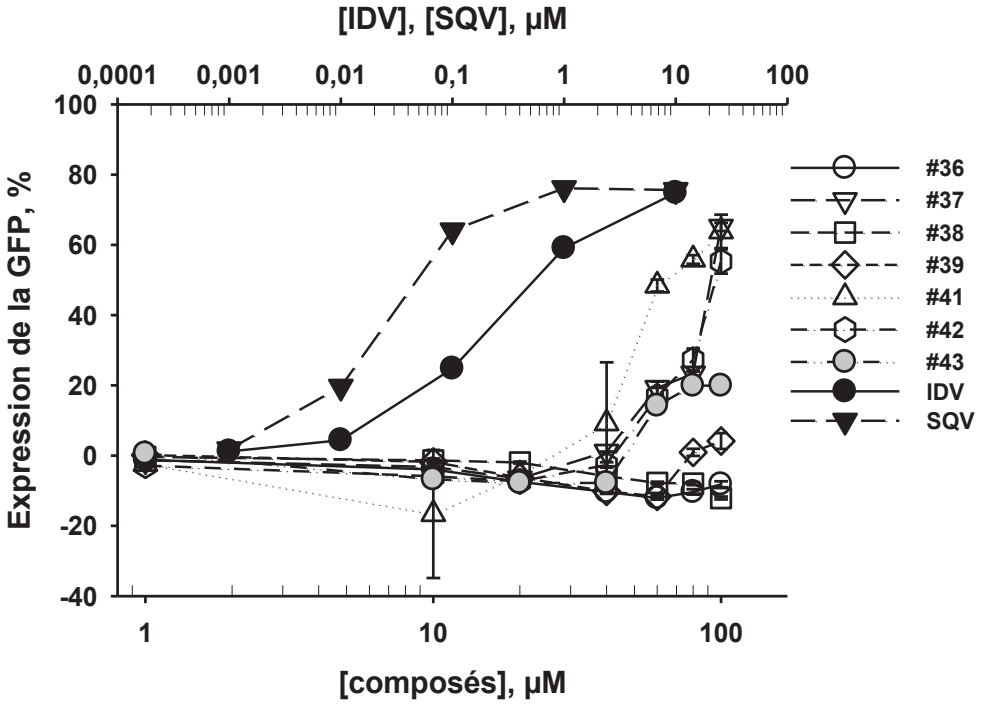


Figure 4 : Inhibition de la protéase HIV-1 par les composés 36 à 43. L'indinavir (IDV) et saquinavir (SQV) ont été utilisés comme contrôles positifs.

Pour déterminer si ces composés sont expulsés hors de la cellule par la P-gp ou ABCG2, leurs toxicités dans les cellules contrôles (NIH3T3 ou HEK293 pcDNA3) ou dans les cellules exprimant la P-gp ou ABCG2 (NIH3T3-P-gp ou HEK293 pcDNA3 ABCG2) ont été estimées à l'aide du test MTT. L'expression de la P-gp ou ABCG2 n'a pas modifié les concentrations efficaces demi-maximales (EC_{50}) de ces composés, ce qui suggère qu'ils ne sont pas transportés par ces pompes (Tableau 2).

Composé	P-gp		ABCG2	
	Cellules exprimant	Cellules contrôle,	Cellules exprimant	Cellules contrôle
36	100±1,8	68,8±2,1	57,2±0,8	56,3±0,9
37	48,6±0,8	42,6±0,9	34,8±2,1	37,6±1,1
38	---	79,8±1,2	---	---
39	71,2±1,7	58,6±1,6	47,9±1,6	52,4±0,4
41	43,5±0,7	42,1±1,5	23,1±1,1	35,3±1,5
42	46,8±0,9	39,4±0,7	37,2±0,9	42,6±0,7
43	12,8±0,4	11,2±0,2	12,8±0,2	17,6±0,4

Tableau 2 : Cytotoxicité des composés en µmoles.

Pour évaluer leur capacité à inhiber ces transporteurs, l'accumulation de la mitoxantrone dans les cellules a été mesurée par cytométrie en flux. Les pourcentages d'inhibition des composés contrôles (GF120918 pour P-gp et Ko143 pour ABCG2) ont montré une inhibition de l'efflux de mitoxantrone par la P-gp ou ABCG2 d'environ 100 % à faibles concentrations (2 µM). Par contre, les

composés 36 à 43 n'ont montré aucune prévention de l'efflux de mitoxantrone par la P-gp ou ABCG2, ce qui suggère que ces composés ne sont probablement pas des inhibiteurs de ces transporteurs (Tableau 3).

Composé	% d'inhibition de l'efflux par la P-gp à		% d'inhibition de l'efflux par ABCG2 à	
	2 µM	10 µM	2 µM	10 µM
36	2,3±0,5	8,1±0,3	0,2±0,03	8,1±0,3
37	5,5±0,5	6,4±0,7	0,2±0,05	0,3±0,01
38	7,1±0,3	11,5±1,2	0,2±0,02	3,3±1,7
39	1,1±0,1	3,0±0,7	0,9±0,6	2,5±0,9
41	4,6±0,2	4,8±0,1	2,5±0,6	8,4±1,1
42	7,7±0,1	22,7±0,5	0,9±0,5	8,1±0,3
43	1,8±0,2	15±2,0	1,8±0,2	8,8±1,1
GF 120918	110±2,5			
Ko143			100±1,5	

Tableau 3 : Pourcentage d'inhibition de l'efflux de mitoxantrone par la P-gp ou ABCG2.

Ces composés ne sont pas capables d'inhiber à la fois la protéase de VIH-1 et le transporteur, mais restent très intéressants car ils ne sont pas reconnus par les transporteurs P-gp ou ABCG2. L'optimisation de ces molécules pourrait améliorer leur potentiel d'inhibition. À cette fin, des structures aux rayons-X de la protéase du VIH-1 en complexe avec ces inhibiteurs pourraient fournir les éléments moléculaires pour les perfectionner. Également, les structures de la P-gp ou ABCG2 en complexe avec ces inhibiteurs pourrait conduire à développer efficacement ces agents doubles, agissant comme inhibiteurs de la protéase du VIH et réversant la résistance liée aux transporteurs P-gp ou ABCG2.

Ces résultats font l'objet d'une publication dans le Journal *Bioorganic & medicinal chemistry* **2013**, *21* (17), 5407-5413 (Publication I).

Chapitre 2. Étude structurale de la P-gp

La P-gp est une protéine de 170 kDa composée de deux moitiés pseudo-symétriques, chacune contenant un domaine de liaison et de hydrolyse de l'ATP appelé «Nucleotide Binding Domain» (NBD) et un domaine transmembranaire appelé «transmembrane domain» (TMD) intervenant dans la reconnaissance des substrats et constitué de six segments de type hélices alpha. Afin d'élucider le mécanisme par lequel ce transporteur évacue les molécules à travers la membrane, le structure des transporteurs ABC en complexe avec des substrats de transport ainsi que leurs intermédiaires conformationnels sont nécessaires.

Partie I : P-gp en complexe

Afin de cristalliser la P-gp, plusieurs substrats (daunorubicine, Hoechst 33342, mitoxantrone et rhodamine 123), des inhibiteurs de notre laboratoire (CT1347, CT1357, CT1336, CT1364 ...) et des inhibiteurs de la protéase (PI) du VIH ont été utilisés (dauronavir, amprénavir, indinavir et le nouveau PI de la publication I: composés 36 à 43). L'expression, la purification et l'état de cristallisation ont été appliqués comme décrit dans la section Matériels et Méthodes.

Parmi toutes les conditions testées pour chaque composé (0,1 M Hepes pH 7-8, 50 mM de sulfate de lithium, 10 mM EDTA, et 24 à 29,5 % w/v PEG 600 Da à 4 °C) les meilleurs sont présentées dans la Figure 5. La vaste majorité des cristaux ont été obtenus à pH: 7-8 et PEG 600: 25,5 à 28 %.

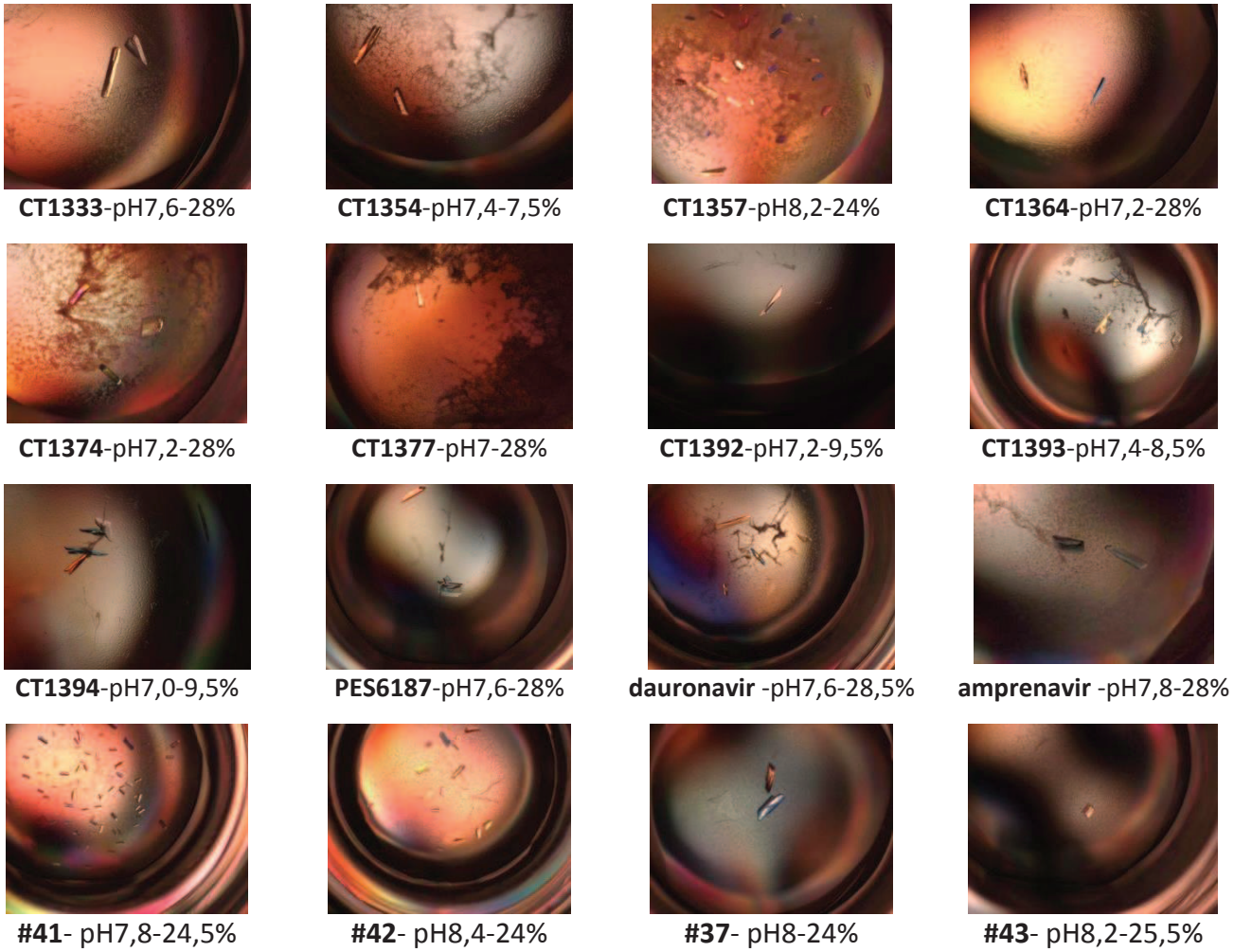
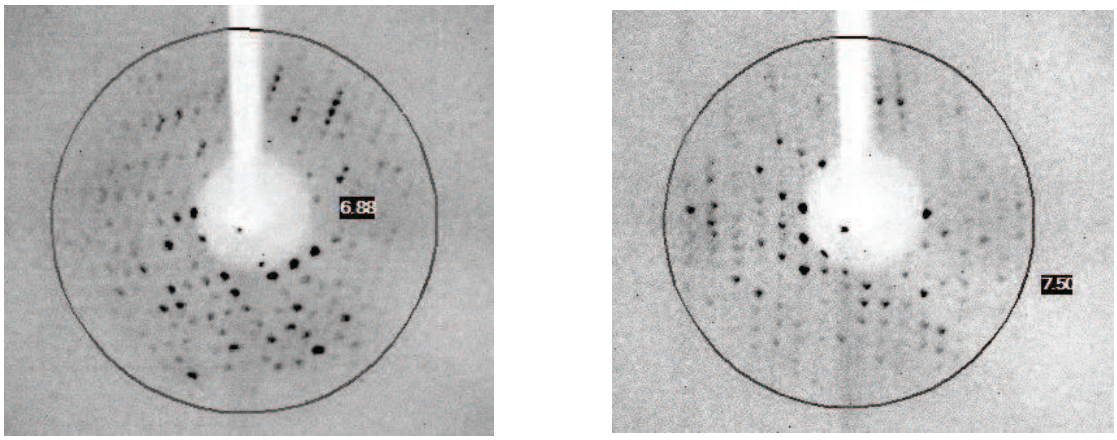


Figure 5 : Cristaux obtenus,
Nom du composé avec sa meilleure condition de cristallisation (pH et % de PEG 600 Da),

Malgré la formation de cristaux avec des morphologies différentes (Figure 5), la résolution toujours supérieure à 6 Å, Figure 6, de ces cristaux est insuffisante pour localiser les composés. En effet, une résolution comprise entre 3 et 6 Å peut en effet conduire l'expérimentateur à commettre des erreurs dans l'interprétation de la carte de densité électronique.



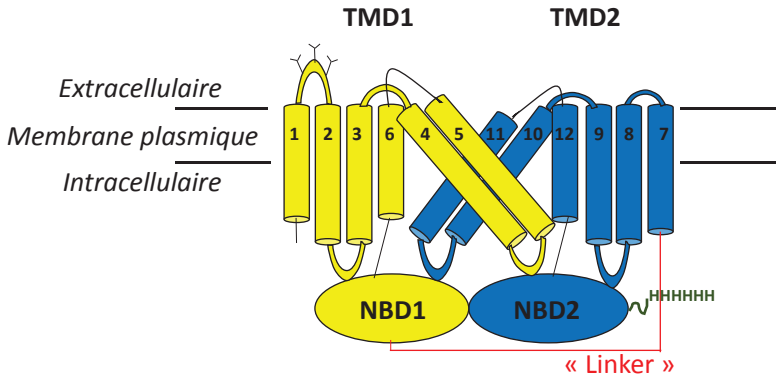
CT1394-pH 7, 29,5 % PEG 600 **CT1374**-pH 7,0, 29,5 % PEG 600

Figure 6 : Clichés de diffraction de CT1394 et CT1374, obtenus à l'aide du diffractomètre « D8 VENTURE crystal X-ray diffraction –Bruker ».

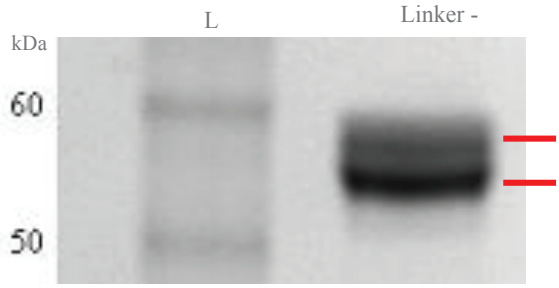
Partie II : la P-gp dans la conformation orientée vers l'extérieur

Une construction de la P-gp présentant une délétion de 40 acides aminés (illustrée par la ligne rouge sur la Figure 7A) entre NBD1 et TM7 (Dr. Ina Urbatsch, TTUHSC, Texas, USA), nommée « P-gp-Linkerless » a été conçue. Après purification, la migration de cette protéine sur SDS PAGE coloré au bleu de Coomassie (Figure 7B) révèle la présence de deux bandes (au lieu d'une à 144 kDa) de 55 à 58 kDa, correspondant aux deux moitiés de la P-gp. Alors que le western blot anti-histidine montre une seule bande à 144 kDa pour P-gp sauvage et de 55 kDa pour P-gp-Linkerless (Figure 7C). Ainsi, deux moitiés de la P-gp sont exprimées et seulement une moitié a l'étiquette histidine. Cette construction était supposée améliorer la résolution cristallographique car que la connexion entre les deux moitiés est probablement trop flexible. Cette construction a également été utilisée pour piéger la P-gp dans la conformation tournée vers l'extérieur «outward-facing conformation», car sans cette union les deux NBD pourront interagir plus facilement.

A) Construction P-gp-Linkerless



B) Gel SDS-PAGE 10 % coloré au bleu de Coomassie



C) Détection de l'étiquette histidine par Western blot

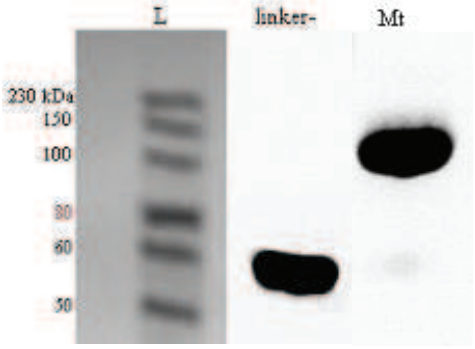


Figure 7 : « P-gp-Linkerless ».
L: marqueurs de poids moléculaire indiqués en kDa; Linker -: fraction de membrane de la construction « P-gp-Linkerless »; Mt: fraction de membrane de P-gp sauvage.

De nombreuse conditions (576) ont été réalisé pour cristalliser la construction « P-gp-Linkerless » à partir des conditions précédentes (0,1 M Hepes pH 7-8, 50 mM de sulfate de lithium, 10 mM EDTA, et 24 à 29,5 % w/v PEG 600 Da à 4 °C) en collaboration avec Rupak Doshi, Mark Villaluz et Paul Szewczyk (Tableau 4). Ces conditions utilisent soit des analogues non hydrolysables de l'ATP (par exemple, AMPPNP) ou de l'ATP en présence de magnésium pour favoriser l'hydrolyse de l'ATP. Le sodium orthovanadate (Na_3VO_4) est un analogue du phosphate qui fige un complexe ADP•Pi•Vi dans un état post-hydrolyse. De ce fait, la protéine en conformation tournée vers l'extérieur avec NBD dimérisés pourrait être obtenue. Le palmitoyl-2-oléoyl-sn-glycéro-3-phosphoéthanolamine (POPE) a été utilisé pour améliorer l'activité catalytique de la P-gp, en présence de vérapamil.

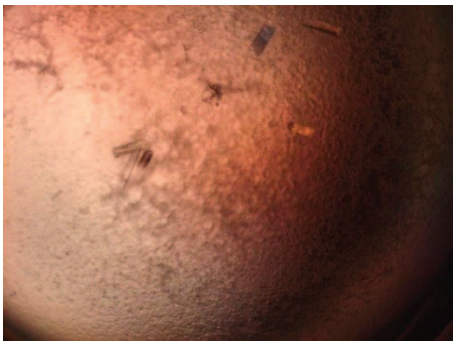
Condition	Agents de précipitation et % testés	pH	Sel
P-gp-Linkerless + $MgSO_4$ + AMPPNP	w/ 350 MME 15-30% w/ 250 DME 15-30% w/ PEG 600 15-30% w/ PEG 400 15-30% w/ PEG 1000 10-25% w/ PEG 1500 10-25% Δ %= 0,5	0,1 M Hepes pH 6-8.2, Δ pH = 0,2	50 mM of $(NH_4)_2SO_4$ or NaCl
P-gp-Linkerless + $MgSO_4$ + Sodium Orthovanadate (Na_3VO_4) + ATP			
P-gp-Linkerless + POPE + verapamil + $MgSO_4$ + Na_3VO_4 + ATP			

Tableau 4 : Tests de cristallisation de P-gp-linkerless.

POPE: palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine; 350 MME : monométhyléther polyéthylène glycol 350; 250 DME: Polyglycol DME 250; PEG 600: polyéthylène glycol 600; PEG 400: polyéthylène glycol 400; PEG 1000: polyéthylène glycol 1000; PEG 1500: polyéthylène glycol 1500;
Concentration finale: $MgSO_4$, Na_3VO_4 et ATP à 10 mM; POPE à 0.1 mg/ml; verapamil à 0.1 mM.

La plupart des conditions donnent des cristaux (P-gp-Linkerless seul, en présence soit d'AMPPNP ou en présence d'ATP + Na3VO4). Malheureusement, nous pouvions conclure que les cristaux de P-gp-linkerless en présence d'AMPPNP ont une symétrie différente (Monoclinique, P2) compare à P-gp-Linkerless seul (Orthorhombique, P212121) parce que la résolution des cristaux ne permet pas d'aller plus loin (Figure 8).

A) Cristaux obtenus



B) Cliché de diffraction

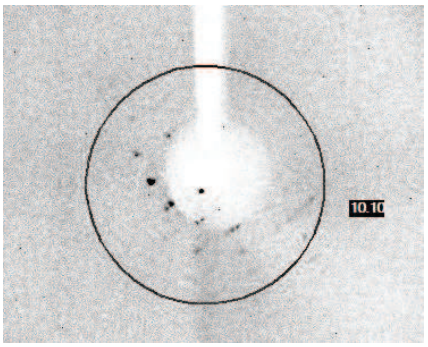


Figure 8 : Meilleure condition pour la crystallisation de la « P-gp-linkerless ».
10 mM AMPPNP, 10 mM $MgSO_4$, 50 mM $(NH_4)_2SO_4$, 18 % PEG 1500, 0.1 M HEPES pH 6,2.

Partie III : validation du modèle de la P-gp

La structure 3D de la P-gp de souris a été résolue à 4 Å, ce qui est une résolution modérée. Afin de valider cette structure, 63 mutants Cystéine de la P-gp de souris ont été exprimés, purifiés et cristallisés (collaboration Pr. Chang). Une fois que le cristal a été formé, il a été trempé dans du chlorure d'éthyl-mercure. Cette molécule produit un signal anomal (maille rouge dans la Figure 9) permettant de localiser avec une grande précision le résidu cystéine dans la structure 3D. La P-gp possède 7 cystéines (sphères vertes de la Figure 9), ainsi la présence du huitième signal est due au résidu recherché (sphère bleue de la Figure 9).

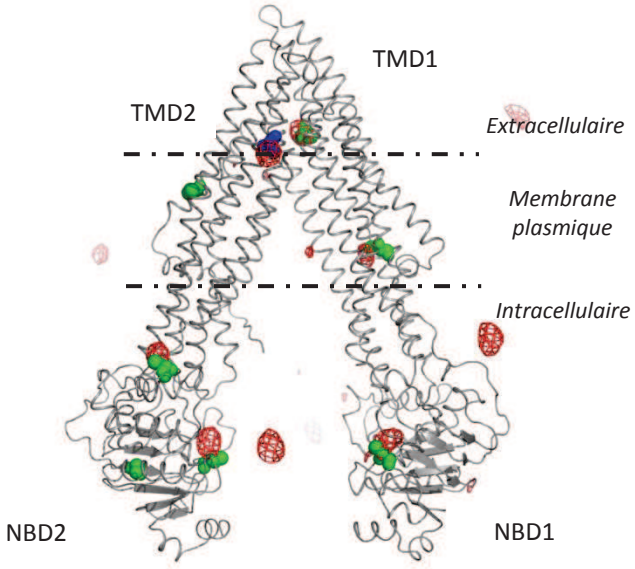


Figure 9 : Résidus cystéines révélés par le mercure.
Les barres horizontales représentent le positionnement approximatif de la bicouche lipidique. «transmembrane domain» (TMD), «Nucleotide Binding Domain» (NBD).

Les positions de 24 acides aminés ont été validées: 17 dans la publication II, soulignés en bleu sur la Figure 10 (S80C, S176C, M1888C, A216C, S218C, A250C, R272C, L283C, G284C, S305C, A309C, G342C, A344C, K730C, S876C, S889C, et S975C) et 7 dans celui publié en 2009 (C133A, C713A, C427A, C952A, C1121A, Cysless1070C, et C1223A [166]).

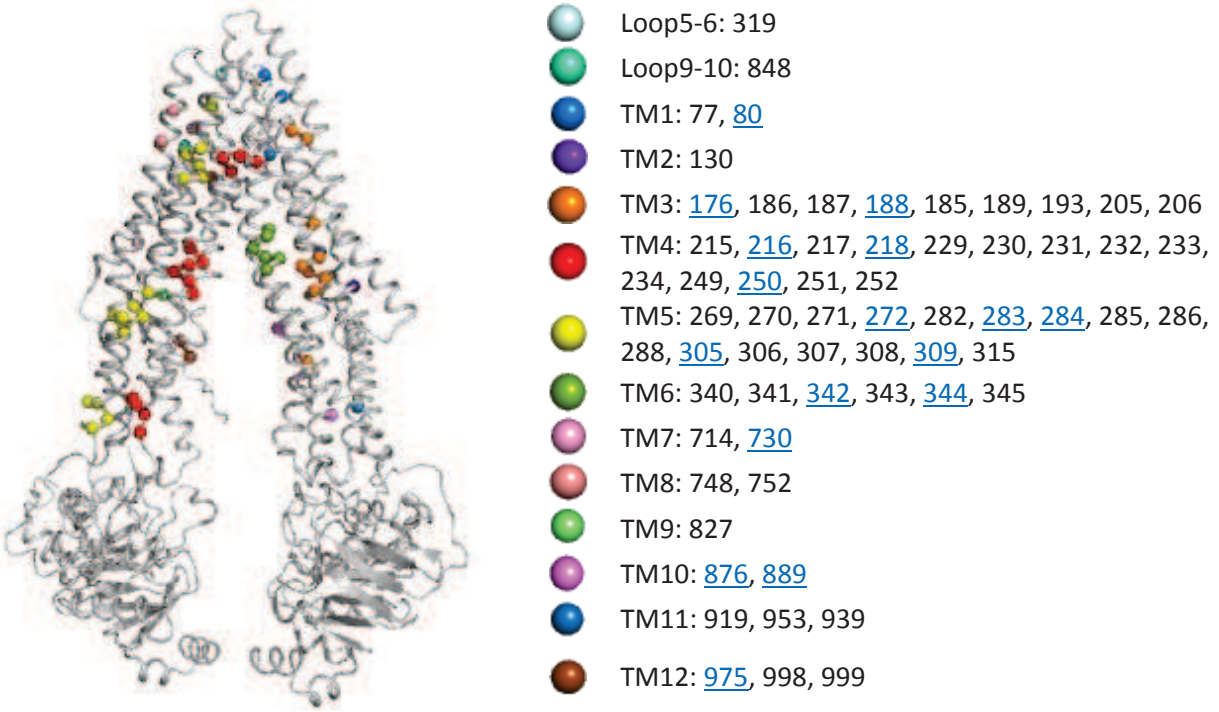


Figure 10 : Positions des acides aminés mutés validées expérimentalement dans la structure de la P-gp souris.
Les sphères représentent les atomes C-α de résidus mutés. En bleu sont montrés les résidus publiés dans l'article II.

Cette série d'expériences de cristallisation et de cristallographie a permis d'identifier trois nouvelles conformations de la P-gp de souris tournées vers l'intérieur (Figure 11A). Deux d'entre elles ont été dérivées de la même forme cristalline. En général, les topologies de ces structures sont les mêmes avec des distances entre le NBD1 et le NBD2 plus importantes (Figure 11A) que celles publiées en 2009 (Figure 11A), mais comparables à d'autres structures des transporteurs ABC publiés [167, 170, 171]. Comme dans le structure originale de la P-gp de souris, les résidus 627 à 683 entre NBD1 et TM7 «linker», importants pour la phosphorylation, ne sont pas bien observées et supposés désordonnés.

Nous avons également établi la structure de la P-gp complexée avec un nano-anticorps (Nb), appelé Nb592 (Figure 11B). Ce dernier se fixe au premier NBD. Il peut être utilisé comme un chaperon de cristallisation pour stabiliser la conformation de la protéine. Il s'agit également d'un inhibiteur très puissant de l'activité ATPasique de la P-gp qui empêche la dimérisation des NBD (Figure 3 de la publication II), essentiel pour l'hydrolyse de l'ATP et le transport des molécules. Un alignement des séquences protéiques entre la P-gp de souris et l'humaine a révélé que les deux protéines sont très semblables au niveau de l'épitope (site de liaison avec le nano-anticorps). De ce fait, ce type de molécules présente un grand potentiel thérapeutique. Des complexes innovants tels que des peptides ou nanoparticules peuvent passer à travers la membrane cellulaire pour inhiber efficacement et spécifiquement la P-gp.

Ces résultats font l'objet d'une publication dans *Proceedings of the National Academy of Sciences* **2013**, *110* (33), 13386-13391 (Publication II).

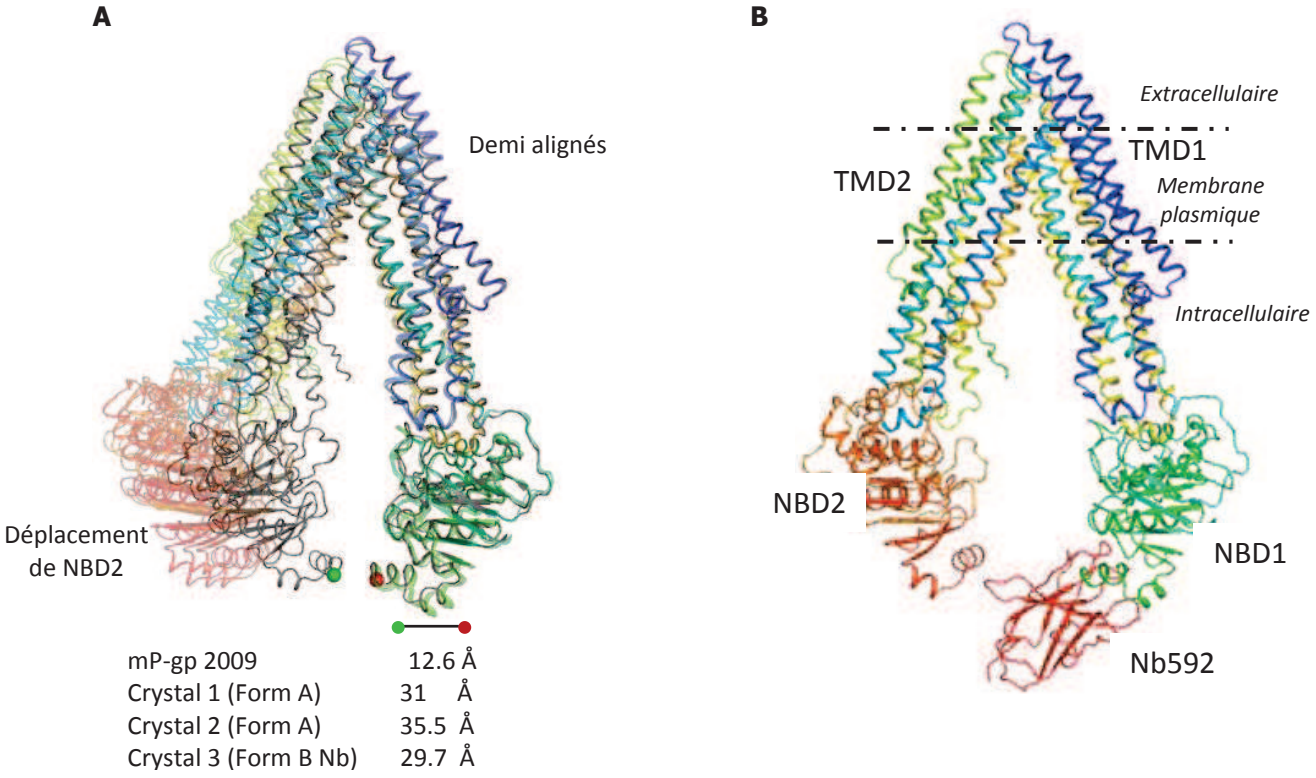


Figure 11 : Structures de la P-gp de souris.

A) Superpositions des structures de la P-gp de souris de 2013 et Apo de 2009 (code PDB: 3G5U).

B) Structure aux rayons-X de la P-gp complexée avec le Nb592 (en rouge).

3G5U est présentée en gris. Les nouvelles structures de la P-gp sont colorées en arc-en-ciel (N-terminal en bleu, C-terminal en rouge).

"Demi alignés": structures X-ray alignés en utilisant les résidus dans TMD1 et NBD1 (résidus 33-209, 852-961, 320-626 et); La distance entre C-term de NBD1 (résidu 626 en rouge) à C-terme de NBD2 (résidu 1271 en vert) sont indiqués.

Les barres horizontales représentent le positionnement approximatif de la bicouche lipidique. «transmembrane domain» (TMD), «Nucleotide Binding Domain» (NBD).

Chapitre 3. Localisation moléculaire de deux sites de transport dans la P-gp

La P-glycoprotéine est une protéine membranaire capable d'interagir avec de nombreuses molécules de structures chimiques variées, dont le seul point commun est de pouvoir diffuser passivement à travers les membranes biologiques. Cette spécificité multiple ou polyspécificité la distingue des autres transporteurs et récepteurs membranaires qui reconnaissent avec une grande affinité un ligand spécifique. Un des défis actuels est de pouvoir déterminer les caractéristiques de la spécificité de reconnaissance des substrats de la P-glycoprotéine. Une partie de cette polyspécificité s'explique par l'existence d'au moins trois sites de fixation pour les différentes molécules [227, 228]. Le site R lie préférentiellement la rhodamine 123 et les anthracyclines tels que la daunorubicine et la doxorubicine. Le site H lie le Hoechst 33342 et la colchicine. Le site P fixe la prazosine et la progestérone. L'existence d'un quatrième site liant seulement des modulateurs non-transportés (GF120918 ou nicardipine) a également été proposée [226]. La zone de fixation des différents molécules forme une grande région flexible composée de plusieurs sous-sites qui lient sélectivement les substrats à travers des interactions hydrophobes, liaisons hydrogènes, ou des interactions électrostatiques [189].

En 2009, la P-gp de souris a été cristallisée seule ou complexée à deux inhibiteurs hexapeptidiques cycliques (QZ59-RRR et QZ59-SSS). La détermination de cette structure tridimensionnelle a constitué une avancée sans précédent dans le domaine [166]. Toutefois, aucune indication n'a été apportée sur leur mode d'inhibition ni sur la localisation distincte ou commune de leurs sites de fixation avec ceux des médicaments. Cette information est pourtant essentielle à la conception d'une nouvelle génération d'inhibiteurs plus efficaces, qui cibleraient les sites de transport. Afin de répondre à cette question, le mécanisme d'inhibition de ces QZ59 par rapport au transport du Hoechst 33342 et de la daunorubicine a été caractérisé.

Le mode d'inhibition des énantiomères du QZ59 sur le transport de substrats se fixant aux sites P, R ou H a été caractérisé *in cellulo*. L'activité d'efflux de la P-gp pour la bodipy-prazosine (fixant le site P), la daunorubicine (fixant le site R) et le Hoeschst 33342 (fixant le site H), est comparée en absence ou en présence d'inhibiteur. Pour cela, des cellules sont mises en présence de concentrations variées en substrat fluorescent (0 à 6 µM) en absence ou en présence de QZ59 (0 à 4 µM). Après un temps d'incubation permettant l'accumulation des substrats, les cellules sont analysées par cytométrie en flux pour connaître la quantité de substrat accumulée. Deux types de cellules ont été comparés : la lignée cellulaire NIH3T3 (contrôle) ou NIH3T3 P-gp (transfectée avec le gène de la P-gp humaine).

L'efflux due à l'activité de la P-gp est estimé en soustrayant la fluorescence des cellules contrôles NIH3T3 (qui accumulent au maximum le substrat fluorescent) à celles des cellules NIH3T3 exprimant la P-gp (qui efflue en partie le substrat). La vitesse de transport a été rapportée en fonction de la concentration en substrat et en fonction de la concentration en QZ59. Pour chaque condition expérimentale, le modèle cinétique correspondant (selon l'équation de Michaelis & Menten, compétitif, non-compétitif, mixte…) a été recherché. En collaboration avec Dr. Emilie Henin et Pr. M. Tod (Faculté de Médecine Lyon-sud Charles Mérieux, Lyon), chaque modèle a été évalué

selon le critère d'information d'Akaike (AIC). La valeur la plus basse d'AIC indique le modèle le plus approprié.

Ces conditions conduisent à des paramètres de valeur apparente, ainsi toutes les constantes cinétiques ont le terme "app". Ceci est classiquement fait quand on ne peut pas établir de véritables valeurs des paramètres cinétiques (K_m, K_i, ...). Déterminer les vrais paramètres cinétiques, et en particulier le taux initial de drogue transloqué par la P-gp, est plus compliqué et non nécessaire pour établir le mécanisme d'inhibition des inhibiteurs de QZ59s.

Le cas de bodipy-prazosine

La bodipy-prazosine (BP) a été utilisée comme un analogue de la prazosine [223], pour explorer les effets de QZ59-RRR ou QZ59-SSS sur le site P. Le transport de BP médiée par la P-gp est déduite en soustrayant l'accumulation de BP dans les cellules contrôles NIH3T3 (triangles de la Figure 12A) des cellules NIH3T3 exprimant la P-gp (dans Figure 12A). Ce transport indique un modèle complexe, étant maximal à 1 µM et puis décroissant (carrés de la Figure 12A), lequel est caractéristique d'une inhibition causée par un excès de substrat. Régression de données avec l'équation correspondante (eq 2.6 dans le Tableau 1 de la publication III: $V_m S / (K_m + S (1 + S/ K_{SI}))$ a donné un $V_{m,app}$ de 60 ± 9 pmol mg^{-1} s^{-1}, un $K_{m,app}$ de 1,1 ± 0,2 µM et une constante d'inhibition de substrat $K_{si,app}$ de 0,6 ± 0,1 µM.

Pour vérifier que l'effet de BP était une inhibition par le substrat, l'inhibition de la prazosine vers son analogue bodipy a été vérifiée, supposée de type compétitif. Étonnamment, cet effet n'a pas été observé, l'efflux de bodipy-prazosine à 1 µM restant identique jusqu'à 30 µM de prazosine (Figure 12B). Ceci suggère que la bodipy-prazosine ne se comporte pas comme la prazosine. De ce fait, l'analogue fluorescent n'a plus été utilisé dans cette étude.

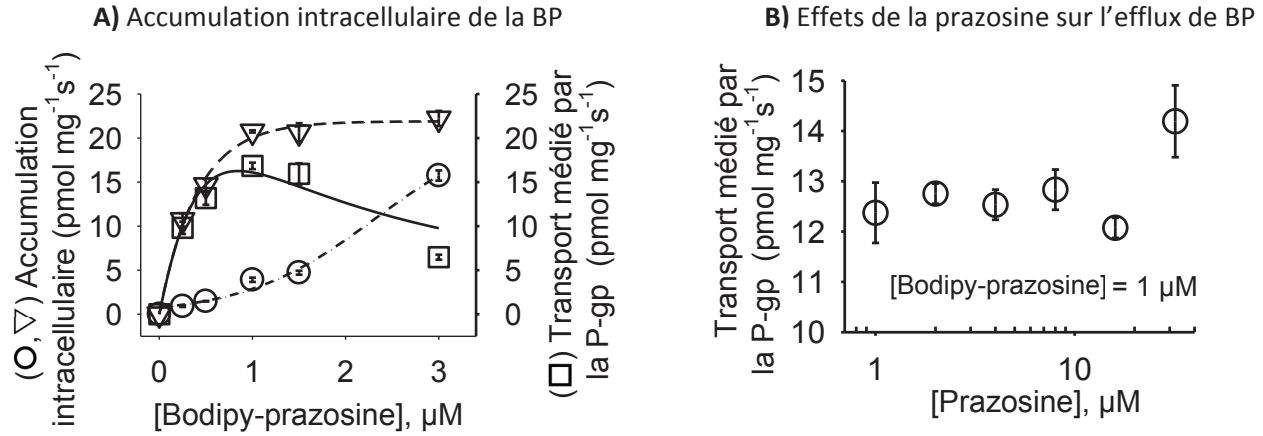


Figure12 : Accumulation intracellulaire de la bodipy-prazosine et transport médié par la P-gp

A) Accumulation intracellulaire de la bodipy-prazosine (BP) dans les cellules contrôles NIH3T3 (triangles) et dans les cellules NIH3T3 exprimant P-gp (cercles). L'efflux de BP par P-gp est la différence entre les triangles et cercles (carrés).

B) Vitesse de transport de la BP en fonction de la prazosine (cercles).

Le cas du Hoechst 33342 et de la daunorubicine

Le choix du modèle d'inhibition du QZ59-RRR vis-à-vis du Hoeschst 33342 est détaillé dans la Figure 13. Les vitesses d'efflux du Hoechst 33342 avec 0 µM (losanges), 1 µM (triangles vers le bas en bleu), 2 µM (triangles vers le haut en vert) et 4 µM (cercles rouges) de QZ59-RRR sont indiquées. La première colonne affiche le modèle testé ainsi que le numéro de l'équation utilisée pour ajuster les données (Tableaux 1 et S1 de la publication III). La deuxième colonne affiche le « Goodness-Of-Fit » (GOF), qui représente la qualité de l'ajustement et qui indique si le modèle ajuste bien les données en les distribuant de manière symétrique autour de la ligne théorique magenta. La troisième, quatrième et cinquième colonne affichent le nombre de points expérimentaux (N), le nombre de paramètres utilisés dans chaque modèle (*npar*) et le critère d'information d'Akaike corrigé du modèle (AIC_c). La sixième colonne montre la différence d'AIC (ΔAIC) entre le modèle de référence, Michaelis & Menten, et le modèle testé. La dernier colonne classe les différents modèles, du plus approprié en rouge jusqu'au moins approprié, en utilisant les ΔAIC.

Comme illustré, chaque modèle d'inhibition donne une AICc allant de 423,9 à 368,6, plus faibles que le modèle de référence qui ne comprend pas d'inhibition (AICc = 543,4). Parmi ces modèles, le modèle d'inhibition mixte donne le plus bas AICc, 368,6 (la valeur la plus basse d'AIC indique le modèle le plus approprié) et comparé au modèle de référence, donne la plus grande différence, ΔAICc de 174,8. Ces scores sont significativement meilleurs que ceux calculés pour le modèle d'inhibition non compétitive (AICc de 371,7 et Δ AICc de 171,7). La différence entre le modèle d'inhibition non compétitive et mixte est de 371,7 à 368,6 = 3,1. Cet écart, même faible reste significatif, une différence de deux étant statistiquement significative pour choisir un modèle. Ainsi le QZ59-RRR est un inhibiteur mixte non-compétitif vis-à-vis du Hoechst 33342.

Ce type d'analyse a été utilisé pour choisir le mode d'inhibition de QZ59-RRR vis-à-vis la daunorubicine (S1B de la publication III). Ainsi que le QZ59-SSS vis-à-vis du Hoechst 33342 (S2A de la publication III) et de la daunorubicine (S2B de la publication III).

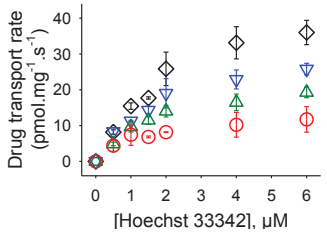
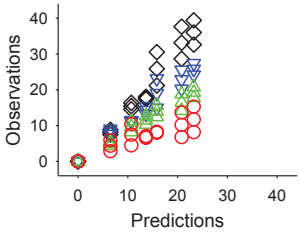
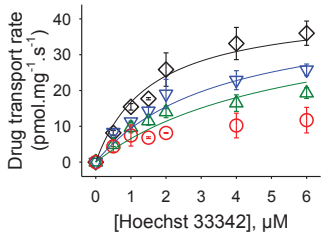
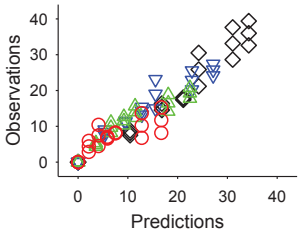
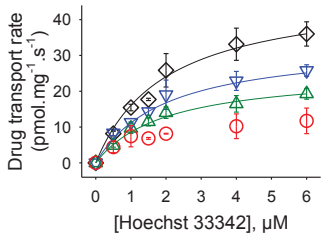
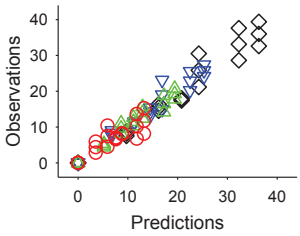
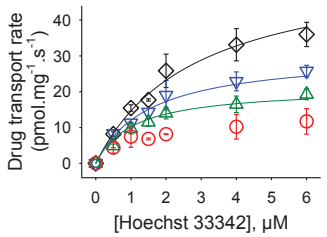
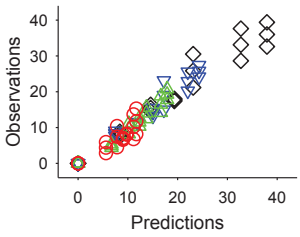
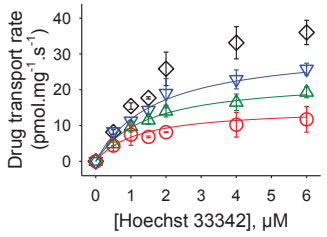
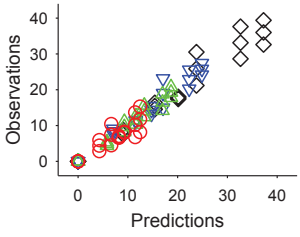
Modèle	GOF	N	*npar*	AICc	ΔAICc (ref-mod)	rank
Pas d'inhibition (eq. 2.1, reference)	 	84	2 K_m V_m	543,4	0	-
Inhibition compétitive (eq. 3.1)	 	84	3 K_m V_m K_I	423,9	119,5	4
Inhibition non-compétitive (eq. 4.1)	 	84	3 K_m V_m K_I	371,7	171,7	2
Inhibition incompétitive (eq. 5.1)	 	84	3 K_m V_m K_I	374,2	169,2	3
Inhibition mixte (eq. 6.1)	 	84	4 K_m V_m K_{i1} K_{i2}	368,6	174,8	1

Figure13 : Modèles d'inhibition du QZ59-RRR vis-à-vis du Hoechst 33342.

Mes résultats montrent que le QZ59-RRR agit comme un inhibiteur mixte avec une tendance non-compétitive vis-à-vis du Hoechst 33342 et comme un inhibiteur non-compétitif vis-à-vis de la daunorubicine. Ceci indique que sa localisation dans la protéine ne correspond pas avec ces sites de transport. Contrairement au QZ59-RRR, le QZ59-SSS rivalise très efficacement avec ces composés,

avec des constantes d'inhibition de 0,15 µM pour le Hoeschst 33342 et 0,3 µM pour la daunorubicine. Ceci suggère que les deux sites de fixation de cet inhibiteur dans la P-gp correspondent aux sites de transport. Une étude de liaison *in silico* nous a permis de distinguer le site du Hoechst 33342 et de la daunorubicine (Figure 14).

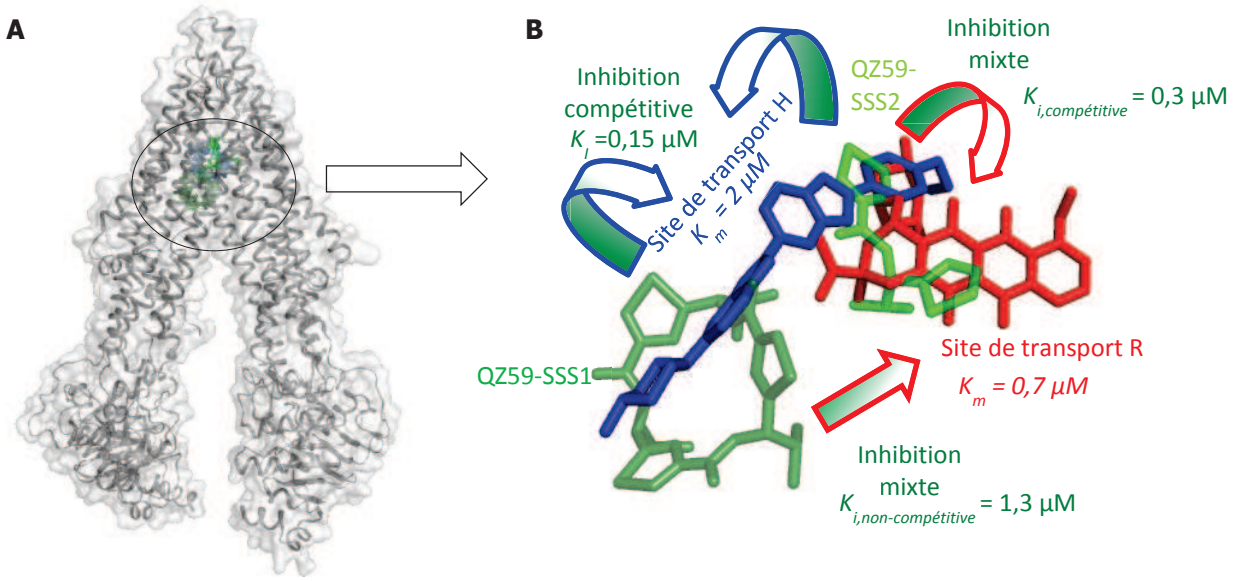


Figure 14 : Localisation moléculaire dans la P-gp du Hoechst 33342 et de la daunorubicine.
A) structure 3D de la P-gp avec les 2 énantiomères de l'inhibiteur hexapeptidique cyclique QZ59. Le QZ59-SSS est colorié en vert (plus haut) et vert obscure (plus bas), le QZ59-RRR est colorié en bleu claire.
B) Localisation des sites. Les couleurs pour les inhibiteurs hexapeptidiques cycliques QZ59 est le même que dans A. Le Hoechst 33342 est colorié en bleu et la daunorubicine en rouge. Les flèches montrent l'effet d'inhibition (avec leurs constantes cinétiques) de chaque QZ59 sur le site de transport du Hoechst 33342 et de la daunorubicine.

Ces résultats font l'objet d'une publication acceptée dans le Journal *FEBS* (Publication III).

Pour conclure

La recherche de la résistance aux médicaments liées à l'activité des pompes à efflux, dans les prochaines années, vont atteindre un tel stade que le mécanisme par lequel ces protéines transportent une grande variété de molécules hors de la cellule seront bien connu. Pour y parvenir, de nouvelles données biochimiques et biophysiques, ainsi que des structures de haute résolution des exportateurs ABC dans des conformations différentes (avec ou sans substrat), sont nécessaires. A ce stade, il sera intéressant de voir comment cette richesse de connaissances sera traduite pour contourner la multi-résistance en clinique.

REFERENCES

1. Simon SM, Schindler M. Cell biological mechanisms of multidrug resistance in tumors. *Proc Natl Acad Sci U S A* 1994,**91**:3497-3504.
2. Latosińska JN, Latosińska M. Anticancer Drug Discovery - From Serendipity to Rational Design. *Drug Discovery, Prof. Hany El-Shemy (Ed.),* 2013,**ISBN: 978-953-51-0906-8, InTech, DOI: 10.5772/52507**
3. Loddo M, Kingsbury SR, Rashid M, Proctor I, Holt C, Young J*, et al.* Cell-cycle-phase progression analysis identifies unique phenotypes of major prognostic and predictive significance in breast cancer. *Br J Cancer* 2009,**100**:959-970.
4. Ullah MF. Cancer multidrug resistance (MDR): a major impediment to effective chemotherapy. *Asian Pac J Cancer Prev* 2008,**9**:1-6.
5. Hanane Akhdar, Claire Legendre, Caroline Aninat, More F. Anticancer Drug Metabolism: Chemotherapy Resistance and New Therapeutic Approaches. *Topics on Drug Metabolism, Dr. James Paxton (Ed.)* 2012,**ISBN: 978-953-51-0099-7, InTech, DOI: 10.5772/30015.**
6. Gottesman MM. Mechanisms of cancer drug resistance. *Annu Rev Med* 2002,**53**:615-627.
7. Gillet JP, Gottesman MM. Overcoming multidrug resistance in cancer: 35 years after the discovery of ABCB1. *Drug Resist Updat* 2012,**15**:2-4.
8. Szakacs G, Paterson JK, Ludwig JA, Booth-Genthe C, Gottesman MM. Targeting multidrug resistance in cancer. *Nat Rev Drug Discov* 2006,**5**:219-234.
9. Barre-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J*, et al.* Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 1983,**220**:868-871.
10. Gallo RC, Sarin PS, Gelmann EP, Robert-Guroff M, Richardson E, Kalyanaraman VS*, et al.* Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). *Science* 1983,**220**:865-867.
11. Klimas N, Koneru AO, Fletcher MA. Overview of HIV. *Psychosom Med* 2008,**70**:523-530.
12. Clavel F, Guetard D, Brun-Vezinet F, Chamaret S, Rey MA, Santos-Ferreira MO*, et al.* Isolation of a new human retrovirus from West African patients with AIDS. *Science* 1986,**233**:343-346.
13. McCutchan FE. Global epidemiology of HIV. *J Med Virol* 2006,**78 Suppl 1**:S7-S12.
14. Teixeira C, Gomes JR, Gomes P, Maurel F, Barbault F. Viral surface glycoproteins, gp120 and gp41, as potential drug targets against HIV-1: brief overview one quarter of a century past the approval of zidovudine, the first anti-retroviral drug. *Eur J Med Chem* 2011,**46**:979-992.
15. Ghosh RK, Ghosh SM, Chawla S. Recent advances in antiretroviral drugs. *Expert Opin Pharmacother* 2011,**12**:31-46.
16. Weiss J, Haefeli WE. Impact of ATP-binding cassette transporters on human immunodeficiency virus therapy. *Int Rev Cell Mol Biol* 2010,**280**:219-279.
17. Kis O, Robillard K, Chan GN, Bendayan R. The complexities of antiretroviral drug-drug interactions: role of ABC and SLC transporters. *Trends Pharmacol Sci* 2010,**31**:22-35.
18. Desai M, Iyer G, Dikshit RK. Antiretroviral drugs: critical issues and recent advances. *Indian J Pharmacol* 2012,**44**:288-298.
19. Johnson VA, Calvez V, Gunthard HF, Paredes R, Pillay D, Shafer RW*, et al.* Update of the drug resistance mutations in HIV-1: March 2013. *Top Antivir Med* 2013,**21**:6-14.
20. Martin LR, Williams SL, Haskard KB, Dimatteo MR. The challenge of patient adherence. *Ther Clin Risk Manag* 2005,**1**:189-199.

21. Learned J. Why HIV drug resistance matters: an overview. *Posit Aware* 2005,**16**:26-30.

22. Amaral L, Engi H, Viveiros M, Molnar J. Review. Comparison of multidrug resistant efflux pumps of cancer and bacterial cells with respect to the same inhibitory agents. *In Vivo* 2007,**21**:237-244.

23. Rees DC, Johnson E, Lewinson O. ABC transporters: the power to change. *Nat Rev Mol Cell Biol* 2009,**10**:218-227.

24. Piddock LJ. Multidrug-resistance efflux pumps - not just for resistance. *Nat Rev Microbiol* 2006,**4**:629-636.

25. Nikaido H. Multidrug resistance in bacteria. *Annu Rev Biochem* 2009,**78**:119-146.

26. van Veen HW, Venema K, Bolhuis H, Oussenko I, Kok J, Poolman B*, et al.* Multidrug resistance mediated by a bacterial homolog of the human multidrug transporter MDR1. *Proc Natl Acad Sci U S A* 1996,**93**:10668-10672.

27. Van Bambeke F, Pages JM, Lee VJ. Inhibitors of bacterial efflux pumps as adjuvants in antibiotic treatments and diagnostic tools for detection of resistance by efflux. *Recent Pat Antiinfect Drug Discov* 2006,**1**:157-175.

28. International Transporter C, Giacomini KM, Huang SM, Tweedie DJ, Benet LZ, Brouwer KL*, et al.* Membrane transporters in drug development. *Nat Rev Drug Discov* 2010,**9**:215-236.

29. Keogh JP. Membrane transporters in drug development. *Adv Pharmacol* 2012,**63**:1-42.

30. Giacomini KM, Huang SM. Transporters in drug development and clinical pharmacology. *Clin Pharmacol Ther* 2013,**94**:3-9.

31. AC. R. Efflux and uptake transporters as determinants of statin response. *Expert Opinion on Drug Metabolism & Toxicology* 2010,**6**:621-632.

32. Tamaki A, Ierano C, Szakacs G, Robey RW, Bates SE. The controversial role of ABC transporters in clinical oncology. *Essays Biochem* 2011,**50**:209-232.

33. Benderra Z, Faussat AM, Sayada L, Perrot JY, Chaoui D, Marie JP*, et al.* Breast cancer resistance protein and P-glycoprotein in 149 adult acute myeloid leukemias. *Clin Cancer Res* 2004,**10**:7896-7902.

34. Clarke R, Leonessa F, Trock B. Multidrug resistance/P-glycoprotein and breast cancer: review and meta-analysis. *Semin Oncol* 2005,**32**:S9-15.

35. Zochbauer-Muller S, Filipits M, Rudas M, Brunner R, Krajnik G, Suchomel R*, et al.* P-glycoprotein and MRP1 expression in axillary lymph node metastases of breast cancer patients. *Anticancer Res* 2001,**21**:119-124.

36. Pavek P, Fendrich Z, Staud F. [Physiologic function of P-glycoprotein]. *Cesk Fysiol* 2002,**51**:99-107.

37. Fletcher JI, Haber M, Henderson MJ, Norris MD. ABC transporters in cancer: more than just drug efflux pumps. *Nat Rev Cancer* 2010,**10**:147-156.

38. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000,**100**:57-70.

39. Amiri-Kordestani L, Basseville A, Kurdziel K, Fojo AT, Bates SE. Targeting MDR in breast and lung cancer: discriminating its potential importance from the failure of drug resistance reversal studies. *Drug Resist Updat* 2012,**15**:50-61.

40. Burse A, Weingart H, Ullrich MS. NorM, an Erwinia amylovora multidrug efflux pump involved in in vitro competition with other epiphytic bacteria. *Appl Environ Microbiol* 2004,**70**:693-703.

41. Jerse AE, Sharma ND, Simms AN, Crow ET, Snyder LA, Shafer WM. A gonococcal efflux pump system enhances bacterial survival in a female mouse model of genital tract infection. *Infect Immun* 2003,**71**:5576-5582.

42. Jaciuk M, Nowak E, Skowronek K, Tanska A, Nowotny M. Structure of UvrA nucleotide excision repair protein in complex with modified DNA. *Nat Struct Mol Biol* 2011,**18**:191-197.

43. Giraud C, Manceau S, Treluyer JM. ABC transporters in human lymphocytes: expression, activity and role, modulating factors and consequences for antiretroviral therapies. *Expert Opin Drug Metab Toxicol* 2010,**6**:571-589.

44. Lucia MB, Cauda R, Landay AL, Malorni W, Donelli G, Ortona L. Transmembrane P-glycoprotein (P-gp/P-170) in HIV infection: analysis of lymphocyte surface expression and drug-unrelated function. *AIDS Res Hum Retroviruses* 1995,**11**:893-901.

45. Meaden ER, Hoggard PG, Maher B, Khoo SH, Back DJ. Expression of P-glycoprotein and multidrug resistance-associated protein in healthy volunteers and HIV-infected patients. *AIDS Res Hum Retroviruses* 2001,**17**:1329-1332.

46. Lucia MB, Rutella S, Leone G, Larocca LM, Vella S, Cauda R. In vitro and in vivo modulation of MDR1/P-glycoprotein in HIV-infected patients administered highly active antiretroviral therapy and liposomal doxorubicin. *J Acquir Immune Defic Syndr* 2002,**30**:369-378.

47. Jorajuria S, Clayette P, Dereuddre-Bosquet N, Benlhassan-Chahour K, Thiebot H, Vaslin B*, et al.* The expression of P-glycoprotein and cellular kinases is modulated at the transcriptional level by infection and highly active antiretroviral therapy in a primate model of AIDS. *AIDS Res Hum Retroviruses* 2003,**19**:307-311.

48. Gollapudi S, Gupta S. Human immunodeficiency virus I-induced expression of P-glycoprotein. *Biochem Biophys Res Commun* 1990,**171**:1002-1007.

49. Andreana A, Aggarwal S, Gollapudi S, Wien D, Tsuruo T, Gupta S. Abnormal expression of a 170-kilodalton P-glycoprotein encoded by MDR1 gene, a metabolically active efflux pump, in CD4+ and CD8+ T cells from patients with human immunodeficiency virus type 1 infection. *AIDS Res Hum Retroviruses* 1996,**12**:1457-1462.

50. Gupta S, Gollapudi S. P-glycoprotein (MDR 1 gene product) in cells of the immune system: its possible physiologic role and alteration in aging and human immunodeficiency virus-1 (HIV-1) infection. *J Clin Immunol* 1993,**13**:289-301.

51. Lucia MB, Cauda R, Malorni W, Rainaldi G, Tumbarello M, Tacconelli E*, et al.* P-170 glycoprotein (P-170) is involved in the impairment of natural killer cell-mediated cytotoxicity in HIV+ patients. *Immunol Lett* 1995,**47**:223-226.

52. Lee CG, Ramachandra M, Jeang KT, Martin MA, Pastan I, Gottesman MM. Effect of ABC transporters on HIV-1 infection: inhibition of virus production by the MDR1 transporter. *FASEB J* 2000,**14**:516-522.

53. Speck RR, Yu XF, Hildreth J, Flexner C. Differential effects of p-glycoprotein and multidrug resistance protein-1 on productive human immunodeficiency virus infection. *J Infect Dis* 2002,**186**:332-340.

54. George AM, Jones PM. Perspectives on the structure-function of ABC transporters: the Switch and Constant Contact models. *Prog Biophys Mol Biol* 2012,**109**:95-107.

55. Juliano RL, Ling V. A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants. *Biochim Biophys Acta* 1976,**455**:152-162.

56. Hyde SC, Emsley P, Hartshorn MJ, Mimmack MM, Gileadi U, Pearce SR*, et al.* Structural model of ATP-binding proteins associated with cystic fibrosis, multidrug resistance and bacterial transport. *Nature* 1990,**346**:362-365.

57. Higgins CF. ABC transporters: physiology, structure and mechanism--an overview. *Res Microbiol* 2001,**152**:205-210.

58. Higgins CF. ABC transporters: from microorganisms to man. *Annu Rev Cell Biol* 1992,**8**:67-113.

59. Chakraburtty K. Translational regulation by ABC systems. *Res Microbiol* 2001,**152**:391-399.

60. Belfield GP, Ross-Smith NJ, Tuite MF. Translation elongation factor-3 (EF-3): an evolving eukaryotic ribosomal protein? *J Mol Evol* 1995,**41**:376-387.

61. Goosen N, Moolenaar GF. Role of ATP hydrolysis by UvrA and UvrB during nucleotide excision repair. *Res Microbiol* 2001,**152**:401-409.

62. Geourjon C, Orelle C, Steinfels E, Blanchet C, Deleage G, Di Pietro A*, et al.* A common mechanism for ATP hydrolysis in ABC transporter and helicase superfamilies. *Trends Biochem Sci* 2001,**26**:539-544.

63. Hopfner KP, Tainer JA. Rad50/SMC proteins and ABC transporters: unifying concepts from high-resolution structures. *Curr Opin Struct Biol* 2003,**13**:249-255.

64. Saurin W, Hofnung M, Dassa E. Getting in or out: early segregation between importers and exporters in the evolution of ATP-binding cassette (ABC) transporters. *J Mol Evol* 1999,**48**:22-41.

65. Quazi F, Lenevich S, Molday RS. ABCA4 is an N-retinylidene-phosphatidylethanolamine and phosphatidylethanolamine importer. *Nat Commun* 2012,**3**:925.

66. Vasiliou V, Vasiliou K, Nebert DW. Human ATP-binding cassette (ABC) transporter family. *Hum Genomics* 2009,**3**:281-290.

67. Dean M, Hamon Y, Chimini G. The human ATP-binding cassette (ABC) transporter superfamily. *J Lipid Res* 2001,**42**:1007-1017.

68. Holland IB, Blight MA. ABC-ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans. *J Mol Biol* 1999,**293**:381-399.

69. Dassa E, Bouige P. The ABC of ABCS: a phylogenetic and functional classification of ABC systems in living organisms. *Res Microbiol* 2001,**152**:211-229.

70. Dean M, Rzhetsky A, Allikmets R. The human ATP-binding cassette (ABC) transporter superfamily. *Genome Res* 2001,**11**:1156-1166.

71. Tusnady GE, Sarkadi B, Simon I, Varadi A. Membrane topology of human ABC proteins. *FEBS Lett* 2006,**580**:1017-1022.

72. Cole SP, Bhardwaj G, Gerlach JH, Mackie JE, Grant CE, Almquist KC*, et al.* Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line. *Science* 1992,**258**:1650-1654.

73. Doyle LA, Yang W, Abruzzo LV, Krogmann T, Gao Y, Rishi AK*, et al.* A multidrug resistance transporter from human MCF-7 breast cancer cells. *Proc Natl Acad Sci U S A* 1998,**95**:15665-15670.

74. Allikmets R, Schriml LM, Hutchinson A, Romano-Spica V, Dean M. A human placenta-specific ATP-binding cassette gene (ABCP) on chromosome 4q22 that is involved in multidrug resistance. *Cancer Res* 1998,**58**:5337-5339.

75. Miyake K, Mickley L, Litman T, Zhan Z, Robey R, Cristensen B*, et al.* Molecular cloning of cDNAs which are highly overexpressed in mitoxantrone-resistant cells: demonstration of homology to ABC transport genes. *Cancer Res* 1999,**59**:8-13.

76. Siest G, Marteau JB, Maumus S, Berrahmoune H, Jeannesson E, Samara A*, et al.* Pharmacogenomics and cardiovascular drugs: need for integrated biological system with phenotypes and proteomic markers. *Eur J Pharmacol* 2005,**527**:1-22.

77. Gillet JP, Efferth T, Remacle J. Chemotherapy-induced resistance by ATP-binding cassette transporter genes. *Biochim Biophys Acta* 2007,**1775**:237-262.

78. Loo TW, Clarke DM. Mutational analysis of ABC proteins. *Arch Biochem Biophys* 2008,**476**:51-64.

79. Linton KJ. Structure and function of ABC transporters. *Physiology (Bethesda)* 2007,**22**:122-130.

80. Ozvegy C, Litman T, Szakacs G, Nagy Z, Bates S, Varadi A*, et al.* Functional characterization of the human multidrug transporter, ABCG2, expressed in insect cells. *Biochem Biophys Res Commun* 2001,**285**:111-117.

81. Graf GA, Yu L, Li WP, Gerard R, Tuma PL, Cohen JC*, et al.* ABCG5 and ABCG8 are obligate heterodimers for protein trafficking and biliary cholesterol excretion. *J Biol Chem* 2003,**278**:48275-48282.

82. Boscoboinik D, Debanne MT, Stafford AR, Jung CY, Gupta RS, Epand RM. Dimerization of the P-glycoprotein in membranes. *Biochim Biophys Acta* 1990,**1027**:225-228.

83. Jette L, Potier M, Beliveau R. P-glycoprotein is a dimer in the kidney and brain capillary membranes: effect of cyclosporin A and SDZ-PSC 833. *Biochemistry* 1997,**36**:13929-13937.

84. Poruchynsky MS, Ling V. Detection of oligomeric and monomeric forms of P-glycoprotein in multidrug resistant cells. *Biochemistry* 1994,**33**:4163-4174.

85. Soszynski M, Kaluzna A, Rychlik B, Sokal A, Bartosz G. Radiation inactivation suggests that human multidrug resistance-associated protein 1 occurs as a dimer in the human erythrocyte membrane. *Arch Biochem Biophys* 1998,**354**:311-316.

86. Dezi M, Fribourg PF, Di Cicco A, Arnaud O, Marco S, Falson P*, et al.* The multidrug resistance half-transporter ABCG2 is purified as a tetramer upon selective extraction from membranes. *Biochim Biophys Acta* 2010,**1798**:2094-2101.

87. Xu J, Liu Y, Yang Y, Bates S, Zhang JT. Characterization of oligomeric human half-ABC transporter ATP-binding cassette G2. *J Biol Chem* 2004,**279**:19781-19789.

88. Dawson RJ, Hollenstein K, Locher KP. Uptake or extrusion: crystal structures of full ABC transporters suggest a common mechanism. *Mol Microbiol* 2007,**65**:250-257.

89. Dawson RJ, Locher KP. Structure of a bacterial multidrug ABC transporter. *Nature* 2006,**443**:180-185.

90. Ward A, Reyes CL, Yu J, Roth CB, Chang G. Flexibility in the ABC transporter MsbA: Alternating access with a twist. *Proc Natl Acad Sci U S A* 2007,**104**:19005-19010.

91. Hollenstein K, Frei DC, Locher KP. Structure of an ABC transporter in complex with its binding protein. *Nature* 2007,**446**:213-216.

92. Oldham ML, Khare D, Quiocho FA, Davidson AL, Chen J. Crystal structure of a catalytic intermediate of the maltose transporter. *Nature* 2007,**450**:515-521.

93. Kadaba NS, Kaiser JT, Johnson E, Lee A, Rees DC. The high-affinity E. coli methionine ABC transporter: structure and allosteric regulation. *Science* 2008,**321**:250-253.

94. Locher KP, Lee AT, Rees DC. The E. coli BtuCD structure: a framework for ABC transporter architecture and mechanism. *Science* 2002,**296**:1091-1098.

95. Pinkett HW, Lee AT, Lum P, Locher KP, Rees DC. An inward-facing conformation of a putative metal-chelate-type ABC transporter. *Science* 2007,**315**:373-377.

96. Hollenstein K, Dawson RJ, Locher KP. Structure and mechanism of ABC transporter proteins. *Curr Opin Struct Biol* 2007,**17**:412-418.

97. Locher KP. Review. Structure and mechanism of ATP-binding cassette transporters. *Philos Trans R Soc Lond B Biol Sci* 2009,**364**:239-245.

98. Borbat PP, Surendhran K, Bortolus M, Zou P, Freed JH, McHaourab HS. Conformational motion of the ABC transporter MsbA induced by ATP hydrolysis. *PLoS Biol* 2007,**5**:e271.

99. Raviv Y, Pollard HB, Bruggemann EP, Pastan I, Gottesman MM. Photosensitized labeling of a functional multidrug transporter in living drug-resistant tumor cells. *J Biol Chem* 1990,**265**:3975-3980.

100. Higgins CF, Gottesman MM. Is the multidrug transporter a flippase? *Trends Biochem Sci* 1992,**17**:18-21.

101. Seeger MA, van Veen HW. Molecular basis of multidrug transport by ABC transporters. *Biochim Biophys Acta* 2009,**1794**:725-737.

102. Gutmann DA, Ward A, Urbatsch IL, Chang G, van Veen HW. Understanding polyspecificity of multidrug ABC transporters: closing in on the gaps in ABCB1. *Trends Biochem Sci* 2010,**35**:36-42.

103. Procko E, O'Mara ML, Bennett WF, Tieleman DP, Gaudet R. The mechanism of ABC transporters: general lessons from structural and functional studies of an antigenic peptide transporter. *FASEB J* 2009,**23**:1287-1302.

104. Sharom FJ. ABC multidrug transporters: structure, function and role in chemoresistance. *Pharmacogenomics* 2008,**9**:105-127.

105. Vetter IR, Wittinghofer A. Nucleoside triphosphate-binding proteins: different scaffolds to achieve phosphoryl transfer. *Q Rev Biophys* 1999,**32**:1-56.

106. Hung LW, Wang IX, Nikaido K, Liu PQ, Ames GF, Kim SH. Crystal structure of the ATP-binding subunit of an ABC transporter. *Nature* 1998,**396**:703-707.

107. Jones PM, O'Mara ML, George AM. ABC transporters: a riddle wrapped in a mystery inside an enigma. *Trends Biochem Sci* 2009,**34**:520-531.

108. Smith PC, Karpowich N, Millen L, Moody JE, Rosen J, Thomas PJ*, et al.* ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimer. *Mol Cell* 2002,**10**:139-149.

109. Moody JE, Millen L, Binns D, Hunt JF, Thomas PJ. Cooperative, ATP-dependent association of the nucleotide binding cassettes during the catalytic cycle of ATP-binding cassette transporters. *J Biol Chem* 2002,**277**:21111-21114.

110. Chen J, Lu G, Lin J, Davidson AL, Quiocho FA. A tweezers-like motion of the ATP-binding cassette dimer in an ABC transport cycle. *Mol Cell* 2003,**12**:651-661.

111. Diederichs K, Diez J, Greller G, Muller C, Breed J, Schnell C*, et al.* Crystal structure of MalK, the ATPase subunit of the trehalose/maltose ABC transporter of the archaeon Thermococcus litoralis. *EMBO J* 2000,**19**:5951-5961.

112. Hopfner KP, Karcher A, Shin DS, Craig L, Arthur LM, Carney JP*, et al.* Structural biology of Rad50 ATPase: ATP-driven conformational control in DNA double-strand break repair and the ABC-ATPase superfamily. *Cell* 2000,**101**:789-800.

113. Jones PM, George AM. The ABC transporter structure and mechanism: perspectives on recent research. *Cell Mol Life Sci* 2004,**61**:682-699.

114. Kerr ID. Structure and association of ATP-binding cassette transporter nucleotide-binding domains. *Biochim Biophys Acta* 2002,**1561**:47-64.

115. Oldham ML, Chen J. Snapshots of the maltose transporter during ATP hydrolysis. *Proc Natl Acad Sci U S A* 2011,**108**:15152-15156.

116. Orelle C, Dalmas O, Gros P, Di Pietro A, Jault JM. The conserved glutamate residue adjacent to the Walker-B motif is the catalytic base for ATP hydrolysis in the ATP-binding cassette transporter BmrA. *J Biol Chem* 2003,**278**:47002-47008.

117. Hanekop N, Zaitseva J, Jenewein S, Holland IB, Schmitt L. Molecular insights into the mechanism of ATP-hydrolysis by the NBD of the ABC-transporter HlyB. *FEBS Lett* 2006,**580**:1036-1041.

118. Zaitseva J, Jenewein S, Jumpertz T, Holland IB, Schmitt L. H662 is the linchpin of ATP hydrolysis in the nucleotide-binding domain of the ABC transporter HlyB. *EMBO J* 2005,**24**:1901-1910.

119. Zhou Y, Ojeda-May P, Pu J. H-loop Histidine Catalyzes ATP Hydrolysis in the E. coli ABC-Transporter HlyB. *Physical Chemistry Chemical Physics* 2013.

120. Jones PM, George AM. Role of the D-loops in allosteric control of ATP hydrolysis in an ABC transporter. *J Phys Chem A* 2012,**116**:3004-3013.

121. Janas E, Hofacker M, Chen M, Gompf S, van der Does C, Tampe R. The ATP hydrolysis cycle of the nucleotide-binding domain of the mitochondrial ATP-binding cassette transporter Mdl1p. *J Biol Chem* 2003,**278**:26862-26869.

122. Higgins CF, Linton KJ. The ATP switch model for ABC transporters. *Nat Struct Mol Biol* 2004,**11**:918-926.

123. Linton KJ, Higgins CF. Structure and function of ABC transporters: the ATP switch provides flexible control. *Pflugers Arch* 2007,**453**:555-567.

124. Senior AE, al-Shawi MK, Urbatsch IL. The catalytic cycle of P-glycoprotein. *FEBS Lett* 1995,**377**:285-289.

125. Sauna ZE, Kim IW, Nandigama K, Kopp S, Chiba P, Ambudkar SV. Catalytic cycle of ATP hydrolysis by P-glycoprotein: evidence for formation of the E.S reaction intermediate with ATP-gamma-S, a nonhydrolyzable analogue of ATP. *Biochemistry* 2007,**46**:13787-13799.

126. Jones PM, George AM. Nucleotide-dependent allostery within the ABC transporter ATP-binding cassette: a computational study of the MJ0796 dimer. *J Biol Chem* 2007,**282**:22793-22803.

127. Jones PM, George AM. 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* 2009,**75**:387-396.

128. Parcej D, Tampe R. ABC proteins in antigen translocation and viral inhibition. *Nat Chem Biol* 2010,**6**:572-580.

129. Jardetzky O. Simple allosteric model for membrane pumps. *Nature* 1966,**211**:969-970.

130. Zou P, Bortolus M, McHaourab HS. Conformational cycle of the ABC transporter MsbA in liposomes: detailed analysis using double electron-electron resonance spectroscopy. *J Mol Biol* 2009,**393**:586-597.

131. Mehmood S, Domene C, Forest E, Jault JM. Dynamics of a bacterial multidrug ABC transporter in the inward- and outward-facing conformations. *Proc Natl Acad Sci U S A* 2012,**109**:10832-10836.

132. Coleman JA, Quazi F, Molday RS. Mammalian P4-ATPases and ABC transporters and their role in phospholipid transport. *Biochim Biophys Acta* 2013,**1831**:555-574.

133. Oldham ML, Davidson AL, Chen J. Structural insights into ABC transporter mechanism. *Curr Opin Struct Biol* 2008,**18**:726-733.

134. Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H, Phillips DC. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* 1958,**181**:662-666.

135. White SH. The progress of membrane protein structure determination. *Protein Sci* 2004,**13**:1948-1949.

136. Ose T, Fujie T, Yao M, Watanabe N, Tanaka I. Crystal structure of the ATP-binding cassette of multisugar transporter from Pyrococcus horikoshii OT3. *Proteins* 2004,**57**:635-638.

137. Verdon G, Albers SV, van Oosterwijk N, Dijkstra BW, Driessen AJ, Thunnissen AM. Formation of the productive ATP-Mg2+-bound dimer of GlcV, an ABC-ATPase from Sulfolobus solfataricus. *J Mol Biol* 2003,**334**:255-267.

138. Verdon G, Albers SV, Dijkstra BW, Driessen AJ, Thunnissen AM. Crystal structures of the ATPase subunit of the glucose ABC transporter from Sulfolobus solfataricus: nucleotide-free and nucleotide-bound conformations. *J Mol Biol* 2003,**330**:343-358.

139. Scheffel F, Demmer U, Warkentin E, Hulsmann A, Schneider E, Ermler U. Structure of the ATPase subunit CysA of the putative sulfate ATP-binding cassette (ABC) transporter from Alicyclobacillus acidocaldarius. *FEBS Lett* 2005,**579**:2953-2958.

140. Gaudet R, Wiley DC. Structure of the ABC ATPase domain of human TAP1, the transporter associated with antigen processing. *EMBO J* 2001,**20**:4964-4972.

141. Schmitt L, Benabdelhak H, Blight MA, Holland IB, Stubbs MT. Crystal structure of the nucleotide-binding domain of the ABC-transporter haemolysin B: identification of a variable region within ABC helical domains. *J Mol Biol* 2003,**330**:333-342.

142. Zaitseva J, Jenewein S, Oswald C, Jumpertz T, Holland IB, Schmitt L. A molecular understanding of the catalytic cycle of the nucleotide-binding domain of the ABC transporter HlyB. *Biochem Soc Trans* 2005,**33**:990-995.

143. Ramaen O, Leulliot N, Sizun C, Ulryck N, Pamlard O, Lallemand JY*, et al.* Structure of the human multidrug resistance protein 1 nucleotide binding domain 1 bound to Mg2+/ATP reveals a non-productive catalytic site. *J Mol Biol* 2006,**359**:940-949.

144. Karpowich N, Martsinkevich O, Millen L, Yuan YR, Dai PL, MacVey K*, et al.* Crystal structures of the MJ1267 ATP binding cassette reveal an induced-fit effect at the ATPase active site of an ABC transporter. *Structure* 2001,**9**:571-586.

145. Yuan YR, Blecker S, Martsinkevich O, Millen L, Thomas PJ, Hunt JF. The crystal structure of the MJ0796 ATP-binding cassette. Implications for the structural consequences of ATP hydrolysis in the active site of an ABC transporter. *J Biol Chem* 2001,**276**:32313-32321.

146. Lewis HA, Buchanan SG, Burley SK, Conners K, Dickey M, Dorwart M*, et al.* Structure of nucleotide-binding domain 1 of the cystic fibrosis transmembrane conductance regulator. *EMBO J* 2004,**23**:282-293.

147. Lewis HA, Zhao X, Wang C, Sauder JM, Rooney I, Noland BW*, et al.* Impact of the deltaF508 mutation in first nucleotide-binding domain of human cystic fibrosis transmembrane conductance regulator on domain folding and structure. *J Biol Chem* 2005,**280**:1346-1353.

148. Lamers MH, Perrakis A, Enzlin JH, Winterwerp HH, de Wind N, Sixma TK. The crystal structure of DNA mismatch repair protein MutS binding to a G x T mismatch. *Nature* 2000,**407**:711-717.

149. Obmolova G, Ban C, Hsieh P, Yang W. Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature* 2000,**407**:703-710.

150. Oswald C, Holland IB, Schmitt L. The motor domains of ABC-transporters. What can structures tell us? *Naunyn Schmiedebergs Arch Pharmacol* 2006,**372**:385-399.

151. Rosenberg MF, Mao Q, Holzenburg A, Ford RC, Deeley RG, Cole SP. The structure of the multidrug resistance protein 1 (MRP1/ABCC1). crystallization and single-particle analysis. *J Biol Chem* 2001,**276**:16076-16082.

152. Rosenberg MF, Callaghan R, Ford RC, Higgins CF. Structure of the multidrug resistance P-glycoprotein to 2.5 nm resolution determined by electron microscopy and image analysis. *J Biol Chem* 1997,**272**:10685-10694.

153. Rosenberg MF, Kamis AB, Callaghan R, Higgins CF, Ford RC. Three-dimensional structures of the mammalian multidrug resistance P-glycoprotein demonstrate major conformational changes in the transmembrane domains upon nucleotide binding. *J Biol Chem* 2003,**278**:8294-8299.

154. Rosenberg MF, Callaghan R, Modok S, Higgins CF, Ford RC. Three-dimensional structure of P-glycoprotein: the transmembrane regions adopt an asymmetric configuration in the nucleotide-bound state. *J Biol Chem* 2005,**280**:2857-2862.

155. Chami M, Steinfels E, Orelle C, Jault JM, Di Pietro A, Rigaud JL*, et al.* Three-dimensional structure by cryo-electron microscopy of YvcC, an homodimeric ATP-binding cassette transporter from Bacillus subtilis. *J Mol Biol* 2002,**315**:1075-1085.

156. Ferreira-Pereira A, Marco S, Decottignies A, Nader J, Goffeau A, Rigaud JL. Three-dimensional reconstruction of the Saccharomyces cerevisiae multidrug resistance protein Pdr5p. *J Biol Chem* 2003,**278**:11995-11999.

157. Hvorup RN, Goetz BA, Niederer M, Hollenstein K, Perozo E, Locher KP. Asymmetry in the structure of the ABC transporter-binding protein complex BtuCD-BtuF. *Science* 2007,**317**:1387-1390.

158. Korkhov VM, Mireku SA, Hvorup RN, Locher KP. Asymmetric states of vitamin B(1)(2) transporter BtuCD are not discriminated by its cognate substrate binding protein BtuF. *FEBS Lett* 2012,**586**:972-976.

159. Korkhov VM, Mireku SA, Locher KP. Structure of AMP-PNP-bound vitamin B12 transporter BtuCD-F. *Nature* 2012,**490**:367-372.

160. Woo JS, Zeltina A, Goetz BA, Locher KP. X-ray structure of the Yersinia pestis heme transporter HmuUV. *Nat Struct Mol Biol* 2012,**19**:1310-1315.

161. Gerber S, Comellas-Bigler M, Goetz BA, Locher KP. Structural basis of trans-inhibition in a molybdate/tungstate ABC transporter. *Science* 2008,**321**:246-250.

162. Khare D, Oldham ML, Orelle C, Davidson AL, Chen J. Alternating access in maltose transporter mediated by rigid-body rotations. *Mol Cell* 2009,**33**:528-536.

163. Oldham ML, Chen J. Crystal structure of the maltose transporter in a pretranslocation intermediate state. *Science* 2011,**332**:1202-1205.

164. Johnson E, Nguyen PT, Yeates TO, Rees DC. Inward facing conformations of the MetNI methionine ABC transporter: Implications for the mechanism of transinhibition. *Protein Sci* 2012,**21**:84-96.

165. Dawson RJ, Locher KP. Structure of the multidrug ABC transporter Sav1866 from Staphylococcus aureus in complex with AMP-PNP. *FEBS Lett* 2007,**581**:935-938.

166. Aller SG, Yu J, Ward A, Weng Y, Chittaboina S, Zhuo R*, et al.* Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science* 2009,**323**:1718-1722.

167. Jin MS, Oldham ML, Zhang Q, Chen J. Crystal structure of the multidrug transporter P-glycoprotein from Caenorhabditis elegans. *Nature* 2012.

168. Ward AB, Szewczyk P, Grimard V, Lee C-W, Martinez L, Doshi R*, et al.* Structures of P-glycoprotein reveal its conformational flexibility and an epitope on the nucleotide-binding domain. *Proceedings of the National Academy of Sciences* 2013,**110**:13386-13391.

169. Li J, Jaimes KF, Aller SG. Refined structures of mouse P-Glycoprotein. *Protein Sci* 2014,**23**:34-46.

170. Shintre CA, Pike AC, Li Q, Kim JI, Barr AJ, Goubin S*, et al.* Structures of ABCB10, a human ATP-binding cassette transporter in apo- and nucleotide-bound states. *Proc Natl Acad Sci U S A* 2013,**110**:9710-9715.

171. Hohl M, Briand C, Grutter MG, Seeger MA. Crystal structure of a heterodimeric ABC transporter in its inward-facing conformation. *Nat Struct Mol Biol* 2012,**19**:395-402.

172. Lugo MR, Sharom FJ. Interaction of LDS-751 and rhodamine 123 with P-glycoprotein: evidence for simultaneous binding of both drugs. *Biochemistry* 2005,**44**:14020-14029.

173. Loo TW, Bartlett MC, Clarke DM. Simultaneous binding of two different drugs in the binding pocket of the human multidrug resistance P-glycoprotein. *J Biol Chem* 2003,**278**:39706-39710.

174. Gottesman MM, Ambudkar SV, Xia D. Structure of a multidrug transporter. *Nat Biotechnol* 2009,**27**:546-547.

175. Kerr ID, Jones PM, George AM. Multidrug efflux pumps: the structures of prokaryotic ATP-binding cassette transporter efflux pumps and implications for our understanding of eukaryotic P-glycoproteins and homologues. *FEBS J* 2010,**277**:550-563.

176. Linton KJ, Holland IB. *The ABC Transporters of Human Physiology and Disease: The Genetics and Biochemistry of ATP Binding Cassette Transporters*: World Scientific Publishing Company, Incorporated; 2011.

177. Wen PC, Verhalen B, Wilkens S, McHaourab HS, Tajkhorshid E. On the origin of large flexibility of P-glycoprotein in the inward-facing state. *J Biol Chem* 2013,**288**:19211-19220.

178. Borst P, Elferink RO. Mammalian ABC transporters in health and disease. *Annu Rev Biochem* 2002,**71**:537-592.

179. Gottesman MM, Hrycyna CA, Schoenlein PV, Germann UA, Pastan I. Genetic analysis of the multidrug transporter. *Annu Rev Genet* 1995,**29**:607-649.

180. Li Y, Yuan H, Yang K, Xu W, Tang W, Li X. The structure and functions of P-glycoprotein. *Curr Med Chem* 2010,**17**:786-800.

181. Liu R, Siemiarczuk A, Sharom FJ. Intrinsic fluorescence of the P-glycoprotein multidrug transporter: sensitivity of tryptophan residues to binding of drugs and nucleotides. *Biochemistry* 2000,**39**:14927-14938.

182. Liu R, Sharom FJ. Site-directed fluorescence labeling of P-glycoprotein on cysteine residues in the nucleotide binding domains. *Biochemistry* 1996,**35**:11865-11873.

183. Sharom FJ. The P-glycoprotein multidrug transporter. *Essays Biochem* 2011,**50**:161-178.

184. Ferte J. Analysis of the tangled relationships between P-glycoprotein-mediated multidrug resistance and the lipid phase of the cell membrane. *Eur J Biochem* 2000,**267**:277-294.

185. Desai PV, Raub TJ, Blanco MJ. How hydrogen bonds impact P-glycoprotein transport and permeability. *Bioorg Med Chem Lett* 2012,**22**:6540-6548.

186. Fu D, Arias IM. Intracellular trafficking of P-glycoprotein. *Int J Biochem Cell Biol* 2012,**44**:461-464.

187. Breier A, Gibalova L, Seres M, Barancik M, Sulova Z. New insight into p-glycoprotein as a drug target. *Anticancer Agents Med Chem* 2013,**13**:159-170.

188. Gottesman MM, Fojo T, Bates SE. Multidrug resistance in cancer: role of ATP-dependent transporters. *Nat Rev Cancer* 2002,**2**:48-58.

189. Eckford PD, Sharom FJ. ABC efflux pump-based resistance to chemotherapy drugs. *Chem Rev* 2009,**109**:2989-3011.

190. Boumendjel A, Boutonnat J, Robert J. *ABC Transporters and Multidrug Resistance*: Wiley; 2009.

191. Hipfner DR, Almquist KC, Leslie EM, Gerlach JH, Grant CE, Deeley RG*, et al.* Membrane topology of the multidrug resistance protein (MRP). A study of glycosylation-site mutants reveals an extracytosolic NH2 terminus. *J Biol Chem* 1997,**272**:23623-23630.

192. Borst P, Evers R, Kool M, Wijnholds J. A family of drug transporters: the multidrug resistance-associated proteins. *J Natl Cancer Inst* 2000,**92**:1295-1302.

193. Rothnie A, Conseil G, Lau AY, Deeley RG, Cole SP. Mechanistic differences between GSH transport by multidrug resistance protein 1 (MRP1/ABCC1) and GSH modulation of MRP1-mediated transport. *Mol Pharmacol* 2008,**74**:1630-1640.

194. Keppler D. Multidrug resistance proteins (MRPs, ABCCs): importance for pathophysiology and drug therapy. *Handb Exp Pharmacol* 2011:299-323.

195. Desuzinges-Mandon E, Arnaud O, Martinez L, Huche F, Di Pietro A, Falson P. ABCG2 transports and transfers heme to albumin through its large extracellular loop. *J Biol Chem* 2010,**285**:33123-33133.

196. Mohrmann K, van Eijndhoven MA, Schinkel AH, Schellens JH. Absence of N-linked glycosylation does not affect plasma membrane localization of breast cancer resistance protein (BCRP/ABCG2). *Cancer Chemother Pharmacol* 2005,**56**:344-350.

197. Henriksen U, Fog JU, Litman T, Gether U. Identification of intra- and intermolecular disulfide bridges in the multidrug resistance transporter ABCG2. *J Biol Chem* 2005,**280**:36926-36934.

198. Zhou S, Zong Y, Ney PA, Nair G, Stewart CF, Sorrentino BP. Increased expression of the Abcg2 transporter during erythroid maturation plays a role in decreasing cellular protoporphyrin IX levels. *Blood* 2005,**105**:2571-2576.

199. Woodward OM, Kottgen A, Coresh J, Boerwinkle E, Guggino WB, Kottgen M. Identification of a urate transporter, ABCG2, with a common functional polymorphism causing gout. *Proc Natl Acad Sci U S A* 2009,**106**:10338-10342.

200. Imai Y, Asada S, Tsukahara S, Ishikawa E, Tsuruo T, Sugimoto Y. Breast cancer resistance protein exports sulfated estrogens but not free estrogens. *Mol Pharmacol* 2003,**64**:610-618.

201. Robey RW, Honjo Y, Morisaki K, Nadjem TA, Runge S, Risbood M*, et al.* Mutations at amino-acid 482 in the ABCG2 gene affect substrate and antagonist specificity. *Br J Cancer* 2003,**89**:1971-1978.

202. Ozvegy-Laczka C, Koblos G, Sarkadi B, Varadi A. Single amino acid (482) variants of the ABCG2 multidrug transporter: major differences in transport capacity and substrate recognition. *Biochim Biophys Acta* 2005,**1668**:53-63.

203. Sarkadi B, Homolya L, Szakacs G, Varadi A. Human multidrug resistance ABCB and ABCG transporters: participation in a chemoimmunity defense system. *Physiol Rev* 2006,**86**:1179-1236.

204. Miwa M, Tsukahara S, Ishikawa E, Asada S, Imai Y, Sugimoto Y. Single amino acid substitutions in the transmembrane domains of breast cancer resistance protein (BCRP) alter cross resistance patterns in transfectants. *Int J Cancer* 2003,**107**:757-763.

205. Jonker JW, Smit JW, Brinkhuis RF, Maliepaard M, Beijnen JH, Schellens JH*, et al.* Role of breast cancer resistance protein in the bioavailability and fetal penetration of topotecan. *J Natl Cancer Inst* 2000,**92**:1651-1656.

206. Doyle L, Ross DD. Multidrug resistance mediated by the breast cancer resistance protein BCRP (ABCG2). *Oncogene* 2003,**22**:7340-7358.

207. Allen JD, Schinkel AH. Multidrug resistance and pharmacological protection mediated by the breast cancer resistance protein (BCRP/ABCG2). *Mol Cancer Ther* 2002,**1**:427-434.

208. Zhou S, Schuetz JD, Bunting KD, Colapietro AM, Sampath J, Morris JJ*, et al.* The ABC transporter Bcrp1/ABCG2 is expressed in a wide variety of stem cells and is a molecular determinant of the side-population phenotype. *Nat Med* 2001,**7**:1028-1034.

209. Scharenberg CW, Harkey MA, Torok-Storb B. The ABCG2 transporter is an efficient Hoechst 33342 efflux pump and is preferentially expressed by immature human hematopoietic progenitors. *Blood* 2002,**99**:507-512.

210. Vlaming ML, Lagas JS, Schinkel AH. Physiological and pharmacological roles of ABCG2 (BCRP): recent findings in Abcg2 knockout mice. *Adv Drug Deliv Rev* 2009,**61**:14-25.

211. Polli JW, Wring SA, Humphreys JE, Huang L, Morgan JB, Webster LO*, et al.* Rational use of in vitro P-glycoprotein assays in drug discovery. *J Pharmacol Exp Ther* 2001,**299**:620-628.

212. Taub ME, Podila L, Ely D, Almeida I. Functional assessment of multiple P-glycoprotein (P-gp) probe substrates: influence of cell line and modulator concentration on P-gp activity. *Drug Metab Dispos* 2005,**33**:1679-1687.

213. Melchior DL, Sharom FJ, Evers R, Wright GE, Chu JW, Wright SE*, et al.* Determining P-glycoprotein-drug interactions: evaluation of reconstituted P-glycoprotein in a liposomal system and LLC-MDR1 polarized cell monolayers. *J Pharmacol Toxicol Methods* 2012,**65**:64-74.

214. Bikadi Z, Hazai I, Malik D, Jemnitz K, Veres Z, Hari P*, et al.* Predicting P-glycoprotein-mediated drug transport based on support vector machine and three-dimensional crystal structure of P-glycoprotein. *PLoS One* 2011,**6**:e25815.

215. Chen L, Li Y, Yu H, Zhang L, Hou T. Computational models for predicting substrates or inhibitors of P-glycoprotein. *Drug Discov Today* 2012,**17**:343-351.

216. Palmeira A, Sousa E, Vasconcelos H, Pinto M, Fernandes MX. Structure and ligand-based design of P-glycoprotein inhibitors: a historical perspective. *Curr Pharm Des* 2012,**18**:4197-4214.

217. Szakacs G, Varadi A, Ozvegy-Laczka C, Sarkadi B. The role of ABC transporters in drug absorption, distribution, metabolism, excretion and toxicity (ADME-Tox). *Drug Discov Today* 2008,**13**:379-393.

218. Hegedus C, Szakacs G, Homolya L, Orban TI, Telbisz A, Jani M*, et al.* Ins and outs of the ABCG2 multidrug transporter: an update on in vitro functional assays. *Adv Drug Deliv Rev* 2009,**61**:47-56.

219. Sissung TM, Baum CE, Kirkland CT, Gao R, Gardner ER, Figg WD. Pharmacogenetics of membrane transporters: an update on current approaches. *Mol Biotechnol* 2010,**44**:152-167.

220. Ambudkar SV, Kim IW, Sauna ZE. The power of the pump: mechanisms of action of P-glycoprotein (ABCB1). *Eur J Pharm Sci* 2006,**27**:392-400.

221. Borgnia MJ, Eytan GD, Assaraf YG. Competition of hydrophobic peptides, cytotoxic drugs, and chemosensitizers on a common P-glycoprotein pharmacophore as revealed by its ATPase activity. *J Biol Chem* 1996,**271**:3163-3171.

222. Orlowski S, Mir LM, Belehradek J, Jr., Garrigos M. Effects of steroids and verapamil on P-glycoprotein ATPase activity: progesterone, desoxycorticosterone, corticosterone and verapamil are mutually non-exclusive modulators. *Biochem J* 1996,**317 (Pt 2)**:515-522.

223. Dey S, Ramachandra M, Pastan I, Gottesman MM, Ambudkar SV. Evidence for two nonidentical drug-interaction sites in the human P-glycoprotein. *Proc Natl Acad Sci U S A* 1997,**94**:10594-10599.

224. Garrigos M, Mir LM, Orlowski S. Competitive and non-competitive inhibition of the multidrug-resistance-associated P-glycoprotein ATPase--further experimental evidence for a multisite model. *Eur J Biochem* 1997,**244**:664-673.

225. Pascaud C, Garrigos M, Orlowski S. Multidrug resistance transporter P-glycoprotein has distinct but interacting binding sites for cytotoxic drugs and reversing agents. *Biochem J* 1998,**333 (Pt 2)**:351-358.

226. Martin C, Berridge G, Higgins CF, Mistry P, Charlton P, Callaghan R. Communication between multiple drug binding sites on P-glycoprotein. *Mol Pharmacol* 2000,**58**:624-632.

227. Shapiro AB, Ling V. Positively cooperative sites for drug transport by P-glycoprotein with distinct drug specificities. *Eur J Biochem* 1997,**250**:130-137.

228. Shapiro AB, 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* 1999,**259**:841-850.

229. Tang F, Ouyang H, Yang JZ, Borchardt RT. Bidirectional transport of rhodamine 123 and Hoechst 33342, fluorescence probes of the binding sites on P-glycoprotein, across MDCK-MDR1 cell monolayers. *J Pharm Sci* 2004,**93**:1185-1194.

230. Sharom FJ, Yu X, DiDiodato G, Chu JW. Synthetic hydrophobic peptides are substrates for P-glycoprotein and stimulate drug transport. *Biochem J* 1996,**320 (Pt 2)**:421-428.

231. Pajeva IK, Wiese M. Pharmacophore model of drugs involved in P-glycoprotein multidrug resistance: explanation of structural variety (hypothesis). *J Med Chem* 2002,**45**:5671-5686.

232. Loo TW, Bartlett MC, Clarke DM. 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* 2003,**278**:13603-13606.

233. Loo TW, Clarke DM. Location of the rhodamine-binding site in the human multidrug resistance P-glycoprotein. *J Biol Chem* 2002,**277**:44332-44338.

234. Qu Q, Sharom FJ. Proximity of bound Hoechst 33342 to the ATPase catalytic sites places the drug binding site of P-glycoprotein within the cytoplasmic membrane leaflet. *Biochemistry* 2002,**41**:4744-4752.

235. Pajeva IK, Globisch C, Wiese M. Structure-function relationships of multidrug resistance P-glycoprotein. *J Med Chem* 2004,**47**:2523-2533.

236. Clark R, Kerr ID, Callaghan R. Multiple drugbinding sites on the R482G isoform of the ABCG2 transporter. *Br J Pharmacol* 2006,**149**:506-515.

237. Baguley BC. Multidrug resistance in cancer. *Methods Mol Biol* 2010,**596**:1-14.

238. Jain R, Majumdar S, Nashed Y, Pal D, Mitra AK. Circumventing P-glycoprotein-mediated cellular efflux of quinidine by prodrug derivatization. *Mol Pharm* 2004,**1**:290-299.

239. van Vlerken LE, Duan Z, Seiden MV, Amiji MM. Modulation of intracellular ceramide using polymeric nanoparticles to overcome multidrug resistance in cancer. *Cancer Res* 2007,**67**:4843-4850.

240. Namanja HA, Emmert D, Davis DA, Campos C, Miller DS, Hrycyna CA*, et al.* Toward eradicating HIV reservoirs in the brain: inhibiting P-glycoprotein at the blood-brain barrier with prodrug abacavir dimers. *J Am Chem Soc* 2012,**134**:2976-2980.

241. Hall MD, Handley MD, Gottesman MM. Is resistance useless? Multidrug resistance and collateral sensitivity. *Trends Pharmacol Sci* 2009,**30**:546-556.

242. Pluchino KM, Hall MD, Goldsborough AS, Callaghan R, Gottesman MM. Collateral sensitivity as a strategy against cancer multidrug resistance. *Drug Resist Updat* 2012,**15**:98-105.

243. Trompier D, Chang XB, Barattin R, du Moulinet D'Hardemare A, Di Pietro A, Baubichon-Cortay H. Verapamil and its derivative trigger apoptosis through glutathione extrusion by multidrug resistance protein MRP1. *Cancer Res* 2004,**64**:4950-4956.

244. Perrotton T, Trompier D, Chang XB, Di Pietro A, Baubichon-Cortay H. (R)- and (S)-verapamil differentially modulate the multidrug-resistant protein MRP1. *J Biol Chem* 2007,**282**:31542-31548.

245. Borowski E, Bontemps-Gracz MM, Piwkowska A. Strategies for overcoming ABC-transporters-mediated multidrug resistance (MDR) of tumor cells. *Acta Biochim Pol* 2005,**52**:609-627.

246. Srivalli KMR, Lakshmi PK. Overview of P-glycoprotein inhibitors: a rational outlook. *Brazilian Journal of Pharmaceutical Sciences* 2012,**48**:353-367.

247. Lee CH. Reversing agents for ATP-binding cassette drug transporters. *Methods Mol Biol* 2010,**596**:325-340.

248. Binkhathlan Z, Lavasanifar A. P-glycoprotein inhibition as a therapeutic approach for overcoming multidrug resistance in cancer: current status and future perspectives. *Curr Cancer Drug Targets* 2013,**13**:326-346.

249. Kimchi-Sarfaty C, Oh JM, Kim IW, Sauna ZE, Calcagno AM, Ambudkar SV*, et al.* A "silent" polymorphism in the MDR1 gene changes substrate specificity. *Science* 2007,**315**:525-528.

250. Holmstock N, Annaert P, Augustijns P. Boosting of HIV protease inhibitors by ritonavir in the intestine: the relative role of cytochrome P450 and P-glycoprotein inhibition based on Caco-2 monolayers versus in situ intestinal perfusion in mice. *Drug Metab Dispos* 2012,**40**:1473-1477.

251. Wu CP, Calcagno AM, Ambudkar SV. Reversal of ABC drug transporter-mediated multidrug resistance in cancer cells: evaluation of current strategies. *Curr Mol Pharmacol* 2008,**1**:93-105.

252. Lespine A, Ménez C, Bourguinat C, Prichard RK. P-glycoproteins and other multidrug resistance transporters in the pharmacology of anthelmintics: Prospects for reversing transport-dependent anthelmintic resistance. *International Journal for Parasitology: Drugs and Drug Resistance* 2012,**2**:58-75.

253. Tao H, Weng Y, Zhuo R, Chang G, Urbatsch IL, Zhang Q. Design and synthesis of Selenazole-containing peptides for cocrystallization with P-glycoprotein. *Chembiochem* 2011,**12**:868-873.

254. Hilton BJ, Wolkowicz R. An assay to monitor HIV-1 protease activity for the identification of novel inhibitors in T-cells. *PLoS One* 2010,**5**:e10940.

255. Cardarelli CO, Aksentijevich I, Pastan I, Gottesman MM. Differential effects of P-glycoprotein inhibitors on NIH3T3 cells transfected with wild-type (G185) or mutant (V185) multidrug transporters. *Cancer Res* 1995,**55**:1086-1091.

256. Rajakuberan C, Hilton BJ, Wolkowicz R. Protocol for a mammalian cell-based assay for monitoring the HIV-1 protease activity. *Methods Mol Biol* 2012,**903**:393-405.

257. Arnaud OK, A. Ettouati, L. Terreux, R. Alame, G. Grenot, C. Dumontet, C. Di Pietro, A. Paris, J. Falson, P. Potent and fully noncompetitive peptidomimetic inhibitor of multidrug resistance P-glycoprotein. *J Med Chem* 2010,**53**:6720-6729.

258. Mosmann T. Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J Immunol Methods* 1983,**65**:55-63.

259. Bai J, Swartz DJ, Protasevich, II, Brouillette CG, Harrell PM, Hildebrandt E*, et al.* A gene optimization strategy that enhances production of fully functional P-glycoprotein in Pichia pastoris. *PLoS One* 2011,**6**:e22577.

260. Miteva MA. *In Silico Lead Discovery*: Bentham Science Publishers; 2011.

261. Alejandra Hernández-Santoyo, Aldo Yair Tenorio-Barajas, Victor Altuzar, Héctor Vivanco-Cid, Mendoza-Barrera C. Protein-Protein and Protein-Ligand Docking. *Protein Engineering - Technology and Application, Dr. Tomohisa Ogawa (Ed.)* 2013.

262. Huang SY, Zou X. Advances and challenges in protein-ligand docking. *Int J Mol Sci* 2010,**11**:3016-3034.

263. Yuriev E, Agostino M, Ramsland PA. Challenges and advances in computational docking: 2009 in review. *J Mol Recognit* 2011,**24**:149-164.

264. Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov* 2004,**3**:935-949.

265. Ranaldi F, Vanni P, Giachetti E. What students must know about the determination of enzyme kinetic parameters. *Biochemical Education* 1999,**27**:87-91.

266. Michaelis L, Menten ML, Johnson KA, Goody RS. The original Michaelis constant: translation of the 1913 Michaelis-Menten paper. *Biochemistry* 2011,**50**:8264-8269.

267. Rogers A, Gibon Y. Enzyme Kinetics: Theory and Practice. In: *Plant Metabolic Networks*. Edited by Schwender J: Springer New York; 2009. pp. 71-103.

268. Sauro HM. *Enzyme Kinetics for Systems Biology*: Ambrosius Publishing; 2012.

269. Garrett RH, Grisham CM. *Biochemistry, Update*.

270. Hill AV. The possible effects of the aggregation of the molecules of hæmoglobin on its dissociation curves. *The Journal of Physiology* 1910,**40**:i-vii.

271. Hofmeyr JH, Cornish-Bowden A. The reversible Hill equation: how to incorporate cooperative enzymes into metabolic models. *Comput Appl Biosci* 1997,**13**:377-385.

272. Adair GS, Bock WtcoAV, H. Field J. THE HEMOGLOBIN SYSTEM: VI. THE OXYGEN DISSOCIATION CURVE OF HEMOGLOBIN. *Journal of Biological Chemistry* 1925,**63**:529-545.

273. Monod J, Wyman J, Changeux JP. On the Nature of Allosteric Transitions: A Plausible Model. *J Mol Biol* 1965,**12**:88-118.

274. Koshland DE, Jr., Nemethy G, Filmer D. Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry* 1966,**5**:365-385.

275. Changeux JP. 50th anniversary of the word "allosteric". *Protein Sci* 2011,**20**:1119-1124.

276. Wagenmakers EJ, Farrell S. AIC model selection using Akaike weights. *Psychon Bull Rev* 2004,**11**:192-196.

277. Motulsky H, Melbourne ACPDPU. *Fitting Models to Biological Data Using Linear and Nonlinear Regression : A Practical Guide to Curve Fitting: A Practical Guide to Curve Fitting*: Oxford University Press, USA; 2004.

278. Team RDC. R: A language and environment for statistical computing. *R Foundation for Statistical Computing: Vienna, Austria.* 2010.

279. Coburger C, Lage H, Molnar J, Langner A, Hilgeroth A. Multidrug resistance reversal properties and cytotoxic evaluation of representatives of a novel class of HIV-1 protease inhibitors. *J Pharm Pharmacol* 2010,**62**:1704-1710.

280. Zhou T, Ohkoshi E, Shi Q, Bastow KF, Lee KH. Anti-AIDS agents 89. Identification of DCX derivatives as anti-HIV and chemosensitizing dual function agents to overcome P-gp-mediated drug resistance for AIDS therapy. *Bioorg Med Chem Lett* 2012,**22**:3219-3222.

281. Hu C, Xu D, Du W, Qian S, Wang L, Lou J*, et al.* Novel 4 beta-anilino-podophyllotoxin derivatives: design synthesis and biological evaluation as potent DNA-topoisomerase II poisons and anti-MDR agents. *Mol Biosyst* 2010,**6**:410-420.

282. Palmeira A, Vasconcelos MH, Paiva A, Fernandes MX, Pinto M, Sousa E. Dual inhibitors of P-glycoprotein and tumor cell growth: (re)discovering thioxanthones. *Biochem Pharmacol* 2012,**83**:57-68.

283. Bouvier G, Evrard-Todeschi N, Girault JP, Bertho G. Automatic clustering of docking poses in virtual screening process using self-organizing map. *Bioinformatics* 2010,**26**:53-60.

284. Primard C, Rochereau N, Luciani E, Genin C, Delair T, Paul S*, et al.* Traffic of poly(lactic acid) nanoparticulate vaccine vehicle from intestinal mucus to sub-epithelial immune competent cells. *Biomaterials* 2010,**31**:6060-6068.

285. Brik A, Wong CH. HIV-1 protease: mechanism and drug discovery. *Org Biomol Chem* 2003,**1**:5-14.

286. Waibel M, Pitrat D, Hasserodt J. On the inhibition of HIV-1 protease by hydrazino-ureas displaying the N-->C=O interaction. *Bioorg Med Chem* 2009,**17**:3671-3679.

287. Ferreira RJ, Ferreira MJ, dos Santos DJ. Molecular docking characterizes substrate-binding sites and efflux modulation mechanisms within P-glycoprotein. *J Chem Inf Model* 2013,**53**:1747-1760.

288. Pajeva IK, Sterz K, Christlieb M, Steggemann K, Marighetti F, Wiese M. Interactions of the Multidrug Resistance Modulators Tariquidar and Elacridar and their Analogues with P-glycoprotein. *ChemMedChem* 2013,**8**:1701-1713.

289. Loo TW, Clarke DM. Drug Rescue Distinguishes between Different Structural Models of Human P-Glycoprotein. *Biochemistry* 2013,**52**:7167-7169.

290. Pluta K, Kacprzak MM. Use of HIV as a gene transfer vector. *Acta Biochim Pol* 2009,**56**:531-595.

291. Alcami J. The HIV replication cycle. Established therapeutic targets and potential targets. *Enferm Infecc Microbiol Clin* 2008,**26 Suppl 12**:3-10.

292. Perno CF. The discovery and development of HIV therapy: the new challenges. *Ann Ist Super Sanita* 2011,**47**:41-43.

293. Klein I, Sarkadi B, Varadi A. An inventory of the human ABC proteins. *Biochim Biophys Acta* 1999,**1461**:237-262.

294. Schmitz G, Liebisch G, Langmann T. Lipidomic strategies to study structural and functional defects of ABC-transporters in cellular lipid trafficking. *FEBS Lett* 2006,**580**:5597-5610.

295. Wu CP, Hsieh CH, Wu YS. The emergence of drug transporter-mediated multidrug resistance to cancer chemotherapy. *Mol Pharm* 2011,**8**:1996-2011.

296. Healthcare G. *Purifying Challenging Proteins: Principles and Methods*: GE Healthcare; 2007.

297. Bill RM, Henderson PJ, Iwata S, Kunji ER, Michel H, Neutze R*, et al.* Overcoming barriers to membrane protein structure determination. *Nat Biotechnol* 2011,**29**:335-340.

298. Wiener MC. A pedestrian guide to membrane protein crystallization. *Methods* 2004,**34**:364-372.

299. Newby ZE, O'Connell JD, 3rd, Gruswitz F, Hays FA, Harries WE, Harwood IM*, et al.* A general protocol for the crystallization of membrane proteins for X-ray structural investigation. *Nat Protoc* 2009,**4**:619-637.

300. Caffrey M. Membrane protein crystallization. *J Struct Biol* 2003,**142**:108-132.

301. Caffrey M, Cherezov V. Crystallizing membrane proteins using lipidic mesophases. *Nat Protoc* 2009,**4**:706-731.

302. Matar-Merheb R, Rhimi M, Leydier A, Huche F, Galian C, Desuzinges-Mandon E*, et al.* Structuring detergents for extracting and stabilizing functional membrane proteins. *PLoS One* 2011,**6**:e18036.

APPENDIX

Tables

Table 1. Chemotherapeutic agent.

Group		Common Name	Primary Target	Mechanism	Similar Drugs
Chemotherapy drugs					
Cell cycle specific agents (CCSA)	antimetabolites	5-fluorouracil	thymdylate synthase, also incorporated into RNA and DNA	mimics deoxyuridine triphosphate	5-azacytidine cytarabine 6-mercaptopurine
		Etoposide	topoisomerase II	inhibits reconnection of DNA	teniposide
		Topotecan	topoisomerase I	inhibits reconnection of DNA	irinotecan
	antifoliates	methotrexate	dihydrofolate reductase	mimics folate	Trimetrexate
	plant alkaloids and mitotic inhibitors	Taxanes	microtubules	inhibits tubule depolymerization	paclitaxel (Taxol®) and docetaxel (Taxotere®)
		Vinca alkaloids	microtubules	inhibits tubule assembly	vinblastine (Velban®), vincristine (Oncovin®), and vinorelbine (Navelbine®)
	some cytotoxic antibiotics	bleomycin	DNA	breaks in the DNA strands within the cell	Blenoxane®
CELL CYCLE NONSPECIFIC AGENTS	alkylating agents,	cyclophosphamide	DNA	nitrogen mustard	ifosfamide melphalan
		carmustine	DNA	nitrosoureas	streptozocin and lomustine
		dacarbazine	DNA	triazines	temozolomid
	platinum compounds	cisplatin	DNA	platinum coordination	carboplatin
	cytotoxic antibiotics	doxorubicin (Adriamycin),	DNA	polycyclic rings allow intercalation	bleomycin mitomycin C daunorubicin mitoxantrone

Other therapies

TARGETED THERAPIES	inhibit a specific protein involved in cancer progression	imatinib	receptor tyrosine kinase (RTK)	inhibits BCR-ABL enzyme	gefitinib, erlotinib, sunitinib, and bortezomib
		Bortezomib	Proteasome inhibitor	activation of apoptosis (programmed cell death)	
HORMONES	Corticosteroids	Dexamethasone	glucocorticoid receptor	modify DNA transcription	hydrocortisone prednisone cortisone
	Estrogens	Diethylstilbestrol	estrogen receptors	change hormonal balance	estradiol modified estrogens
	anti-estrogens	Tamoxifen	estrogen receptors	blocks receptors in receptor-positive breast cancer	torimenifene
IMMUNOTHERAPY	Monoclonal antibody	rituximab (Rituxan®)		use antibodies to attack the cancer cell	alemtuzumab (Campath®)
	Non-specific immunotherapies	BCG		boost the immune response	interleukin-2 (IL-2), and interferon-alfa
	Immune-modulating	thalidomide		inhibits different molecular targets(e.g., angiogenesis)	lenalidomide (Revlimid®)
	Cancer vaccines	Provenge®	advanced prostate cancer	stimulate the body's own immune system	

Non-exhaustive list adapted from: ***http://www.chemistryexplained.com/Ce-Co/Chemotherapy.html#b***

Table 2. Genome of HIV-1 with functions of viral proteins.

Genome	Category	Gene / origin	protein produced	Protein function
	Structural	*gag*/ viral protease-me- diated cleavage of Gag polyprotein (Pr55) during vi-rion maturation	p17, matrix (MA)	forms the outer layer of the core protein
			p24, capsid (CA)	forms the inner layer of the core protein
			p7, nucleocapsid (NC)	binds directly to the RNA genome
			SP1	Regulating the rate of cleavage of the fusion with the cell membrane and the formation of the virion
			SP2	
			P6	Ribonucleoprotein, viral packaging
		pol / cleavage of the Gag-Pol polyprotein (160 Pr) PR during maturation of the virion	P6*	stabilizing interactions Gag-Gag-Pol and activation of PR.
	Enzymes		p66/p51 and P15, reverse transcriptase (RT)	viral cDNA synthesis
			p11, protease (PR)	Cleaves the p160 protein. Virion maturation
			p31, integrase (IN)	integrase
	envelope proteins	*env* / gp160 cleavage by cellular proteases	gp41, (TM)	is a membrane protein associated to gp120, necessary for the fusion
			gp 120, (SU)	It occupies a peripheral position of the envelope level and binds to the CD4 receptor
	regulatory proteins	***tat***	p14	Activates transcription of proviral DNA, promoting the increased level of synthesis of viral proteins
		rev	p19	Allows increased mRNAs corresponding to gag, pol and env and other genes such as *vif*, *nef, vpu* ou *vpr*
	accessory proteins	***vif***	p23	Is the origin of the infectivity of the virus particle
		nef	p27	Increases viral replication reduces the number of host cells
		vpu	p16	Is necessary for efficient viral assembly and budding
		vpr	p15	Active low transcription of proviral DNA

LTR: Long Terminal Repeat, the RNA sequences repeated at both ends of HIV's genetic material
Adapted from [290, 291].

Table 3. Therapeutic anti-HIV drugs in clinical use.

Classes of drugs	Drug	Abbreviation / (commercial name)	Year of approval by the FDA	drugs abandoned
Reverse transcriptase inhibitors (RTI)	Nucleoside/nucleotide analog reverse transcriptase inhibitors, NRTI: They mimic the nucleic acid bases (in parenthesis), bind competitively the reverse transcriptase resulting in the termination of the DNA chain			
	Zidovudine (thymidine)	AZT/ (Retrovir®)	1987	
	Didanoside (adenosine)	ddI / (Videx®)	1991	
	Zalcitabine (cytosine)	ddC / (Hivid®)	1992	×
	Stavudine (thymidine)	d4T / (Zerit®)	1994	
	Lamivudine (cytidine)	3TC / (Epivir®)	1995	
	Abacavir (guanine)	ABC / (Ziagen®)	1998	
	Tenofovir (adenosine)	TDF / (Viread®)	2001	Replaced by TDF
	Emcitrabine (cytidine)	FTC / (Emtriva®)	2004	
	Non-nucleoside analog reverse transcriptase inhibitors, NNRTI: Bind an allosteric site, P66 subunit (non-competitively) of the reverse transcriptase, induces a conformational change and prevent substrate binding.			
	Nevirapine	NVP / (Viramune®)	1996	
	Efavirenz	EFV / (Sustiva®)	1997	
	Delavirdine	DLV / (Rescriptor®)	1998	
	Etravirine	ETR / (Intelence®)	2008	
Protease inhibitors (PI)	Bind to HIV protease and prevent subsequent cleavage of polypeptides. Thus, the immature viruses cannot bud from the cell and infect new cells.			
	Saquinavir	SQV/(Fortovase®,Invirase®)	1995	
	Ritonavir	RTV / (Norvir®)	1996	
	Indinavir	IND / (Crixivan®)	1996	
	Nelfinavir	NFV / (Viracept®)	1997	
	Amprenavir	APV / (Agenerase®)	1999	Replaced by FSV
	Lopinavir	LPV / (avec RTV Kaletra®)	2000	
	Atazanavir	ATV / (Reyataz®)	2003	
	Fosamprenavir	FSV / (Lexiva®)	2003	
	Tipranavir	TPV / (Aptivus®)	2005	
	Darunavir	DRV / (Prezista®)	2006	
Integrase inhibitors (INI)	They prevent the incorporation of the proviral DNA into the DNA of the host cell			
	Raltegravir	RAL / (Isentress®)	2007	
	Elvitegravir	EVG/(GS-9137)	Phase III	
	CX00287: blocks the interaction of LEDGF/p75 with Integrase		Under development	

Entry inhibitors	Fusion inhibitors: It binds to the triple helical region in gp41 of HIV, stabilizing its conformation. So it blocks the fusion and prevents infection of target cells.		
	Enfuvirtide	T20 / (Fuzeon®)	2003
	Attachment inhibitors: humanized monoclonal IgG4 antibody against CD4 receptors		
	Ibalizumab	TMB-355	Phase II
	CCR5 antagonists: It prevents the gp120/CCR5 interaction necessary for viral entry		
	Maraviroc	MVC / (Selzentry®)	2007
	PRO140		Phase III
Maturation inhibitors	Target gag polyprotein precursor, which is vital in final stage of virus development for a mature and infectious virion		
	Bevirimat	PA-457	Phase II
Capsid assembly inhibitors	Bind in a unique pocket on capsid affecting the integrity and survival of the virus in the assembly of viral particles and in the infection of host cells.		
	PF-3450071 and PF-3450074		Under development

TDF:Tenofovir disoproxil fumarate; LEDGF: Lens epithelium derived growth factor inhibitor

Adapted from [14, 16-18, 292]

Table 4. Associated mutations in viral proteins associated with resistance to anti-HIV drugs.

Mutations in	Associated with resistance to	Amino acid: M Wild-type / 300 position / R or R/S/T substitution(s)	Drug
Reverse transcriptase gene	Nucleoside/nucleotide analog reverse transcriptase inhibitors, NRTI	M41L, A62V, 69 inserts, K70R, L210W, T215Y/F, K219Q/E	Multi-nRTI Resistance: *69 inserts* affects all nRTIs
		A62V, V75I, F77L, F116V, Q151M	Multi-nRTI Resistance: *151 Complex* affects all nRTIs (except Tenofovir)
		M41L, D67N, K70R, L210W, T215Y/F, K219Q/E	Multi-nRTI Resistance: *Thymidine Analogue* affects all nRTIs
		K65R, L74V, Y115F, M184V	Abacavir
		K65R, L74V	Didanosine
		K65R, M184V/I	Emtricitabine and Lamivudine
		M41L, K65R, D67N, K70R, L210W, T215Y/F, K219Q/E	Stavudine
		K65R, K70E	Tenofovir
		M41L, D67N, K70R, L210W, T215Y/F, K219Q/E	Zidovudine
	Non-nucleoside analog reverse transcriptase inhibitors, NNRTI	L100I, K101P, K103N/S, V106M,V108I, Y181C/I, Y188L, G190S/A, P225H, M230L	Efavirenz
		V90I, A98G, L100I, K101E/H/P, V106I, E138A/G/K/Q, V179D/F/T, Y181C/I/V, G190S/A, M230L	Etravirine
		L100I, K101P, K103N/S, V106A/M, V108I, Y181C/I, Y188C/L/H, G190A, M230L	Nevirapine
		K101E/P, E138A/G/K/Q/R, Y179L, Y181C/I/V, Y188L H221Y, F227C, M230I/L	Rilpivirine
Protease gene	Protease inhibitors (PI)	L10 I/F/V/C/G, G16E, K20R/M/I/T/V, L24I, V32I, L33I/F/V, E34Q, M36I/L/V, M46I/L, G48V, I50L, F53L/Y, I54L/V/M/TA, D60E, I62V, I64L/M/V, A71V/I/T/L, G73C/S/T/A, V82A/T/F/I, I84V, I85V, N88S, L90M, I93L/M	Atazanavir +/- ritonavir
		V11I,V32I,L33F, I47V, I50V, I54M/L, T74P, L76V, I84V, L89V L10F/I/R/V, V32I, M46I/L, I47V, I50V, I54L/V/M, G73S, L76V, V82A,/F/S/T, I84V, L90M	Darunavir / ritonavir Fosamprenavir / ritonavir
		L10I/R/V, K20M/R, L24I, V32I,M36I, M46I/L, I54V, A71V/T,G73S/A, L76V, V77I, V82A/F/T, I84V, L90M	Indinavir / ritonavir
		L10F/I/R/V, K20M/R, L24I, V32I, L33F, M46I/L, I47V/A, I50V, F53L, I54V/L/A/M/T/S, L63P, A71V/T, G73S, L76V, V82A/F/T/S, I84V, L90M	Lopinavir / ritonavir
		L10F/I, D30N, M36I, M46I/L, A71V/T, V77V/T, V77I, V82A/F/T/S, I84V, N88D/S, L90M	Nelfinavir
		L10I/R/V, L24I, G48V, I54V/L, I62V, A71V/T, G73S, V77I, V82A/F/T/S, I84V, L90M	Saquinavir / ritonavir
		L10V, L33F, M36I/L/V, K43T,M46L, I47V, I54A/M/V, Q58E, H69K/R, T74P, V82L/T, N83D, I84V, L89I/M/V	Tipranavir / ritonavir
envelope gene	entry inhibitors	G36D/S, I37V, V38A/M/E, Q39R, Q40H, N42T,N43D	Enfuvirtide
		- mutations are found in the V3 loop of gp120 - overgrowth of CXCR4, X4-tropic strains	Maraviroc
integrase gene	integrase inhibitors	L74M, E92Q, T97A, E138A/K, G140A/S, Y143R/H/C, Q148H/K/R, N155H	Raltegravir
		T66,I/A/K, E92Q/G, T97A, S147G, Q148R/H/K, N155H	Elvitegravir

Amino acid abbreviations:
A: alanine; C: cysteine; D: aspartate; E: glutamate; F: phenylalanine; G: glycine; H: histidine; I: isoleucine; K: lysine; L: leucine; M: methionine; N: asparagine; P: proline; Q: glutamine; R: arginine; S: serine; T: threonine; V: valine; W: tryptophan; Y: tyrosine.
Non-exhaustive list adapted from [19]

Table 5. Locations, functions and pathologies associated with ABC transporters

	Members (alias), **AA/ molecular weight in Da** and **domain arrangement**	**Expression**	**Function**	**Mendelian disorders**
A / ABC1	**ABCA1** (ABC1, CERT, TGD), 2261/254286, (TMD-NBD)2	Ubiquitous	Cholesterol efflux onto HDL	Tangier disease, familial hypoapoproteinemia
	ABCA2 (ABC2, STGD), 2436/269974, (TMD-NBD)2	Brain, kidney, lung, heart Lung,	**Drug resistance**, role in macrophage lipid metabolism and neutral development, trafficking of LDL	Alzheimer's disease
	ABCA3 (ABC3), 1704/191362, (TMD-NBD)2	Lung, lamellar bodies in type II cells	**Drug resistance**, surfactants secretion, implication in MDR suggested	Neonatal surfactant deficiency
	ABCA4 (ABCR), 2273/255944, (TMD-NBD)2	Rod and cone photoreceptors	N-Retinylidiene-PE efflux	Stargardt, fundus flavimaculatis, retinitis pigmentosum, cone-rod dystrophy
	ABCA5 (ABC13), 1642/186508, (TMD-NBD)2	Brain, lung, heart, and thyroid gland, lysosomes and late endosomes	Unknown yet. Importance in cardiomyocites and follicular cells. Relationship suggested with lysosomal diseases, including DCM (dilated cardiomyopathy)	
	ABCA6, 1617/184286, (TMD-NBD)2	Ubiquitous, highest expression in liver, heart and brain	Potentially involved in macrophage lipid homeostasis	
	ABCA7 (ABCX), 2146/234364, (TMD-NBD)2	Spleen, thymus, bone marrow, peripheral leukocytes	Mediates apo-lipoprotein-derived generation of HDL, role in the clearance of apoptic cells	Sjogren syndrome
	ABCA8, 1581/179245, (TMD-NBD)2	In most organs, especially in heart, skeletal muscle, and liver	Function unknown	
	ABCA9, 1624/184362, (TMD-NBD)2	Heart, brain, fetal tissue	Likely involved in monocyte different differentiation and macrophage lipid homeostasis	
	ABCA10, 1543/175746, (TMD-NBD)2	Heart, brain, gastro-intestinal tract, monocytes	Involvement suggested in macrophage lipid homeostasis	
	ABCA12, 2595/293251, (TMD-NBD)2	Lamellar granules and lamellar bosies	Lipid metabolism	Lamellar ichtyosis type II, Harlequin ichtyosis
	ABCA13, 5058/576239, (TMD-NBD)2	Human trachea, testes, bone marrow	Candidate for several pathologies such as Shwachman–Diamond syndrome and cancers (CNS, prostate, leukemia)	
B / MDR - TAP	**ABCB1** (MDR1, P-gp, PGY1), 1280/141463, (TMD-NBD)2	Many tissues especially those with barrier functions: liver, BBB, kidney, intestine, placenta, apical membranes	**Multidrug resistance**. Provide protection against hydrophobic xenobiotic	Ivermectin susceptibility, ulcerative colitis
	ABCB2 (TAP1), 748/80965, TMD-NBD	Most cells, ER	Peptide transport	Immune deficiency
	ABCB3 (TAP2), 686/75664, TMD-NBD	Most cells, ER	Peptide transport	Immune deficiency

	ABCB4 (MDR2-3, PFIC3, PGY3), 1279/140682, (TMD-NBD)2	Liver, apical membrane	**Drug resistance**, phosphatidylcholine transport	Progressive familial intrahepatic cholestasis type 3 (PFIC-3)
	ABCB5, 812/89831, (TMD-NBD)2	Testes, breast, melanocytes, melanomas	**Drug resistance**, membrane potential and regulator of cell fusion in physiologic skin progenitor cells, melanomas	
	ABCB6 (MTABC3), 842/93886, TMD-NBD	Mitochondria	Heme transport and	
	ABCB7 (ABC7, Atmlp), 752/82641, TMD-NBD	Mitochondria	Fe/S cluster transport, biosynthesis of heme, essential for Hematopoiesis	X-linked sideroblastosis and anemia (XLSA/A)
	ABCB8 (MABC1), 735/79948, TMD-NBD	Mitochondria	Function not determined yet. This protein may also play a role in the transport of phospholipids, heme as well as peptides	
	ABCB9 (TAPL), 766, TMD-NBD	Testes, moderately in the brain and spinal	Peptide transport	
	ABCB10 (MTABC2), 738/79009, TMD-NBD	Mitochondria	Heme biosynthesis	
	ABCB11 (BSEP, PFIC2, SP-GP), 1321/146393, (TMD-NBD)2	Liver	**Drug resistance**, bile salts transport	Progressive familial intrahepatic cholestasis type 2 (PFIC-2)
C / MRP	**ABCC1** (MRP1), 1531/171561, TMD0(TMD-NBD)2	Lung, testes, Peripheral Blood Mononuclear Cell, lateral membrane	**Multidrug resistance**, organic anion transporter (involved in inflammation mediated by LTC4)	
	ABCC2 (MRP2, CMOAT), 1545/174191, TMD0(TMD-NBD)2	Liver, intestine, kidney, apical membrane	**Drug resistance**, organic anion efflux including bile salt transport	Dubbin-Johnson Syndrome (DJS)
	ABCC3 (MRP3, CMOATP2), 1527/169343, TMD0(TMD-NBD)2	Lung, intestine, liver, kidney, lateral membrane	**Drug resistance**, organic anion efflux including bile salt	
	ABCC4 (MRP4, MOATB), 1325/149541, (TMD-NBD)2	Many tissues	**Drug resistance**, organic anion efflux including bile acids and salts	
	ABCC5 (MRP5, MOATC), 1437/160660, (TMD-NBD)2	Many tissues	**Drug resistance**, organic anion efflux	
	ABCC6 (MRP6, MOATE, PXE), 1503/164904, TMD0(TMD-NBD)2	Kidney, liver, lateral membranes	**Drug resistance**, physiological function undetermined	Pseudoxanthoma elasticum (PXE)
	ABCC7 (CFTR), 1480/168174, (TMD-NBD)2	Exocrine tissue, apical membranes	Chloride ion channel	Cystic Fibrosis, congenital bilateral absence of the vas deferens
	ABCC8 (SUR1), 1580/176891, TMD0 (TMD-NBD)2	Pancreas	Sulfonylurea receptor, modulator of ATP-sensitive K+ channels and insulin release;	Persistent hyperinsulinemic hypoglycemia of infancy (PHHI)
	ABCC9 (SUR2), 1549/174260, TMD0 (TMD-NBD)2	Skeletal muscle, heart	K(ATP) channel regulation	Dilated cardiomyopathy with ventricular tachycardia

	ABCC10 (MRP7), 1492/161629, TMD0(TMD-NBD)2	Low expression in all tissues except pancreas	**Drug resistance**, E(2)17betaG transport	
	ABCC11 (MRP8), 1382/154301, (TMD-NBD)2	Low expression in all tissues except kidney, spleen, colon, brain	**Drug resistance**, E(2)17betaG, steroid sulfates, glutathione conjugates and monoanionic bile acids	
	ABCC12 (MRP9), 1359/152244, (TMD-NBD)2	Breast, testes, brain, skeletal, and ovary	Function unknown. Potential target for the immunotherapy of breast cancer	
	ABCC13 (PRED6), 325,			
D / ALD	**ABCD1** (ALD, ALDP), 745/82937, TM-NBD	Peroxisome	VLCFA transport regulation	Adrenoleukodystrophy (ALD)
D / ALD	**ABCD2** (ALDL1, ALDR), 740/83233, TM-NBD	Peroxisome	Function unknown. Modify the ALD phenotype	
D / ALD	**ABCD3** (PXMP1, PMP70), 659/75476, TM-NBD	Peroxisome	Peroxisome biogenesis	
D / ALD	**ABCD4** (PMP69, PXMP1L), 606/68596, TM-NBD	Peroxisome	Function not determined yet. May modify the ALD phenotype. May also play a role in the process of peroxisome biogenesis	
E/OABP	**ABCE1** (RTI, RNS4I, HP68), 599/67314, (NBD)2	Ubiquitous	Inhibition of RNase L, required for the assembly of HIV-1 and other lentiviruses, role in the initiation of translation	
F/GCN20	**ABCF1** (ABC50), 845/95926, (NBD)2	Ubiquitous	mRNA translation	
F/GCN20	**ABCF2**, 623/71290, (NBD)2	Ubiquitous	mRNA translation	
F/GCN20	**ABCF3**, 709/79745, (NBD)2	Ubiquitous	mRNA translation	
G/WHITE	**ABCG1** (WHITE1), 678/75592, NBD-TMD	Brain, spleen, lung, liver, macrophages	Cholesterol efflux to HDL and tissue lipid homeostasis	
G/WHITE	**ABCG2** (BCRP, MXR, ABCP), 655/72314, NBD-TMD	Placenta, intestine, liver, breast, apical membranes	**Multidrug resistance**, Riboflavin (vitamin B2) pump, heme efflux	Protoporphyria IX
G/WHITE	**ABCG3**			
G/WHITE	**ABCG4** (WHITE2), 646/71896, NBD-TMD	Macrophage, brain, eye, spleen	Cholesterol efflux to HDL	
G/WHITE	**ABCG5** (Sterolin 1), 673/75679, NBD-TMD	Liver, small intestine	Sterol transport	β-Sitosterolemia
G/WHITE	**ABCG8** (Sterolin), 673/75679, NBD-TMD	Liver, small intestine	Sterol 2 transport	β-Sitosterolemia

HDL: high density lipoprotein; LDL: low density lipoprotein; VLCFA: very long chain fatty acid; LTC4: Leukotriene C4; CNS: Central nervous system; OABP: oligoadenylate binding protein; TMD: transmembrane domain; NBD: nucleotide binding domain; AA: number of amino acids; apical membrane: portion of the cell exposed to the lumen; basolateral membrane: rest of the cell (i.e. sides and base)

Adapted from [77, 293, 294]. More information: http://nutrigene.4t.com/humanabc.htm.

Table 6. % of sequence identity of human and prokaryotic ABC transporters.

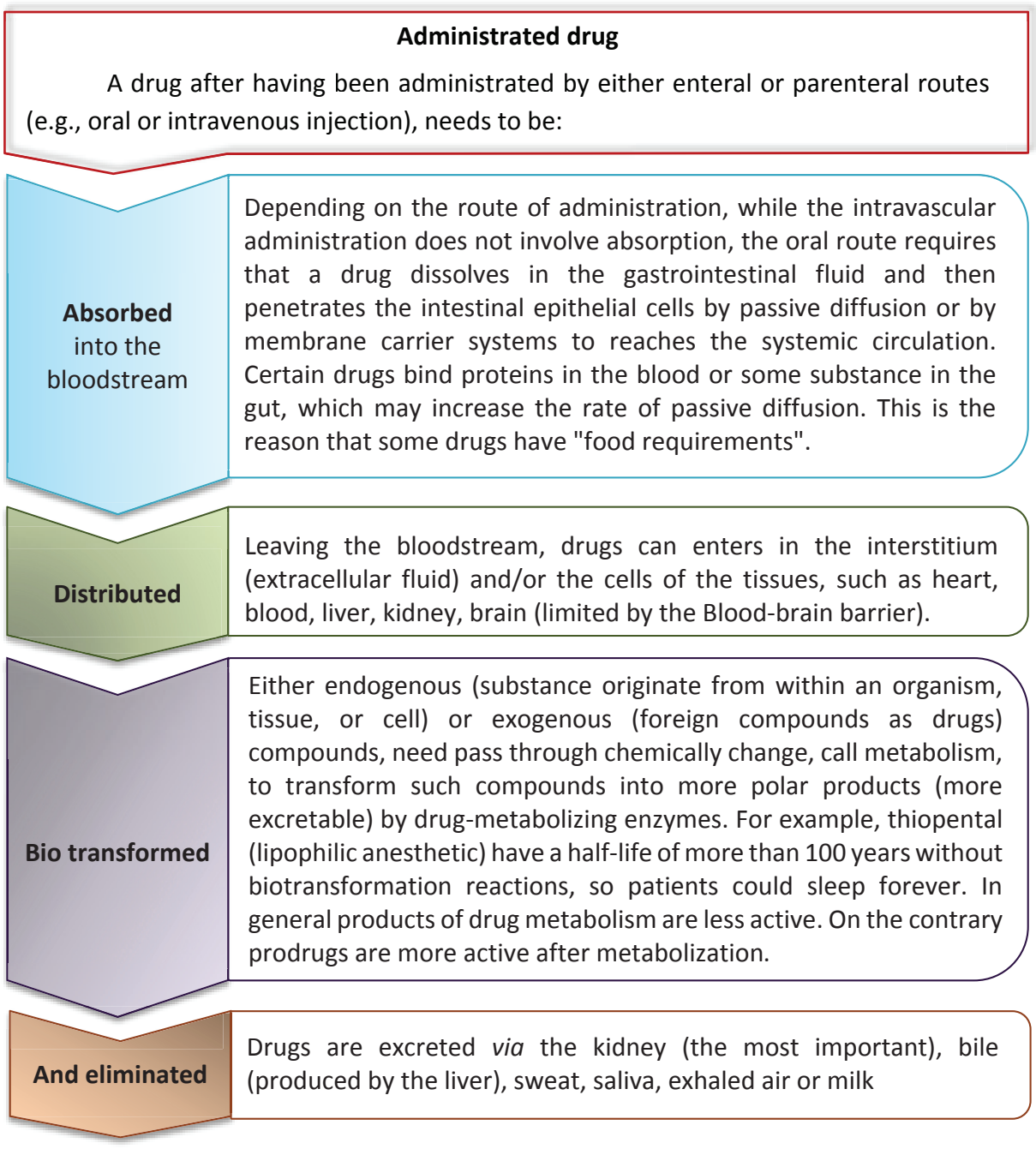
		P-gp		MRP1		ABCG2	MsbA	LmrA	Sav1866
		NBD1	NBD2	NBD1	NBD2	NBD			
P-gp	NBD1	100							
	NBD2	61	100						
MRP1	NBD1	28	29	100					
	NBD2	32	32	26	100				
ABCG2	NBD	18	21	12	14	100			
MsbA		53	48	32	35	13	100		
LmrA		44	40	24	35	24	40	100	
Sav1866		49	47	33	38	23	55	50	100

		P-gp		MRP1		ABCG2	MsbA	LmrA	Sav1866
		TMD1	TMD2	TMD1	TMD2	TMD			
P-gp	TMD1	100							
	TMD2	26	100						
MRP1	TMD1	8	7	100					
	TMD2	10	9	11	100				
ABCG2	TMD	4	3	8	8	100			
MsbA		17	16	8	14	8	100		
LmrA		16	25	11	8	6	16	100	
Sav1866		13	15	13	13	6	18	20	100

According to [175]

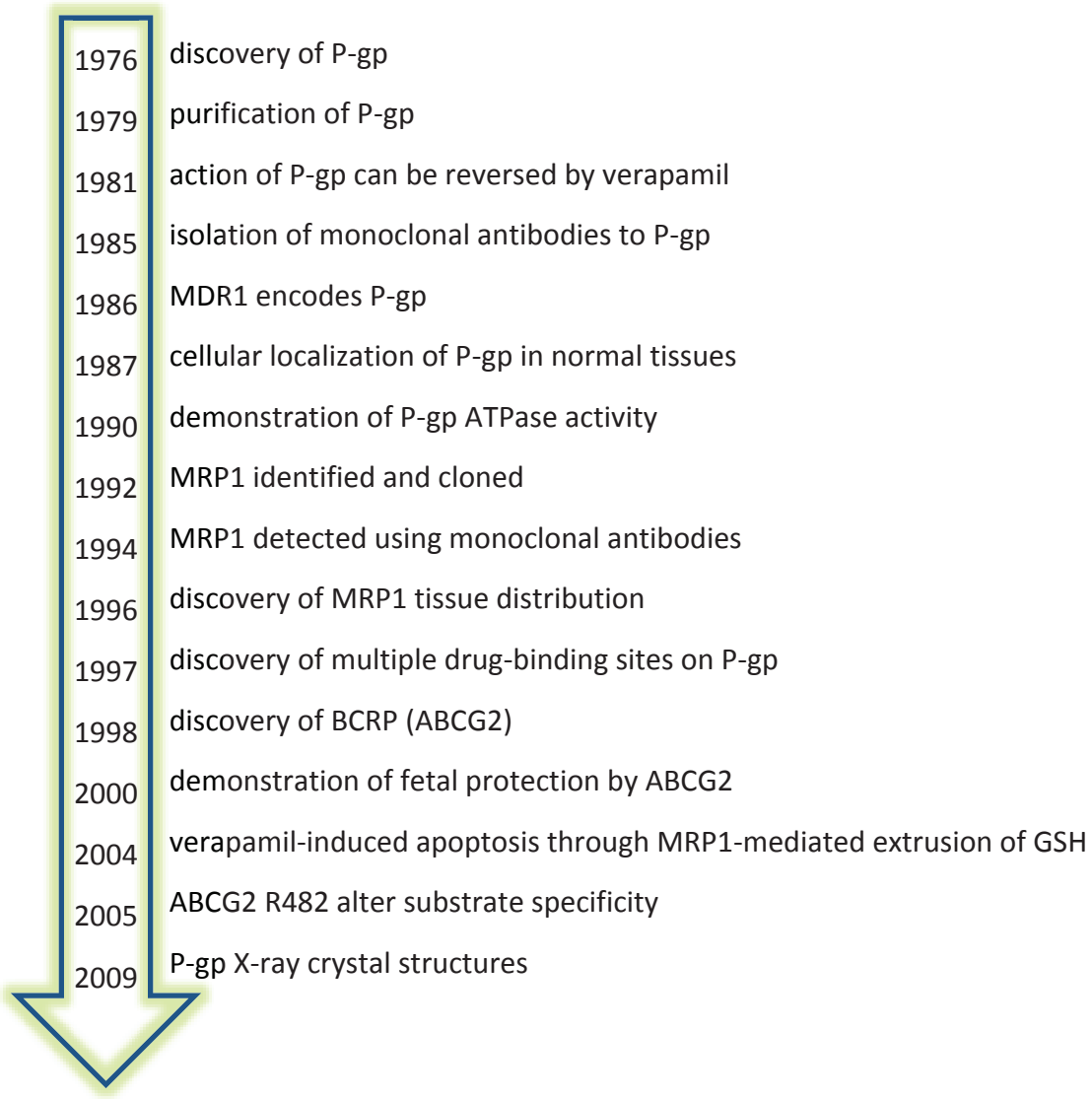
Figures

Figure 1. Drug absorption, distribution, metabolism and excretion (ADME).



Adapted from the website:
http://www.columbia.edu/itc/gsas/g9600/2004/GrazianoReadings/Drugabs.pdf

Figure 2. Timeline of Major Discoveries of MDR-Linked Human ABC Transporters: P-gp, MRP1 and ABCG2



According to [295].

Figure 3. AutoDock-vina scripts

```
Input:
  --receptor arg        rigid part of the receptor (PDBQT)
  --flex arg            flexible side chains, if any (PDBQT)
  --ligand arg          ligand (PDBQT)

Search space (required):
  --center_x arg        X coordinate of the center
  --center_y arg        Y coordinate of the center
  --center_z arg        Z coordinate of the center
  --size_x arg          size in the X dimension (Angstroms)
  --size_y arg          size in the Y dimension (Angstroms)
  --size_z arg          size in the Z dimension (Angstroms)

Output (optional):
  --out arg             output models (PDBQT), the default is chosen based on
                        the ligand file name
  --log arg             optionally, write log file

Misc (optional):
  --cpu arg                 the number of CPUs to use (the default is to try to
                            detect the number of CPUs or, failing that, use 1)
  --seed arg                explicit random seed
  --exhaustiveness arg (=8) exhaustiveness of the global search (roughly
                            proportional to time): 1+
  --num_modes arg (=9)      maximum number of binding modes to generate
  --energy_range arg (=3)   maximum energy difference between the best binding
                            mode and the worst one displayed (kcal/mol)

Configuration file (optional):
  --config arg          the above options can be put here

Information (optional):
  --help                display usage summary
  --help_advanced       display usage summary with advanced options
  --version             display program version
```

According to the website: http://vina.scripps.edu/manual.html

Figure 4. R-scripts written to determine the kinetics parameters in chapter III.

SCRIPT H33342-QZ59RRR.r

```
# MIXTE (Ki1 != Ki2), eq. 6.1
>modelMIX= as.formula("TR~Vm/(1+I/Ki2)*S/(Km*(1+I/Ki1)/(1+I/Ki2)+S)")
>fitMIX=nls(modelMIX, data=data, start=list(Vm=30, Km=0.8, Ki1=0.5,
Ki2=0.5) , )
>summary(fitMIX)
>AIC(fitMIX, k=2)
      #Result:
      #Parameters:
      #     Estimate    Std. Error  t value     Pr(>|t|)
      #Vm   52.0033     2.9391      17.694      < 2e-16 ***
      #Km    2.3747     0.2975       7.983      8.72e-12 ***
      #Ki1   4.7861     2.0487       2.336        0.022 *
      #Ki2   1.6349     0.2436      6.711       2.54e-09 ***
      #Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
      #Residual standard error: 2.086 on 80 degrees of freedom
      #Number of iterations to convergence: 6
      #Achieved convergence tolerance: 3.203e-06
      #AIC = 367.7835
```

SCRIPT_Dauno-QZ59RRR.r

```
# non competitive inhibition by inhibitor + non competitive
inhibition by substrate enhanced by inhibitor,  eq. 4.9
>modelINHIBSUBSTR_NC.NCI=
as.formula("TR~Vm/(1+S/KSI*I/KI)/(1+I/KI)*S/(Km+S)")
>fitINHIBSUBSTR_NC.NCI=nls(modelINHIBSUBSTR_NC.NCI, data=data,
start=list(Vm=30, Km=0.9, KI=1, KSI=5) , )
>summary(fitINHIBSUBSTR_NC.NCI)
>AIC(fitINHIBSUBSTR_NC.NCI)

      #Results:
      #Parameters:
      #     Estimate    Std. Error  t value     Pr(>|t|)
      #Vm   40.7901     1.4815      27.534       < 2e-16 ***
      #Km    1.0088     0.1196      8.434       6.97e-13 ***
      #KI    1.9524     0.3594      5.432       5.11e-07 ***
      #KSI   6.9049     3.1682      2.179        0.032 *
      #Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
      #Residual standard error: 2.015 on 86 degrees of freedom
      #Number of iterations to convergence: 22
      #Achieved convergence tolerance: 9.512e-06
      #AIC = 387.4362
```

SCRIPT H33342-QZ59RRR.r

```
# cooperativité et inhibition competitive + activation du transport
et de la cooperativité
#equation 3.5
>modelIHN.ACT.HILL.MOD=
as.formula("TR~Vm*(1+I/KA)*S**(h*(1+I/KA))/(Km**(h*(1+I/KA))*(1+I/KI)
/(1+I/KA)+S**(h*(1+I/KA)) ) ")
>fitIHN.ACT.HILL.MOD=nls(modelIHN.ACT.HILL.MOD, data=data,
start=list(Vm=30, Km=2, KI=0.1, KA=4, h=1.1) , )
>summary(fitIHN.ACT.HILL.MOD)
>AIC(fitIHN.ACT.HILL.MOD)

      #Result:
      #Parameters:
      #     Estimate    Std. Error  t value     Pr(>|t|)
      #Vm 31.45302    3.53921       8.887       6.10e-13 ***
      #Km  1.82194    0.44134       4.128       0.000103 ***
      #KI  0.15399    0.03667       4.199       8.08e-05 ***
      #KA  4.45655    1.00018       4.456       3.26e-05 ***
      #h   1.18563    0.15282       7.758       6.57e-11 ***
      #Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
      #Residual standard error: 1.713 on 67 degrees of freedom
      #Number of iterations to convergence: 7
      #Achieved convergence tolerance: 1.274e-06
      #AIC = 288.625
```

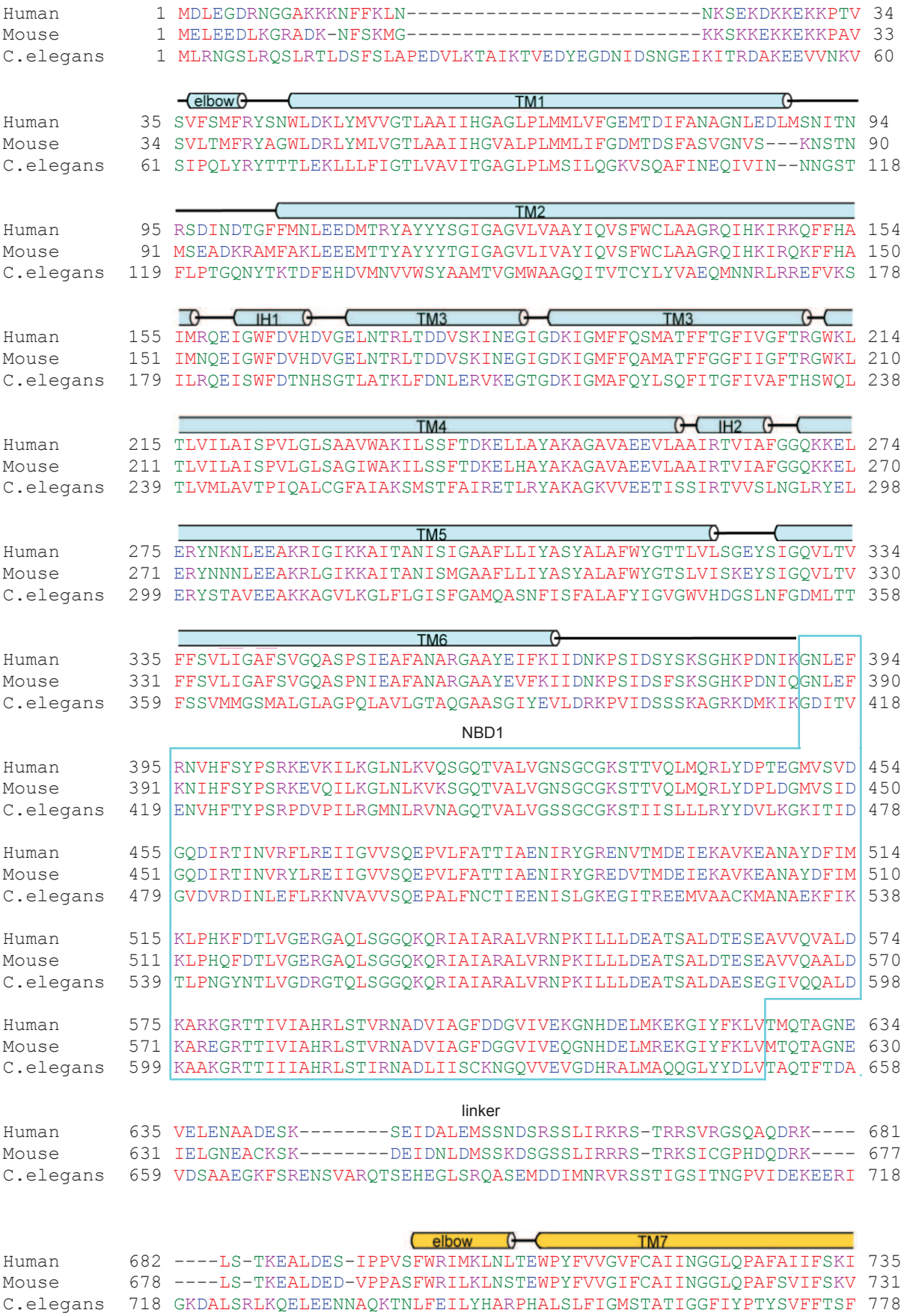
SCRIPT_Dauno-QZ59RRR.r

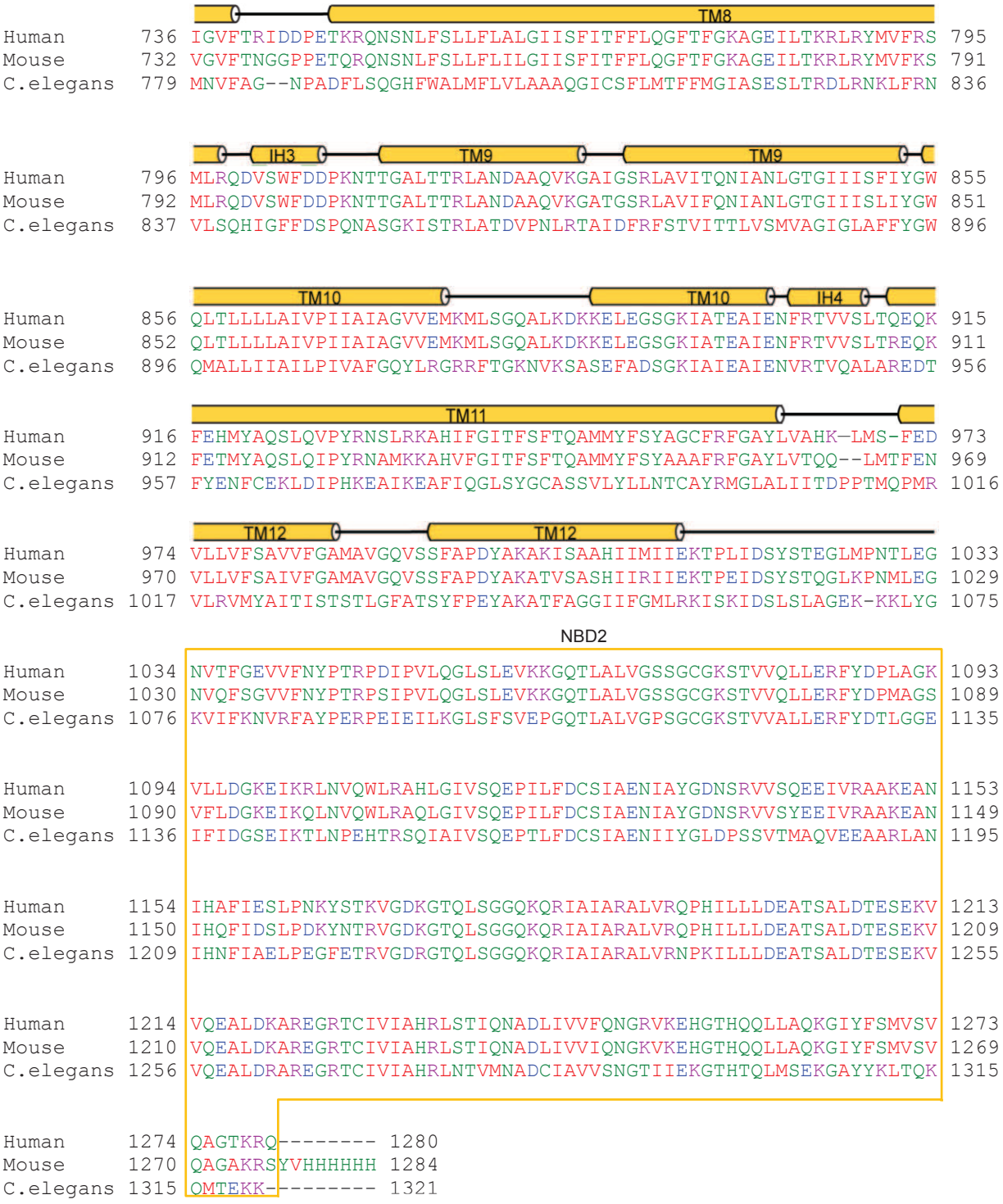
```
# mixed inhibition by inhibitor + non competitive inhibition by
substrate enhanced by inhibitor, eq. 6.9
>modelINHIBSUBSTR_NC.MI=
as.formula("TR~Vm/(1+S/KSI*I/Ki2)/(1+I/Ki2)*S/(Km*(1+I/Ki1)/(1+I/Ki2)
+S)")
>fitINHIBSUBSTR_NC.MI=nls(modelINHIBSUBSTR_NC.MI, data=data,
start=list(Vm=30, Km=0.9, Ki1=5, Ki2=5, KSI=10) , )
>summary(fitINHIBSUBSTR_NC.MI)
>AIC(fitINHIBSUBSTR_NC.MI)

      #Results:
      #Parameters:
      #     Estimate    Std. Error  t value     Pr(>|t|)
      #Vm  39.12378    1.41876      27.576      < 2e-16 ***
      #Km   0.97579    0.11731       8.318      6.44e-12 ***
      #Ki1  0.35454    0.08498       4.172      8.89e-05 ***
      #Ki2  1.28280    0.89252      1.437        0.155
      #KSI 16.09578   25.69466      0.626        0.533
      #Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
      #Residual standard error: 1.796 on 67 degrees of freedom
      #Number of iterations to convergence: 9
      #Achieved convergence tolerance: 7.925e-06
      #AIC = 295.4831
```

Each "line" of the R-script (= each instruction) is indicated by the sign ">"

Figure 5. Amino acid sequence alignment of human, mouse and *C. elegans* P-gp





Secondary structure of P-gp is indicated according to [167].
Sequence identities: human *vs* mouse: 86.7 %, human *vs* c. elegans: 46.2 %, mouse *vs* c. elegans: 43.5 %.

Workflow of crystallographic study of a membrane protein

This part describe the problems posed by the crystallographic study of a membrane protein, since over-expression until crystallization (reviewed in [296-301]). It should be note that strategies used to get the structure are often a case-by-case protein, however a general rule is found: protein sample should be more than 98% pure, more than 95% homogeneous and more than 95% stable when stored unconcentrated at 4°C for two week.

Over-expression

Most of the proteins have to be expressed in other host systems (e.g., *Escherichia coli*, *Pichia pastoris* yeast, Sf9 or High five insect cells, human embryonic kidney HEK293, etc; as shown below) that theirs natural source because they are low abundance there; with exceptions of mammalian and bacterial rhodopsins, aquaporins, respiratory complexes, ATPases, photosynthetic complexes, reaction centers and light-harvesting proteins. One common trouble to use another system is than lipid composition between the natural source and the host are different, making it difficult to properly insert the membrane protein and generally producing unfolded protein. Other issue is that expression of membrane proteins can be toxic to the host strains, limiting the quantities necessary for crystallography. In addition some proteins need specific requirements (e.g., post-translation modifications) that can't be possible to realize in the host cell. Thus, the choice is not easy; but a study of expression of membrane proteins in different systems concluded that the best host is the one most closely related in evolution to the source of the target membrane protein. Understand the biosynthesis machinery in the host cell is also important because this can help to identify the bottlenecks in expression. For example expression of mammalian membrane proteins in *E. coli* requires low-copy-number plasmid, low temperatures during induction, etc; to reduce toxicity and misfolded protein.

Expression system	Advantages	Disadvantages
Escherichia coli	Good system for prokaryotic membrane protein production.	Often not suitable for eukaryotic membrane proteins. No glycosylation and limited post-translational modifications.
Yeast	Can perform some post-translational modification. Several are used to membrane proteins.	*S. cerevisiae* not produce high cell densities and hyperglycosylation can occur. Different lipids compared with mammalian cells.

Insect cells	Growth conditions less complex than mammalian cells. Relatively high expression levels. Glycosylation.	More costly and complex *than E. coli* or yeast. Different lipids compared with mammalian cells.
Mammalian cells	Authentic mammalian protein is produced. HEK293SGnTI$^-$ can be grow in suspension	High cost and low quantity.
Cell free	Allows expression of toxic or degraded proteins. May incorporated label and non-natural amino acids.	High cost. Insertion in the membrane or detergent micelle is not fully developed.

Some systems used in membrane protein production.

Cell disruption and membrane preparation

This step has to be adapted to the host cell because it can damage the membranes. After production, pellet cell are suspended in a lysis buffer usually containing a protease inhibitor cocktail (reduce possible protein degradation) and DNase (reduce viscosity), cell are then disrupted using a method shown below. Membranes are isolated by a set of centrifugations. However, some times is possible to add the detergent directly to the cell lysate, with no prior isolation of membranes.

Technique	Principle	Advantages (+) / Disadvantages (-)
Liquid shear pressure (e.g., French press)	Rapid pressure drop by transferring the sample from a chamber at high pressure through an orifice into a chamber at low pressure.	+ Fast and efficient, also for large volumes. - Causes heating of the sample (cooling is required).
Ultrasonication	Cells disrupted by high frequency sound	+ Simple. - Causes heating of the sample and is difficult to control by cooling. - Proteins may be destroyed by shearing. - Noisy. - Not for large volumes.
Glass bead milling	Agitation of the cells with fine glass beads	+ Useful for cells that are more difficult to disrupt (e.g., yeast). - Somewhat slow and noisy.
Osmotic shock	Change from high to low osmotic medium	+ Simple, inexpensive. - Only useful for disruption of cells with less robust walls (e.g., animal cells).

Repeated freezing and thawing	Cells disrupted by repeated formation of ice crystals; usually combined with enzymatic lysis	+ Simple, inexpensive. + Yields large membrane fragments. - Slow. - May damage sensitive proteins and dissociate membrane protein complexes. - Low yield.
Enzymatic lysis	Often used in combination with other techniques; lysozyme is commonly used to break cell walls of bacteria	+ Gentle. + Yields large membrane fragments. - Slow. - Low yield.

Some techniques to disrupt cell.

Extraction with detergents

Any crystallographic study required a sample in solution. Thus, membrane proteins need to be extracted from their natural environment, the lipid bilayer membrane, to an aqueous environment by the use of amphiphilic molecules such as detergents. During the solubilization the detergent, at a concentration above the CMC (critical micelle concentration), disintegrate the lipid bilayer while incorporate lipids and proteins on its micelles. Generally, an ultracentrifugation follow after this step to remove the unsolubilized material. If the detergent is not good enough to mimic the lateral pressure and charge distribution of the biological membrane, the protein will be inactive or denatured.

The structure of a detergent is a polar (hydrophilic) head group and a nonpolar (hydrophobic) tail. According to the polar head they can be classified in nonionic, ionic (anionic or cationic) and zwitterionic detergents, shown below. While ionic detergents disrupt protein-lipid interaction but not protein-protein interactions, is the contrary to the ionic and zwitterionic detergents.

Class	Advantages (+) / Disadvantages (-)	Name	Abbr.	gr/mole	CMC, mM	1%, mM
Nonionic	+ Generally mild and non-denaturing. +Widely used. - May give low solubilization yields.	n-Dodecyl-β-D-Maltopyranoside	DDM	511	0,12 0,2M Na	20
		Polyethylene glycol tert-octylphenyl ether, X=10	Triton X-100	647	0,01 50 mM Na	16
Ionic	+ Can be extremely efficient in solubilization. - Often denaturing. - Interfere with ion exchange separations.	Sodium dodecylsulfate	SDS	289	2-3 50 mM Na	35

Zwitterionic	+ Often used in membrane protein crystallization. - More denaturing than nonionic detergents.	Lauryldimethylamine N-oxide	LDAO	229	1-3 50 mM Na	44
		Fos choline 12	FC12	352	1,5 -	29
		Fos choline 16	FC16	408	0,013 -	29
		3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate	CHAPS	615	6-10 50 mM Na	16

Classification of detergents.

Abbr: abbreviation; gr/mole: molecular weight; CMC: critical micelle concentration in that condition; 1%, mM: the molar concentration corresponding to 1% of detergent.

Note: the CMC may vary with pH, temperature and ionic strength.

In the laboratory, anionic calix[4]arene-based detergents (C_4C_n, n = 1–12) were designed to extract and stabilize membrane proteins in their native form [302]. This is possible due to their capacity to generate a salt-bridges network between their trianionic hydrophilic head and the bulk of basic residues which are abundant at the cytosol-membrane interface. This new type of interaction is the basis of the innovative concept.

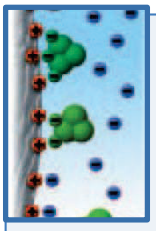
During the following steps (purification and crystallization) the detergent may be change for another one, for example a longer chain detergent may cover more membrane protein during the extraction, whereas the shorter chain may form a more compact protein-detergent complex, and so amenable to crystallization.

Purification and concentration

As soluble proteins, multiple chromatographic steps are carried out to obtain purified protein-detergent complex (shown below). Typically, the first step in a purification protocol is the immobilized metal ion chromatography (IMAC), follow by ion exchange chromatography (IEX) or size-exclusion chromatography (SEC). However, it should be note that not all the matrices are compatible with the presence of detergents (e.g., hydrophobic interaction chromatography, HIC). In general the last steep is the SEC column because this can allow to check the homogeneity of the protein, however others techniques can be used to check the quality of the protein, such as: SDS-PAGE (yield and purity), mass spectrometry or light scattering or SEC (homogeneity), native PAGE (oligomerization), etc.

IMAC

- Reversible interaction between a protein (e.g., histidine tag) and a specific metal attached to a chromatographic matrix (e.g., Nickel).
- In the case of Nickel column, the elution is done with imidazole.
- Sample volume is not limited.



IEX

- Separation according to charge differences.
- A resin having positive groups is known as the "anion exchange", while the contrary is known as "cation exchange".
- Proteins will take off by increasing the ionic strength of the elution buffer.
- Sample volume not limited.

SEC

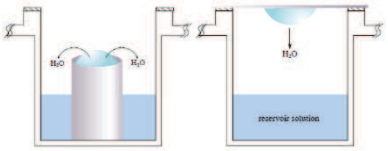
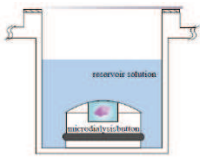
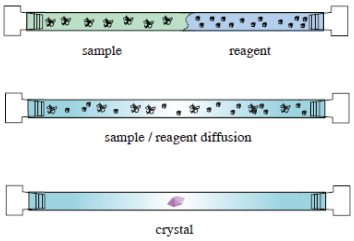
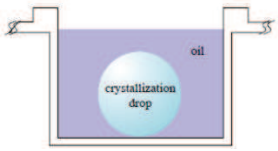
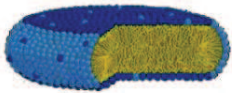
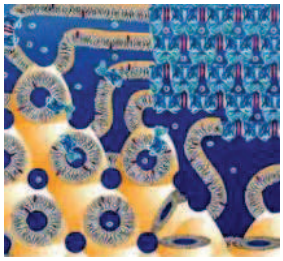
- Remove aggregates and other impurities according to the size while simultaneously perform buffer exchange.
- Ideal to final purification step.
- Limited sample volume and low rate range.

Some principles in chromatographic purification.

In order to achieve the supersaturated solutions necessary for successful crystallization experiments, after purification the protein is concentrated several times. But the detergent, and lipids or glycerol are also concentrated, which have bad consequences because too much detergent can denature the protein or impede the protein-protein contacts essential for the crystallization. Contrary, too low detergent may make insoluble protein and precipitate it. Is therefore preferable to use systems with concentration cutoffs higher than the size of the micelles but obviously less than the size of the protein-detergent complex. Another alternative is adjusted the concentration of the detergent after the concentration of the protein, using Bio-Beads (hydrophobic beads) which are capable of adsorb the detergent.

Crystallization

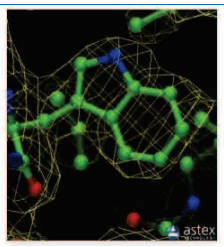
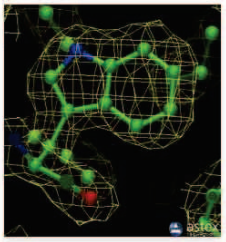
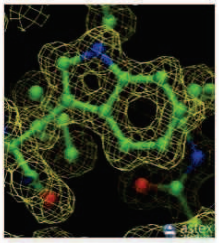
A variety of methodologies can be used to bring a protein solution into a supersaturation state and thus increase the protein-protein contacts necessary for the formation of crystals. They can be divided in two major groups: the *in surfo* and the bilayer methods (shown below). While the first use surfactants to produce mixed micelles that incorporate the protein and detergent in crystallization the second use bimolecular lipids. The common point is the multitude of parameters (pH, temperature, additives as salts and lipids, precipitating agents as polyethylene glycol PEG...), so the determination of crystallization conditions of a protein can be a daunting task.

	Advantages (+) / Disadvantages (-)	Name	Principle	Illustration
in surfo	+ Easy to handle. + Widely used. + Requires small amount of sample. - Potentially destabilizing environment for the detergent micelles.	Vapor diffusion (Sitting & Hanging drop)	The water from the gout is transfer to the reservoir leading to the supersaturation state.	
		Dialysis	Semi-permeable membrane allow water and small molecules to pass but not the macromolecules.	
		Batch	Simple is mixed with concentrated precipitant in a closed vessel to produce a final supersaturated concentration.	
		Micro- batch under oil	The drop is placed under paraffin oil which allows water leave the drop while protein remain.	
Bilayer	+ use lipid bilayer - based (avoiding the hosting medium) + easy to pipette - Recently introduced.	Bicelle	The protein-detergent complex is mixed with the bicelle stock solution (disks formed by detergent and a long-chain lipid) prior to setting up crystallization trials similar to sitting & hanging drop method.	
		Vesicle	The detergent is washed away and spontaneous vesiculation occurs upon subsequent incubation with a second detergent. The crystallization is done by vapor diffusion method.	
	+ use lipidic mesophase (avoiding the hosting medium). - Relative expensive. - Extremely viscous and sticky - Difficult to handle.	*In meso* (cubic and sponge phase)	The mix water-monoacylglycerol form a 3D network in which membrane proteins can freely diffuse in the lipids. By adding a precipitant crystals could be obtained.	

Some principles in crystallization.

After crystallization, energetic X-rays pass through crystals which are deflected by the protein molecule's electron clouds and saved by a photon detector. Since crystals have repeating symmetrical units, the deflected X-rays cause constructive interference with one another construing a pattern of dots, called diffraction pattern by a photon detector. If the protein diffract well, that means from all angles, the diffraction pattern will provide a template of the electron densities within the protein at a determinate resolution (as explain below). In order to determine the three-dimensional structure of the protein, the density map need to be fit with amino acids. Molecular replacement (use the coordinates of a similar protein) or multiple anomalous dispersion (use heavy atoms) are often used to solve the phase problem and thus resolve the structure.

Accuracy and precision

Resolution (Å)	Meaning	Example: residue Trp147 of *Bacilus subtilis* ferrochelatase
>4.0	Individual coordinates meaningless	
3.0 - 4.0	Fold possibly correct, but errors are very likely. Many side chains placed with wrong rotamer.	
2.5 - 3.0	Fold likely correct except that some surface loops might be mismodelled. Several long, thin side chains (lys, glu, gln, etc.) and small side chains (ser, val, thr, etc.) likely to have wrong rotamers.	 PDB code 2H1W, 2.6 Å
2.0 - 2.5	As 2.5 - 3.0, but number of side chains in wrong rotamer is considerably less. Many small errors can normally be detected. Fold normally correct and number of errors in surface loops is small. Water molecules and small ligands become visible.	 PDB code 2AC4, 2.1 Å
1.5 - 2.0	Few residues have wrong rotamer. Many small errors can normally be detected. Folds are extremely rarely incorrect, even in surface loops.	
0.5 - 1.5	In general, structures have almost no errors at this resolution. Rotamer libraries and geometry studies are made from these structures	PDB code 2H1V, 1.2 Å

Meaning of the resolution of protein structures.

According to the websites: http://en.wikipedia.org/wiki/Resolution_(electron_density) and http://www.proteinstructures.com/Experimental/Experimental/electron-density.html

The mesh represents the electron density into the model (ball and sticks representation). As shown in figures the electron density map for the side chain of Trp147 get better moving toward high resolution, the hole in the center of the aromatic ring is visible only at 1.2 Å resolution.

Posters and oral communications

- Seminaries "Bases Moléculaires et Structurales des Systèmes Infectieux" (BMSSI) spring 2014 (February 2014, Lyon, France). Oral presentation: Conformational flexibility of the multidrug resistance P-glycoprotein. Martinez L.

- French-Belgian annual ABC meeting 2013 edition (Octobre 2013, Lyon, France). Communication: Deciphering the polyspecificity of human multidrug resistance p-glycoprotein. Martinez L.

- Gordon conferences, Multi-Drug Efflux Systems (Mars 2013, Ventura, CA). Poster presentation: Binding sites for Hoechst 33342 and QZ59-SSS overlap in human ABCB1. Martinez, L.; Arnaud, O. ; Henin, E. ; Tao, H. ; Chaptal, V. ; Tod, M. ; Di Pietro, A. ; Zhang, Q.; Chang, G. ; and Falson, P.

- 4th FEBS Special Meetingon ABC Proteins, FEBS « Federation of European Biochemical Societies » (March 2012, Innsbruck, Austria). Poster presentation: P-glycoprotein inhibition mechanism of QZ59. Martinez L., Arnaud O., Henin E., Tao H., Chaptal V., Andrieu T., Dussurgey S., Tod M., Di Pietro A., Zhang Q., Chang G., and Falson P.

- Scientific Day ”ARC 1 - Health” (September 2012, Isle d’Abeau, France). Oral presentation: Structural and functional study of human ABC transporters involved in the efflux of anti-retroviral agents. Martinez L.

- Science Day of Cluster 10 “Infectiologie” (January 2012, Lyon, France). Poster presentation: Novel compounds inhibiting human P-glycoprotein and the VIH-1 protease. Martinez L., Hilton B.,Wolkowicz R., Ettouati L., Paris J., Le Borgne M., Terreux R., Andrieu T., Dussurgey S., Di Pietro A ., and Falson P.

Novel compounds inhibiting human P-glycoprotein and the HIV-1 protease

Lorena MARTINEZ[1*], Brett J. HILTON[2], Roland WOLKOWICZ[2], Laurent ETTOUATI[3], Joëlle PARIS[3], Marc le BORGNE[3], Raphaël TERREUX[4], Thibaud ANDRIEU[5], Sébastien DUSSURGEY[5], Attilio Di PIETRO[1] and Pierre FALSON[1*]

[1] *Drug Resistance Mechanism and Modulation Laboratory, BMSSI-IBCP UMR 5086 CNRS/Université Lyon 1, 7 passage du Vercors 69367 Lyon, France*
[2] *Department of Biology, San Diego State University, 5500 Campanile Drive NLS325b San Diego, CA 92182 San Diego, USA*
[3] *Laboratoire de Chimie Thérapeutique, Université Lyon I, Faculté de Pharmacie 8, avenue Rockefeller - 69373 LYON cedex 08*
[4] *Laboratoire de bioinformatique et RMN structurale, BMSSI-IBCP UMR 5086 CNRS/Université Lyon 1, 7 passage du Vercors 69367 Lyon, France*
[5] *UMS3444, Tour Inserm CERVI, 21 Avenue Tony Garnier, 69365, Lyon cedex 07, France*
lmartine@ibcp.fr ; pfalson@ibcp.fr

BACKGROUND

AIDS

- Caused by the human immunodeficiency virus (HIV)
- Not any cure nor preventive treatment yet
- The suppression of HIV replication is achieved through the combination of at least three antiretroviral drugs.

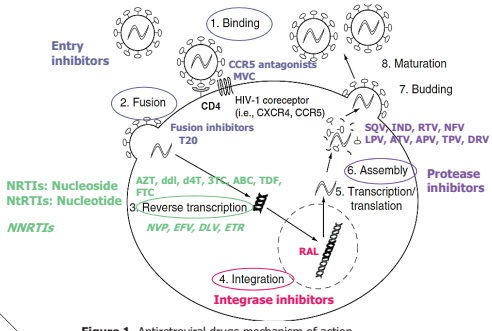


Figure 1. Antiretroviral drugs mechanism of action.

- Therapy fails due to drug toxicity, viral resistances and complex pharmacokinetic properties like drug transport by efflux pumps.

P-glycoprotein (P-gp)[1]:

- Member of the ATP-binding cassette (ABC) transporter family who use the energy of ATP hydrolysis to transport a large variety of drugs across the plasma membrane.

- *"may influence antiretroviral therapy in the target cell level (e.g., lymphocytes), at the bioavailability level (intestine/liver) and at sanctuary sites (brain, fetus, testis)" (Weiss 2010)*[2]

Inhibitors	Target	Pgp IC_{50}, (µM)	
LPV, TPV, NFV, RTV	Protease	5	- Transported by P-gp - Not good P-gp inhibitors
SQV, ATV, DRV		13-38	
APV		*?*	
DLV	Reverse transcriptase	?	
EFV			

AIM OF THE STUDY: Develop new molecules inhibiting both P-gp and HIV-1 protease activity (P-gp inhibition increasing bioavailability).

Approach

Protease inhibitors

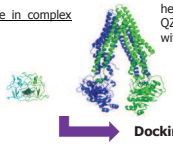
	Resolution, A	PDB ID
APV	1.02	3NU3
SQV	1.05	3DIX
DRV	1.46	3PWM
TPV	1.80	2O4P
ATV	1.80	3EKY
NFV	1.97	3EKX
RTV	2.20	2B60
IDV	2.20	2B7Z
LPV	2.80	1MUI

Table 1. HIV-1 protease in complex with inhibitors

Inhibitors of P-gp

	Resolution, A	PDB ID
QZ59(RRR)	4.4	3g60
QZ59(SSS)	4.35	3g61
CT[3,4]		
Progesterone		
Elacridar (GF120918)		
Tariquidar (XR9576)		
Zosuquidar (LY335979)		
Verapamil		

Table 2. Inhibitors of P-gp. Two cyclic hexapeptide inhibitors QZ59-RRR and QZ59-SSS were crystallized in complex with P-gp in 2009[5].



Docking → mixed inhibitor

1. OPTIMIZATION OF DOCKING PARAMETERS

Choice of receptor

As differents mutations of HIV-1 protease produce a large resistance to the protease inhibitors as chosen a protease without mutation.

Alignment of the sequence of wild-type HIV-1 protease (UniProt) with different proteases from the PDB showed that the protease with the code pdb 1mui has no mutation.

Dataset

9 protease inhibitors, taken from the pdb
CT, inhibitors of P-gp, were created in Sybyl
QZ59s, taken from the pdb
Other P-gp inhibitors were created in Sybyl

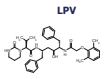
Computational details

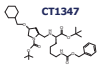
« Sybyl molecular-modeling » (SYBYL 1.3. Tripos Inc)
Surflex-Dock (Tripos) with Cscore

2. DOCKING ANALYSIS

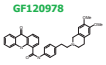
Cscore

Score >10
LPV, CT1316, CT1361, CT1326, CT1327, **CT1347**, CT1338, CT1344, CT1336, CT1342, CT1301, CT1337, CT1302, NFV, CT1346, CT1345, SQV, CT1339

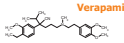




Score 7.5-10
CT1357, TPV, CT1343, CT1329, CT1345, IDV, CT1348, DRV, RTV, CT1340, CT1354, APV, CT1341, CT1300, Zosuquidar, GF120918, CT1356, ATV, CT1355



Score 5-7.5
CT1333, Verapamil, Tariquidar,CT1351, **CT1353**, Progesterone



Score 0-5
QZ59(RRR), QZ59(SSS),

Figure 2. Rank of scored. Four groups were identified, the blue ones is composed of molecules with the best affinity for the HIV-1 protease and red ones represent the worse. Interestingly many of P-gp inhibitors (CT) are in the best range.

Correlation between Score et $Ki_{experimental}$

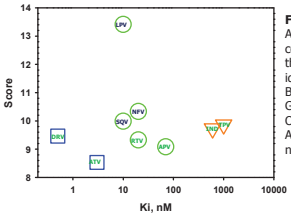


Figure 3. Score vs Ki. According to the constant of inhibition, three groups were identified.
Blue = Ki < 5 nM,
Green = 5 <Ki< 100 nM
Orange = Ki > 600 nM.
A clear correlation was not found.

AuPosSOM[6] « Automatic analysis of Poses using SOM »

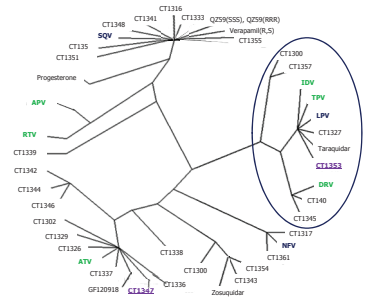


Figure 4. AuPosSOMS. The classification according to their similarity of contacts during docking showed that a few group of CT inhibitors could potentially inhibit the HIV-1 protease

3. TEST *IN CELLULO*[7]

Cells : SupT1 (the human T-cell lymphoblastic lymphoma cell line) transfected with Gal4, PRm/Gal4 or PR/Gal4

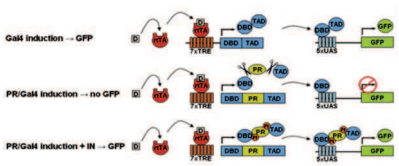


Figure 5. Representation of the System.

The Gal4 protein is a transcription factor
N-terminal DNA-binding domain (DBD: amino-acids 1–147)
C-terminal trans- activation domain (TAD: amino-acids 768–881).

D= doxycycline (bins to rtTA)
rtTA = reverse tetracycline transactivator (rtTA-Dox binds the TRE «Tet-Responsive Element»)
PR = wield type HIV-1 protease.
PRm =inactive version of PR (D25A)

In the absence of Dox no detectable GFP.
Inactive inhibitor = GFP (-)
Active inhibitor = GFP (+)

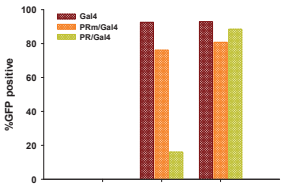


Figure 6. Quantification of eGFP expression by flow cytometry.
In the absence of Dox no detectable eGFP expression was observed. Induction of eGFP expression is observed with 1mg/mL in Gal4 and PRm/Gal4 cell but not in PR/Gal4. However, when PR/Gal4 cells were treated with 10mM IDV, a large induction of eGFP expression was observed.

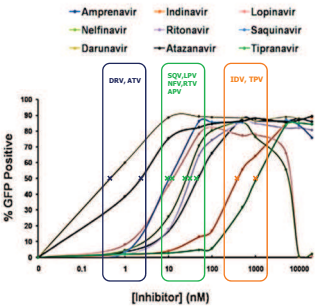


Figure 7. Activity of different protease inhibitors approved by the FDA. According to the constant of inhibition three groups were identified.
The 50% relative fluorescent intensity (IC50) for this inhibitors were agreed to the literature.

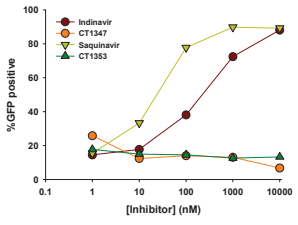


Figure 8. Inhibition of the predicted inhibitors.
In order to confirm the results of docking two compounds were tested, the CT1347 and the CT1353, both of them didn't show inhibition of HIV-1 protease.

CONCLUSION

According to the results of docking, several CT inhibitors may potentially inhibit the HIV-1 protease. These results were not confirmed *ex vivo* for the two tested so far, which let us hypothesize that those compounds, quite hydrophobic, stay in the plasma membrane, and thus cannot reach the HIV -1 PR.

PERSPECTIVES

One short term perspective is to modify the CT to increase their solublity.
In the order to develop more efficient inhibitors, it is of importance to understand the mechanism by which P-gp transports this antiretroviral drugs. P-gp crystallization with selected inhibitors will be carried out according to the protocols of Pr. Chang of SCRIPPS Research Institute thanks to the Explora' 2011 fellowship from *Région Rhône-Alpes*.

REFERENCES

1. Juliano RL and Ling V. (1976), **A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants**, *Biochim Biophys Acta, 455, 152-62*
2. Weiss J and Haefeli WE. (2010), **Impact of ATP-binding cassette transporters on human immunodeficiency virus therapy**, *Int Rev Cell Mol Biol, 280, 219-79*
3. Sharom et al., (1999), **Interaction of the P-glycoprotein multidrug transporter (MDR1) with high affinity peptide chemosensitizers in isolated membranes, reconstituted systems, and intact cells**, *Biochem Pharmacol, 58, 571-86.*
4. Arnaud O et al., (2010), **Modulateurs peptidiques et peptidomimétiques non compétitifs spécifiques de la glycoprotéine P**, *J Med Chem, 53, 6720-9*
5. Aller et al., (2009), **Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding**, *Science, 323(5922):1718-22*
6. Bouvier et al., (2010), **Automatic clustering of docking poses in virtual screening process using self-organizing map**, *Bioinformatics, 26(1):53-60.*
7. Hilton B and Wolkowicz R. (2010), **An Assay to Monitor HIV-1 Protease Activity for the Identification of Novel Inhibitors in T-Cells**, *PLoS One, 5, e10940.*

This work was supported by the Centre National de la Recherche Scientifique (CNRS) and University of Lyon 1 (UMR5086), and Ministère de la Recherche et de l'Enseignement Supérieur (EA3741, Lyon 1). It was funded by the Agence Nationale de la Recherche (ANR-06-BLAN_0420, ANR-06-PCVI-0019-01, ANR piribio09_444706), the Association pour la Recherche sur le Cancer (ARC) and the Ligue Nationale Contre le Cancer (Labellisation 2012). Financial support of Lorena MARTINEZ (PhD student) was from the Cluster d'infectiologie from the **Région Rhône-Alpes**.

Binding sites for Hoechst 33342 and QZ59-*SSS* overlap in human ABCB1

Lorena MARTINEZ[1*], Ophélie ARNAUD[1], Emilie HENIN[2], Houchao TAO[3], Vincent CHAPTAL[1], Michel TOD[2], Attilio Di PIETRO[1], Qinghai ZHANG[3], Geoffrey CHANG[4] and Pierre FALSON[1*]

[1] *Drug Resistance Mechanism and Modulation Laboratory, BMSSI UMR 5086 CNRS/Université Lyon 1, 7 passage du Vercors 69367 Lyon, France*
[2] *EMR3738, Ciblage Thérapeutique en Oncologie, Faculté de Médecine Lyon-sud Charles Mérieux, BP12, 69921 OULLINS Cedex,FRANCE*
[3] *Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, CB105, La Jolla, CA 92037, USA*
[4] *Skaggs School of Pharmacy and Pharmaceutical Sciences, UCSD, 9500 Gilman Dr, San Diego, California 92093, USA*

**lorena.martinez@ibcp.fr ; p.falson@ibcp.fr*

Introduction

Background

- ATP-binding cassette (ABC) transporters are involved in the multidrug resistance (MDR) phenotype.
- Different identified drug-binding sites: H and R sites, respectively binding Hoechst 33342 and anthracyclins (rhodamine 123, daunorubicin, etc)
- The x-ray crystal structure of ABCB1 a (mouse) was recently solved with two cyclic hexapeptide inhibitors, QZ59-RRR and QZ59-SSS, bound in the internal cavity (Figure 1A).

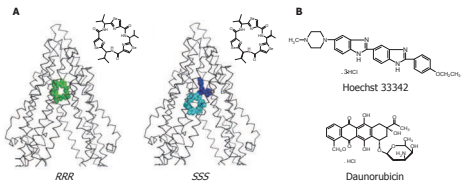


Figure 1. Mouse ABCB1 + QZ59 compounds. (*Aller et al., 2009*).

Aim of the study

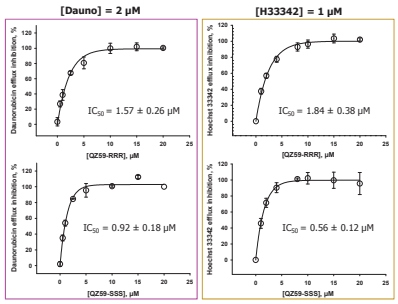
To determine whether QZ59 inhibitors & commonly transport drugs of human ABCB1 share the same binding sites.

Approach

- Kinetic analysis of the inhibition of ABCB1-mediated transport by QZ59 compounds using flow cytometry.
- H site study with Hoechst 33342 (H33342), (Fig 1B).
- R site study with daunorubicin (dauno), (Fig 1B).

QZ59 inhibits human ABCB1

Half maximal inhibition concentration



In general:
RRR IC_{50} =1.5-2.2 µM
SSS IC_{50} =0.5-1.2 µM

Figure 2. IC_{50} of the inhibition of daunorubicin and H33342 efflux by QZ59 compounds.

QZ59 compounds show concentration-dependent inhibition of human ABCB1- mediated transport

H-site

RRR

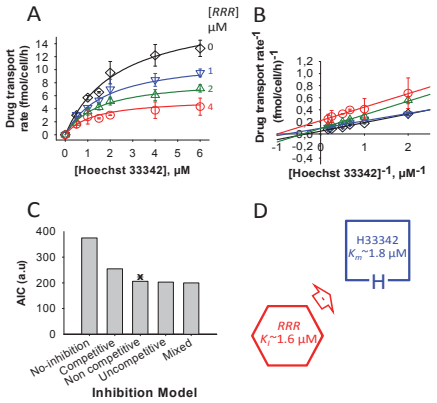


Figure 3. Hoechst 33342 efflux inhibition by QZ59-RRR.
A- Direct plot V=f(S) of drugs efflux. **B-** Lineweaver–Burk double-reciprocal plot *v-1 = f([S]-1)*.
C. Histogram of the Akaike's Information Criterion (AIC) estimated for each inhibition model.
D- Scheme of drug-substrate and inhibitor binding sites interaction.

***RRR* doesn't bind to the H site.**

SSS

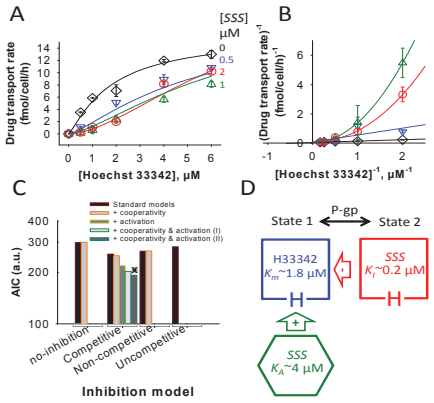


Figure 4. Hoechst 33342 efflux inhibition by QZ59-SSS.
A- Direct plot V=f(S) of drugs efflux. **B-** Lineweaver–Burk double-reciprocal plot *v-1 = f([S]-1)*.
C. Histogram of the Akaike's Information Criterion (AIC) estimated for each inhibition model.
D- Scheme of drug-substrate and inhibitor binding sites interaction.

***SSS* binds to the H site**

Predicted H drug-transport sites of ABCB1

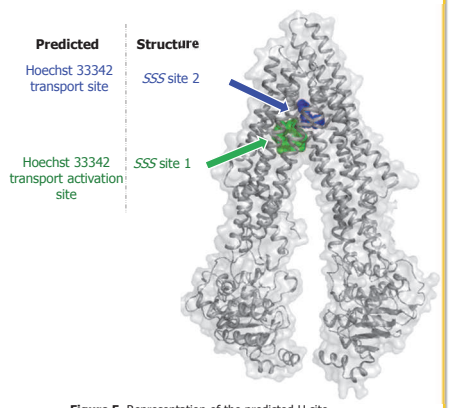


Figure 5. Representation of the predicted H site.
Two QZ59-SSS interact with the pump, one at an inhibition site with a K_i of 0.2 µM and the other at an activation site, with a K_A of 4 µM.

R-site

RRR

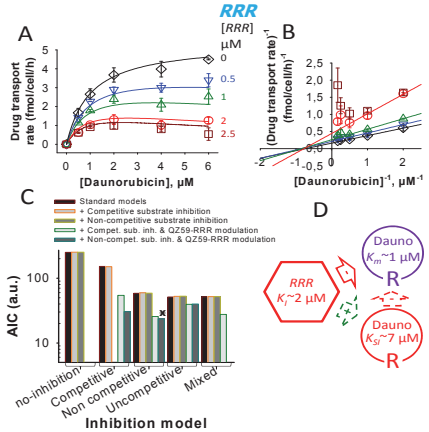


Figure 6. Daunorubicin efflux inhibition by QZ59-RRR.
A- Direct plot V=f(S) of drugs efflux. **B-** Lineweaver–Burk double-reciprocal plot *v-1 = f([S]-1)*.
C. Histogram of the Akaike's Information Criterion (AIC) estimated for each inhibition model.
D- Scheme of drug-substrate and inhibitor binding sites interaction.

***RRR* doesn't bind to the R site.**

SSS

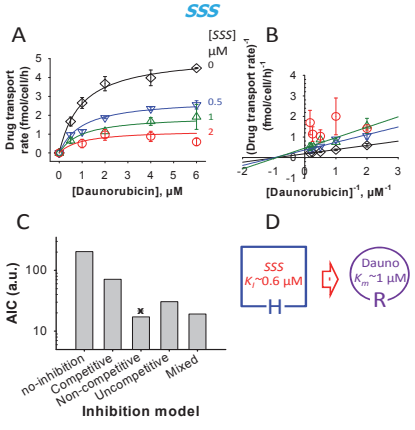


Figure 7. Daunorubicin efflux inhibition by QZ59-SSS.
A- Direct plot V=f(S) of drugs efflux. **B-** Lineweaver–Burk double-reciprocal plot *v-1 = f([S]-1)*.
C. Histogram of the Akaike's Information Criterion (AIC) estimated for each inhibition model.
D- Scheme of drug-substrate and inhibitor binding sites interaction.

***SSS* doesn't bind to the R site**

Equations for the kinetic analysis of the inhibition of ABCB1 by QZ59 R/S

Substrate	Inhibitor	Model	Equation
H33342	*RRR*	Non-competitive inhibition	$v=\frac{V_m/\left(1+\frac{I}{K_I}\right)\cdot S}{K_m+S}$
	SSS	Competitive inhibition + Cooperativity and QZ59S activation	$v=\frac{V_m\cdot\left(1+\frac{I}{K_I}\right)\cdot S^{h\left(1+\frac{I}{K_A}\right)}}{K_m^{h\left(1+\frac{I}{K_A}\right)}\cdot\frac{\left(1+\frac{I}{K_I}\right)}{\left(1+\frac{I}{K_A}\right)}+S^{h\left(1+\frac{I}{K_A}\right)}}$
Dauno	*RRR*	Non-competitive inhibition + Non-competitive Substrate Inhibition and QZ59R modulation	$v=\frac{\frac{V_m}{\left(1+\frac{S}{K_{SI}}+\frac{I}{K_I}\right)}\cdot\left(1+\frac{I}{K_I}\right)\cdot S}{K_m+S}$
	SSS	Non-competitive inhibition	$v=\frac{V_m/\left(1+\frac{I}{K_I}\right)\cdot S}{K_m+S}$

Table 1. QZ59R/S inhibition of Hoechst 33342 and daunorubicin transports.
Equations retained for each model of QZ59R/S inhibition towards Hoechst 33342 and daunorubicin transports.

CONCLUSION

1°/ QZ59-RRR doesn't share the H or R transport sites.

2°/ QZ59-SSS competitively inhibits the Hoechst 33342 efflux at low concentration (< 1 µM) but also activates it at high concentration (>1 µM). This implies the existence of two QZR9-SSS binding sites.

3°/ By predicted the H transport site, our results offer the possibility to target more precisely this site to develop modulators. The discovery of the activation site may also offer another way to modulate the pump.

4°/ Our results show that Hoechst 33342 is transported 3–4 times faster than daunorubicin (V_m=16 fmol/cell/h vs V_m=4.5 fmol/cell/h) with a two-fold lower affinity (K_m=1.8 µM vs K_m=1 µM).

REFERENCES

- Juliano RL and Ling V. (1976), A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants, *Biochim Biophys Acta, 455, 152-62.*
- Aller SG, Yu J, Ward A, Weng Y, Chittaboina S, et al. (2009) Structure of P-Glycoprotein Reveals a Molecular Basis for Poly-Specific Drug Binding. Science 323: 1718-1722.
- Gottesman MM, Ambudkar SV, Xia D (2009) Structure of a multidrug transporter. Nat Biotech 27: 546-547.
- Jardetzky O (1966) Simple allosteric model for membrane pumps. Nature 211: 969-970.
- Dawson RJP, Locher KP (2007) Structure of the multidrug ABC transporter Sav1866 from Staphylococcus aureus in complex with AMP-PNP. FEBS Letters 581: 935-938.
- Hohl M, Briand C, Grütter MG, Seeger MA (2012) Crystal structure of a heterodimeric ABC transporter in its inward-facing conformation. Nat Struct Mol Biol 19: 395-402.
- Jin MS, Oldham ML, Zhang Q, Chen J (2012) Crystal structure of the multidrug transporter P-glycoprotein from Caenorhabditis elegans. Nature advance online publication.

This work was supported by the Centre National de la Recherche Scientifique (CNRS) and University of Lyon 1 (UMR5086), and Ministère de la Recherche et de l'Enseignement Supérieur (EA3741, Lyon 1). It was funded by the Agence Nationale de la Recherche (ANR-06-BLAN_0420, ANR-06-PCVI-0019-01, ANR p1ibio09_444706), the Association pour la Recherche sur le Cancer (ARC) and the Ligue Nationale Contre le Cancer (Labellisation 2012-2014). Financial support of Lorena MARTINEZ (PhD student) was from the Cluster d'infectiologie from the Région Rhône-Alpes.

P-glycoprotein inhibition mechanism of QZ59 stereoisomers

Lorena MARTINEZ[1*], Ophélie ARNAUD[1], Houchao TAO[2], Sarah FERNANDEZ[1], Laura LEPINE[1], Attilio Di PIETRO[1], Thibaud ANDRIEU[3], Sébastien DUSSURGEY[3], Qinghai ZHANG[2], Geoffrey CHANG[2] and Pierre FALSON[1*]

[1] *Drug Resistance Mechanism and Modulation Laboratory, BMSSI-IBCP UMR 5086 CNRS/Université Lyon 1, 7 passage du Vercors 69367 Lyon, France*
[2] *Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, CB105, La Jolla, CA 92037, USA*
[3] *UMS3444, Tour Inserm CERVI, 21 Avenue Tony Garnier, 69365, Lyon cedex 07, France*

lmartine@ibcp.fr ; p.falson@ibcp.fr

Background

P-glycoprotein (P-gp)

- Member of the ATP-binding cassette (ABC) transporter family involved in the multidrug resistance (MDR) phenotype.

- Different identified drug-binding sites: H, R and P sites, respectively binding Hoechst 33342, rhodamine 123 and anthracyclins, and prazosin.

- 2009: resolution of the mouse P-gp x-ray structure with two cyclic hexapeptide inhibitors QZ59-RRR and QZ59-SSS. (Figure 1).

Aim of the study

Determine if the QZ59 inhibitors share the sites occupied by transported drugs (P-R and H sites) and study their mechanism of inhibition.

Approach

The QZ59 inhibitors impact on the efflux of drugs binding to the H, R and P sites was measured by flow cytometry.

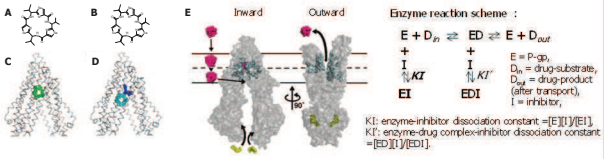


Figure 1. Resolution of Pgp X-ray crystal structures. Chemical structures of **(A)** QZ59-RRR and **(B)** QZ59-SSS. **(C)** Location of one QZ59-RRR (green spheres- pdb 3G60, 4.4 Å) and **(D)** two QZ59-SSS (blue and cyan spheres- pdb 3G61, 4.35 Å) molecules in the P-gp internal cavity. **(E)** Models of drug transport in P-gp, substrate (magenta), ATP (yellow) and the residues in the drug binding site (cyan spheres). *Reprinted from Aller et al., 2009.*

Characteristics

Cytotoxicity

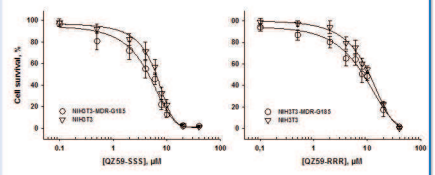


Figure 2. QZ59 cytotoxicity. The QZ59-RRR and QZ59-SSS are cytotoxic with IC_{50} of about 10 and 6.5 µM respectively. No protection is observed with the expression of human P-gp, or even a slight higher sensitivity suggesting that **both QZ59 are not transported by the pump.**

Half maximal inhibition concentration

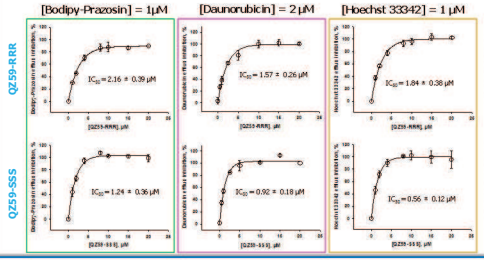


Figure 3. IC_{50} of QZ59 inhibitors on bodipy-prazosin, daunorubicin and Hoescht 33342 efflux.
In general:
QZ59-RRR IC_{50} between 1.5-2.2 µM
QZ59-SSS IC_{50} between 0.5-1.2 µM

Thanks to these results, the concentrations of QZ59 inhibitors in the next experiences were fixed to 0, 0.5, 1, and 2 µM, excepted the QZ59-RRR in the H site which was 0, 1, 2 and 4 µM

P-site

QZ59-RRR

QZ59-SSS

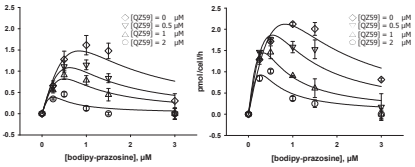


Figure 4. Direct plot V=f(S) of bodipy-prazosin efflux with QZ59-RRR and QZ59-SSS . **The Bodipy-prazosin is transported up to 1 µM and inhibits it own efflux above. The QZ59-RRR and QZ59-SSS behave equally non competitive (at least qualitatively).**

Representation of P drug-transport site of P-gp.

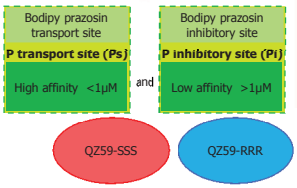


Figure 5. Representation of the P site. The QZ59 inhibitors are in a different site of the Prazosin because they are non competitive inhibitors for prazosin.

R-site

QZ59-RRR

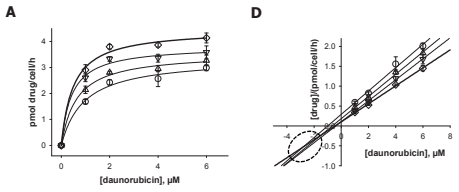


Figure 6. QZ59-RRR daunorubicin efflux inhibition.
A- Direct plot V=f(S) of drugs efflux. **B-** Values of VM and KM deduced from each fit and Plotted in **C. D-** Hannes-Wolf plots. **E-** Enzyme reaction scheme.

QZ59-RRR inhibited the efflux of daunorubicin in a mixed non competitive inhibition way.

QZ59-SSS

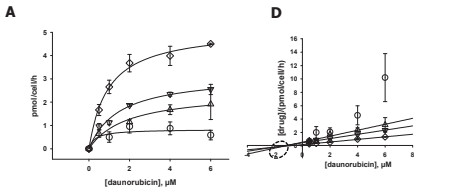


Figure 7. QZ59-SSS daunorubicin efflux inhibition.
A- Direct plot V=f(S) of drugs efflux. **B-** Values of VM and KM deduced from each fit and Plotted in **C. D-** Hannes-Wolf plots. **E-** Enzyme reaction scheme.

QZ59-SSS inhibited the efflux of daunorubicin in a non competitive way with negative cooperativity from 2 µM of QZ59-SSS.

Representation of P and R drug-transport site of P-gp

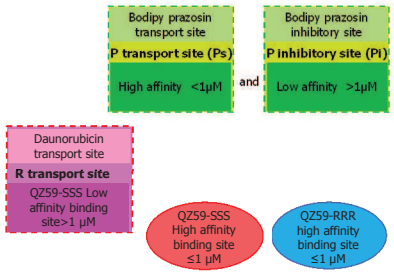


Figure 8. Representation of the P and R site.
Below 1 µM, both QZ59 inhibited the efflux of daunorubicine in a non competitive way, suggesting that they do not bind to the R-sites. In addition the QZ59-SSS became competitive for daunorubicine efflux above 2 µM.

H-site

QZ59-RRR

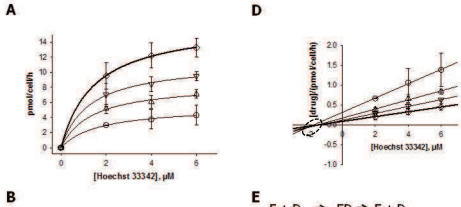


Figure 9. QZ59-RRR Hoechst 33342 efflux inhibition.
A- Direct plot V=f(S) of drugs efflux. B- Values of VM and KM deduced from each fit and Plotted in C. D- Hannes-Wolf plots. E- Enzyme reaction scheme.

The non competitive inhibition displayed by QZ59-RRR towards Hoechst 33342 suggests that the inhibitor does not bind preferentially to the apo or enzyme:substrate complex.

QZ59-SSS

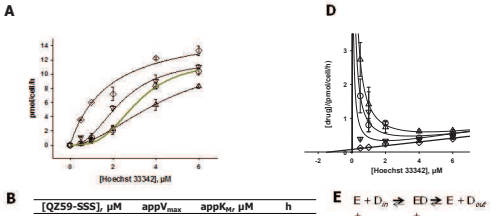


Figure 10. QZ59-SSS Hoechst 33342 efflux inhibition.
A- Direct plot V=f(S) of drugs efflux. B- Values of VM and KM deduced from each fit and Plotted in C. D- Hannes-Wolf plots. E- Enzyme reaction scheme.

The QZ59-SSS display a allosteric inhibition towards Hoechst 33342. The inhibitor bind preferentially to the enzyme:substrate complex and above 2 µM becomes competitive with the H33342 transport site.

Representation of P-R and H drug-transport site of P-gp

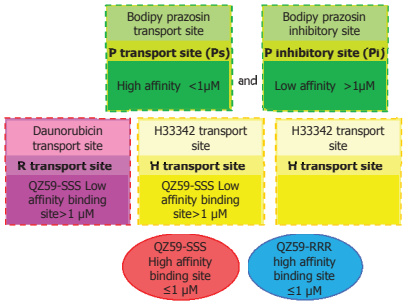


Figure 11. Representation of the P, R and H site.
The number of hill=2, suggesting the existence of two sites for the Hoechst 33342. Below 1 µM both QZ59 inhibited the efflux of Hoechst 33342 in a non competitive way, suggesting that they do not bind to the H-sites. However the QZ59-SSS became competitive for H33342 efflux above 2 µM.

CONCLUSION

This study suggests that P-gp, at steady state, displays :

1° / a high affinity transport for prazosin (Ps) and an inhibitory site of lower affinity (Pi), this results are in accordance with Dey 2001. Both QZ59 inhibited the efflux of prazosin in a non competitive way, suggesting that they do not bind to the P-sites.

2° / a allosteric inhibition of QZ59-SSS toward the Hoechst 33342 transport with a number of hill=2, suggesting the existence of two sites for the Hoechst 33342.

3° / two low affinity binding site for the QZ59-SSS in the R-site and H-site.

Dey et al., (1997) described the presence of at least two nonidentical substrate interaction sites in P-gp. This study confirms the existence of 2 P and H sites. The existence of one another R-site is probable.

REFERENCES

- Juliano RL and Ling V. (1976), **A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants**, *Biochim Biophys Acta, 455, 152-62.*
- Aller et al., (2009), **Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding**, *Science, 323(5922):1718-22.*
- Sharom et al., (1999), **Interaction of the P-glycoprotein multidrug transporter (MDR1) with high affinity peptide chemosensitizers in isolated membranes, reconstituted systems, and intact cells**, *Biochem Pharmacol, 58, 571-86.*
- Sharom FJ. (2011), **The P-glycoprotein multidrug transporter**, Essays Biochem, *50(1):161-78.*
- Shapiro AB and Ling V. (1997), **Positively cooperative sites for drug transport by P-glycoprotein with distinct drug specificities**, *Eur. J. Biochem: 250, 30-137.*
- Shapiro AB and Ling V. (1999), **Stimulation of P-glycoprotein-mediated drug transport by prazosin and progesterone Evidence for a third drug-binding site**, *Eur. J. Biochem; 259, 841-50.*
- Dey et al., (1997).**Evidence for two nonidentical drug-interaction sites in the human P-glycoprotein**, *Proc Natl Acad Sci U S A; 94(20):10594-9.*
- Maky et al., (2006), **Allosteric Modulation Bypasses the Requirement for ATP Hydrolysis in Regenerating Low Affinity Transition State Conformation of Human P-glycoprotein**, *J Biol Chem, 281(16):10769-77.*

This work was supported by the Centre National de la Recherche Scientifique (CNRS) and University of Lyon 1 (UMR5086), and Ministère de la Recherche et de l'Enseignement Supérieur (EA3741, Lyon 1). It was funded by the Agence Nationale de la Recherche (ANR-06-BLAN_0420, ANR-06-PCVI-0019-01, ANR piribio09_444706), the Association pour la Recherche sur le Cancer (ARC) and the Ligue Nationale Contre le Cancer (Labellisation 2012-2014). Financial support of Lorena MARTINEZ (PhD student) was from the Cluster d'infectiologie from the Région Rhône-Alpes.

Title: Structural and functional study of efflux pumps involved in drug resistance

Abstract

L'efficacité des chimiothérapies est limitée par la surexpression de pompes d'efflux adressées à la membrane plasmique des cellules cibles. En effet, celles-ci réduisent le taux intracellulaire des médicaments anticancéreux, antiviraux, antifongiques et antibactériens. La P-gp/ABCB1 est la plus impliquée dans ce phénomène, suivie de MRP1/ABCC1 et de BCRP/ABCG2. Une approche pour surmonter ce phénomène est de développer des médicaments qui ne soient pas expulsés par ces pompes. Dans ce contexte, nous avons développé une nouvelle classe d'inhibiteurs de la protéase du VIH-1 qui ne sont ni transportés par P-gp ni par BCRP. Ils sont ainsi des candidats intéressants aux trithérapies contre le SIDA. Un point clé pour comprendre comment ces transporteurs font traverser les médicaments à travers la membrane est d'identifier des nouvelles structures. Ainsi, nous avons résolu trois structures de P-gp de souris. Une d'entre-elles est complexée à un nano-anticorps lié au premier NBD («nucleotide-binding domain»), qui fige la P-gp dans une nouvelle conformation ouverte vers l'intérieur. Pour finir, nous avons localisé deux sites de liaison de P-gp en caractérisant les modes d'inhibition de deux inhibiteurs précédemment cocristallises avec la pompe. Ceci permette mieux comprendre le mécanisme de translocation et offre la possibilité de cibler plus précisément ces sites pour développer des modulateurs de cette pompe.

Keywords: ABC transporters, P-glycoprotein, screening, X-ray structures, binding sites.

Titre: Étude structurale et fonctionnelle des pompes à efflux impliqués dans la résistance aux médicaments

Résumé

L'efficacité des chimiothérapies est limitée par la surexpression de pompes d'efflux adressées à la membrane plasmique des cellules cibles pour réduire le taux intracellulaire des médicaments anticancéreux, antiviraux, antifongiques et antibactériens. La P-gp/ABCB1 est la plus impliquée dans ce phénomène, suivie de MRP1/ABCC1 et BCRP/ABCG2. Une approche pour surmonter ce phénomène est de développer des médicaments qui ne soient pas expulsés par ces pompes. Dans ce contexte, nous avons développé une nouvelle classe d'inhibiteurs de la protéase du VIH-1 qui ne sont ni transportés par P-gp ni par BCRP. Ils sont ainsi des candidats intéressants aux trithérapies contre le SIDA. Un point clé pour comprendre comment ces transporteurs font traverser les médicaments à travers la membrane est d'identifier des nouvelles structures. Ainsi, nous avons résolu trois structures de P-gp de souris. Une d'entre-elles est complexée à un nano-anticorps lié au premier NBD («nucleotide-binding domain»), qui fige la P-gp dans une nouvelle conformation ouverte vers l'intérieur. Pour finir, nous avons localisé deux sites de liaison de P-gp en caractérisant les modes d'inhibition de deux inhibiteurs précédemment co-cristallises avec la pompe. Ceci permette mieux comprendre le mécanisme de translocation et offre la possibilité de cibler plus précisément ces sites pour développer des modulateurs de cette pompe.

Mots clés : Transporteurs ABC, P-glycoprotéine, criblage, 3D-structures, sites de transport.

Adresse: Mécanisme et Modulation de la Résistance aux Médicaments (DRM2), UMR 5086/CNRS/ Université Claude Bernard LYON 1. 7, passage du Vercors- 69367 LYON cedex 07- France.