Identification and functional characterization of an ABC transporter of Haemonchus contortus, the P-glycoprotein 13

Marion A. David

Université Fédérale

Toulouse Midi-Pyrénées

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)

Cotutelle internationale avec l'Université McGill

Présentée et soutenue par :
Marion DAVID
le vendredi 14 octobre 2016

Titre :

Identification and functional characterization of an ABC transporter of Haemonchus contortus, the P-glycoprotein 13

École doctorale et discipline ou spécialité :
ED SEVAB : Pathologie, Toxicologie, Génétique et Nutrition

Unité de recherche :
INRA ToxAlim UMR 1331, Equipe Transporteurs Membranaires et Résistance

Directeur/trice(s) de Thèse :

Dr. Anne LESPINE et Pr. Roger K. PRICHARD

Jury :
Dr. Jean-Michel JAULT - Rapporteur
Pr. Elias GEORGES - Rapporteur
Pr. Alexis VALENTIN - Président du jury
Dr. François ANDRE - Examinateur
Dr. Anne LESPINE - Directrice de thèse
Pr. Roger K. PRICHARD - Directeur de thèse

Acknowledgements

First, I would like to thank the various fundings that made this work possible: the Natural Sciences and Engineering Research Council of Canada, the FRQNT Centre for Host-Parasite Interactions, Quebec, and EMIDA ERA-NET project CARES n° 11-EMID-003-02. I also thank the Center for host-pathogens interactions (CHPI) that allowed me to extend my financial support for six months, the Université Paul Sabatier (UPS) in Toulouse, and the Frontenac Program of Quebec government that covered most of my travel fees between France and Canada.

I would like to acknowledge Pr. Elias Georges, Dr. Jean-Michel Jault and Pr. Alexis Valentin for accepting to evaluate my PhD work.

Many thanks go to my supervisors Dr. Anne Lespine and Dr. Roger K. Prichard for giving me the opportunity to do this PhD in co-direction between each of their laboratories in France and Canada. This was a very rewarding experience in many ways, both scientific and personal. Thanks for believing in me since the beginning and for always encouraging me to persist while Hco-Pgp-13 was giving us hard times to learn to know it. Thanks for all your advice and for watching over my work so that it followed the right track during all these years. Thanks for helping me with the extension of my PhD, thus allowing it to finish in the best possible conditions.

I also would like to acknowledge my adivsory committee, Pr. Robin Beech and Dr. Cecile Menez, who gave much advice and helped many times with results analysis during the three years of my PhD.

I would like to warmly thank the scientists who have highly contributed to the success of my PhD:

- *Thanks to Dr. François André for welcoming me to its laboratory at the CEA and teaching me with bio/chemo-informatics. Thanks for teaching me patience and the art of perfectionism that work together. Thanks for the extensive help with results analysis, and for revising each part of the manuscripts as precisely as possible.*
- *Thanks to Dr. Stéphane Orlowski for the incredible amount of work on ABC transporter homologies studies. Thanks also for the help with writing and correcting articles. Thank you for all our scientific talks, about my thesis project or other subjects of interest, and for being someone as passionate about research as you are. I hope I learned everything I could from your philosophy of science.*
- *Thanks to Dr. Chantal Lebrun, for being a skilled biochemist with a passion for transmitting her knowledge. Thanks for having always been happy to help with experiments, to advice with strategies and troubleshooting, and to participate in my manuscript correction. Thanks for your contagious love of science and your great humanity.*
- *Thanks to the students who critically helped me during these 4 years. Thanks to Shaima Hashem for helping with my learning of in silico docking and homology modelling. Thanks for your availability and sense of humor. Thanks to Thomas Duguet for helping in many ways with immuno-fluorescence assays. Thanks for always being happy to collaborate and for being the best moaner in the Parasitology building after me, thus contributing to the French reputation. Thanks to Clément Frainay for his availability and the high amount of technical help provided for docking calculations run in Toulouse. Thanks for being a ninja of bioinformatics.*

I also want to thank all other members of my both teams:

- *in Montreal: Thanks to Kathy for teaching me most of techniques in McGill,and for various and extensive technical help during all my PhD, even from the other side of the ocean. Thanks to Hua for her precious advice with the nested PCR and for her sweetness. Thanks to Catherine for many scientific advice, long talks and valuable support. Thanks also to all the people who made this lab such a nice cosmopolitan place: Ludmel, Virginie, Aissatou, Nour, Pablo, Shoaib, Durai and Phillip.*
- *à ToxAlim, INRA : Merci à Jean-François, pour sa présence et pour tout le chocolat salvateur de mon estomac. Merci à Elise, pour « la tristitude » et sa joyeuse folie contagieuse. Merci à Léa, pour ne trouver aucune blague drôle mais pour avoir bien rit depuis son arrivée, et pour m'avoir aidé précieusement sur la dernière ligne droite ! Merci à Mélanie et Alice pour leur aide au tout début de ma thèse et à mon retour de Montréal. Enfin, merci aux étudiantes de licence pour avoir participé à mon projet avec joie : Christina, Elsa et Mélanie.*

A special thanks to the students I met in neighbor teams and who helped me in many ways:

- *at the Institute of Parasitology, McGill: Adeline, thanks for coming back to the airport to take me back to your home on the 13th of November 2015, and for everything else. Krittika, so many thanks for running with me by -13°C, it was so "cool"! Thanks to all of the students of your team, for making me discover both Canadian and US American customs. Thanks to the Fauchers: Nilmini and Hana for making me travel to Sri Lanka and Tunisia just by talking with you. Finally, thanks to the best roommates ever: Sarah and Oscar, for being the kindest Quebecer and cat I could ever meet.*
- *à ToxAlim, INRA : Merci à Alix, Eli et Sarah d'avoir fait leur thèse la porte en face ou au bout du couloir, et d'avoir toujours été prêtes à s'entraider de diverses manières, et en particulier les derniers jours d'écriture, vous avez été géniales. Merci pour tous les précieux souvenirs accumulés pendant ces dernières années, j'espère que de nombreuses occasions d'en rajouter suivront. Merci à Isaura pour ton sourire permanent et les très bons moments partagés et merci aux étudiantes de l'équipe 05 qui y ont également participé.*

Merci à tous les autres, famille ou amis, loin de la science, qui m'ont soutenue de façon toute aussi importante :

- *Au Canada: Merci à Marie-Jo et Dominik, Suzanne, Cynthia and Djé, pour être des exemples de sagesse, de force et de gaieté. Merci à Renata Heidersdorf pour ses merveilleux cours d'art qui m'ont aidé plus qu'il n'est imaginable dans mon projet scientifique.*

- *En France : Merci à mes parents pour avoir toujours accepté de me soutenir financièrement tout au long de ces quatre ans. Merci de m'avoir fait pousser à la campagne et d'avoir arrosé mes racines de votre amour pour les rendre plus solides que je n'aurais cru. Les fruits de mes années de travail sont aussi pour vous. Merci à ma grande famille, auprès de laquelle j'ai aimé me ressourcer. Merci aux précieux amis qui ont toujours été tout près quelquesoit la distance. Merci à Patty pour les trop rares moments partagés ces dernières années et à Léti pour n'avoir jamais hésité à m'accueillir à Paris dès qu'un séjour était nécessaire au CEA. Vous êtes les meilleures. Merci à Nicolas pour le soutien et les encouragements essentiels à ma fin de thèse, merci d'avoir été au bon endroit au bon moment ! Merci infiniment à Marie-Jo, pour les mots simples et percutants, qui reboostent en quelques phrases pour au moins des mois. Enfin, un immense merci à la Team AGL, qui s'agrandit toujours plus, merci à tous ces gens au grand cœur d'être entré dans ma vie ! Un merci spécial à Geneviève, sans qui l'aventure n'aurait jamais eu lieu, et qui veille comme une maman poule sur tous ses petits, et Valérie, inspirante de motivation et de bienveillance, une grande sœur inestimable. Enfin, merci à Ludo pour ton grand cœur fort si complémentaire à mon petit cœur sensible.*

A Grégory Lemarchal,
Pour m'avoir inspiré ma quête :

« Tenter, sans force et sans armure
D'atteindre l'inaccessible étoile
[…]
Je ne sais si je serai ce héros
Mais mon cœur serait tranquille »
La Quête – Jacques Brel

Table of contents

List of figures and tables of the literature review

Abbreviations

ABA = abamectin

ABC = ATP-binding cassette

ACD = actinomycin D

ATP = adenosine tri-phosphate

AH = anthelmintic

BLAST = Basic Local Alignment Search Tool

CAM = calcein-AM

CCH = colchicine

Cel = *Caenorhabditis elegans*

Ceg = *Cylicocylus elongatus*

Cgr = *Cricetulus griseus*

CLO = closantel

CLS = cholesterol

CSP = cyclosporin A

DNR = daunorubicin

DOR= doramectin

DPM = dipyridamole

DXR = doxorubicin

EMD = emodepside

EPR = eprinomectin

Hco = *Haemonchus contortus*

Hsa = *Homo sapiens*

HST = hoechst 33342

IVM = ivermectin

IVA = ivermectin-aglycone

KTC = ketoconazole

LEV = levamisole

MDR = multidrug resistance

ML = macrocyclic lactone(s)

Mmu = *Mus musculus*

MNP = monepantel

MOX = moxidectin

NBD = nucleotide binding domain

Oar = *Ovis aries*

PCT = paclitaxel

PDB = Protein Data Bank

Pgp = P-glycoprotein

PRG = progesterone

QMEAN = Qualitative Model Energy Analysis

RHO = rhodamine 123

RMSD = root mean square deviation

SEL = selamectin

SNP = single nucleotide polymorphism

TBZ = thiabendazole

TCZ = triclabendazole

TMD = transmembrane domain

VBL = vinblastine

VCR = vincristine

VLN = valinomycin

VRP = verapamil

VSP = valspodar

Contribution of authors

The design and execution of experiments in this thesis were carried out by the author under the supervision of Pr. Roger K. Prichard at the Institute of Parasitology, McGill University, Montreal, Canada and Dr. Anne Lespine at the Research Institute in food toxicology (Toxalim), INRA, Toulouse, France, except for the heterologous expression of Hco-pgp-13 in *Pichia pastoris* and several functional assays.

For the first and second manuscripts, as well as part of the third one, a great deal of help came from Dr. François André as he provided teaching of homology modelling and docking calculations. He also helped with interpretation of results and correction of article manuscripts. Dr. Stéphane Orlowski did all the hotspots alignments, and also helped extensively with the interpretation of results as well as with writing of the first article and correction of the thesis manuscript. Shaima Hashem and Clément Frainay helped with technical issues regarding the use of bioinformatics softwares and Unix, respectively. Dr. Fabien Jourdan provided a Unix system for running calculations on the Genotoul Platform. Nicolas Loiseau helped with *in silico* data interpretation.

For the third manuscript, Dr. Robin Beech (McGill University) helped with getting access to the sequences of *H. contortus*, teaching how to use bioinformatics tools and antibody design. Catherine Bourguinat and Hua Che gave thoughfull advices critical for the success of whole *Hco-Pgp-13* cDNA sequencing. Thomas Duguet provided a great help for *H. contortus* cryosectionning, setting up of immunofluorescence assays and use of confocal microscope. Kathy Keller provided the *H. contortus* worm sample and helped technically during DNA/RNA extraction and cloning. She also helped with the shipments between the two partner laboratories of a considerable amount of products and samples. Franck Talmont performed the transfection of *Pischia pastoris* with *Hco-pgp-13* and Dr. Chantal Lebrun performed *in vitro* experiments for characterization of its function. She also participated in result analysis, writing of the article and correction of the thesis manuscript.

Dr. Roger K. Prichard and Anne Lespine supervised the entire processes and provided valuable technical advice and help with correction of manuscripts needed for the smooth completion of work.

This work was supported by the Natural Sciences and Engineering Research Council of Canada, the FRQNT Centre for Host-Parasite Interactions, Quebec, and EMIDA ERA-NET project CARES n° 11-EMID-003-02.

Scientific valorization

Publications

Modeling multispecific drug recognition by *Caenorhabditis elegans* P-glycoprotein 1. David M., Orlowski S., Prichard R.K., Hashem S., André F., and Lespine A. In preparation.

In silico analysis of anthelmintics binding to *Caenorhabditis elegans* P-glycoprotein 1. David M., Orlowski S., Prichard R.K., Hashem S., André F., and Lespine A. 2016. Int. J. for Parasitol., Drug and drug resistance. In Press.

Characterization of *Haemonchus contortus* P-glycoprotein 13. David M., Lebrun C., Duguet T.B., Talmont F., Beech R., Orlowski S., André F., Lespine A. and Prichard R.K. In preparation.

Oral communications at international and national meetings

David M, Beech R., Prichard R.K. and Lespine A. Identification and functional characterization of *Haemonchus contortus* Pgp-13. 11th French-Belgian ABC meeting, October 3, 2014. Paris, France.

David M, Beech R., Prichard R.K. and Lespine A. Identification and functional characterization of *Haemonchus contortus* Pgp-13. "Journée des doctorants" (école doctorale SEVAB), Mars 2015. Toulouse, France.

David M, Orlowski S., Andre F., Prichard R.K. and Lespine A. *In silico* analysis of anthelmintics binding on *Caenorhabditis elegans* P-glycoprotein 1. "Symposium on Anthelmintics : Discovery to Resistance II" Feb 10, 2016, San Diego, USA.

General Introduction

ABC transporters belong to a large family of ATP hydrolyzing proteins involved in the transport of a wide variety of compounds in mammals. Among these transporters, the multi-drug resistance (MDR) ABC transporters are more specifically dedicated to the transport of xenotoxics and they have been clearly involved in resistance to drugs in all living kingdoms (Jones and George, 2005). In mammals, MDR ABC transporters such as P-glycoprotein (MDR1/ABCB1/Pgp) can efflux various, structurally unrelated drugs, and Pgp is responsible for resistance to chemotherapy in mammalian tumor cells (Eckford and Sharom, 2009; Leonard et al., 2003). Noticeably, mammalian Pgp also transports various AHs, including ML and IVM, which contributes to modulating its efficacy in the host (Lespine et al., 2007; Roulet et al., 2003; Schinkel et al., 1994).

Many ABC transporters are present in various kingdoms, including different infectious agents (Koenderink et al., 2010; Lage, 2003). In nematodes, many homologs of ABC transporters are expressed in each organism. *C. elegans* is the model organism for studying resistance of parasitic nematodes to AHs, as it is genetically very close to all roundworms and presents the advantage to harbor a free-living life-cycle. Plus, it was the first animal to have its genome totally sequenced. In total, 60 ABC transporters, among which 14 full Pgp homologs, were found to be expressed in different organs and at various stages of development of this nematode (Zhao et al., 2004). There is little information on their respective functions, except that ML can interact with some of these Pgps. Indeed, induction of expression of individual genes after selection under IVM pressure is associated with increased IVM resistance, and IVM susceptibility can be improved by using mammalian Pgp inhibitors on *C. elegans* (James and Davey, 2009; Lespine et al., 2012). As a more direct observation, the loss of each of the 14 Pgps has been shown to increase susceptibility of *C. elegans* to IVM or MOX to various degrees (Ardelli and Prichard, 2013; Janssen et al., 2013). Otherwise, Jin et al. (2012) established on purified Cel-Pgp-1 that some mammalian Pgp substrate drugs of various chemical structures could stimulate its ATPase activity, providing the most compelling evidence of a multidrug transport function for a nematode Pgp. However, AH drugs have not been tested on this experimental system. Finally, IVM or MOX-selected strains of *C. elegans* showed an increased susceptibility to the three ML IVM, MOX and EPR after exposure with the ABC transporter substrate verapamil (Menez et al., 2016).

In many parasitic nematodes, the precise number of Pgps is not well defined, excepted for *Haemonchus contortus,* whose genome has recently been sequenced and annotated (ftp:// ftp.sanger.ac.uk/pub/pathogens/Haemonchus/contortus) (Laing et al., 2011; Laing et al., 2013). In this species, 10 homologs of Pgps were identified, and a few of them have been localized and their function studied individually. Indeed, in heterologous recombinant systems overexpressing the protein Hco-Pgp-2, Hco-Pgp-16 or Hco-Pgp-9.1, ML were shown to inhibit Pgp-mediated drug transport (Godoy et al., 2015a, 2016; Godoy et al., 2015b). This was also shown for some Pgps of other parasitic ABC transporters : Cylicocylus elongatus Pgp-9 and Dirofilaria immitis Pgp-11 (Kaschny et al., 2015; Mani et al., 2016). Plus, the role of Pgp in IVM transport in parasites has been indicated by the capacity of well-described Pgp inhibitors to increase IVM sensitivity or reverse resistance to IVM in susceptible and resistant strains, respectively. Valspodar, verapamil and pluronic acid indeed increased the susceptibility to ivermectin of sensitive and resistant strains of *Teladorsagia circumcincta* and *H. contortus* (Bartley et al., 2009).

H. contortus is a parasite of critical importance in veterinary medicine. Because of the putative implication of Pgps in AH resistance, it is important to know more about the function of Pgp-like proteins in *H. contortus*. In order to gain insight into the characteristics of Pgps of parasitic nematodes, we focused our interest on *Hco-pgp-13*. Indeed, its closest ortholog in C. elegans, *Cel-pgp-13*, was found to be located in the amphids, which are chemo-sensitive neurons shown to play a role in ML sensitivity (Dent et al., 2000; Freeman et al., 2003; Menez et al., 2016; Urdaneta-Marquez et al., 2014). Although *Cel-pgp-6* is also expressed in this organ, no ortholog was found for this gene in *H. contortus*.

The gene and protein identity, as well as the function of Hco-Pgp-13 and its localization in the nematode were thus investigated.

We first tested the strength of an *in silico* docking strategy of ligands on an ABC transporter by studying the interaction with Cel-Pgp-1 of compounds previously described experimentally as Cel-Pgp-1 substrates. We thus aimed at shedding light on the binding properties of a multispecific substrate protein. Furthermore, our next goal was to describe for the first time the binding of AH drugs, including ML, to a nematode ABC transporter, Cel-Pgp-1.

To characterize Hco-Pgp-13, we then checked experimentally the cDNA sequence of *Hco-pgp-13* predicted by large scale sequencing (ftp://ftp.sanger.ac.uk/pub/pathogens/ Haemonchus/contortus) (Laing et al., 2013) starting from whole worm RNA. We predicted the

protein sequence and 3D structure of Hco-Pgp-13 and investigated its homology to Cel-Pgp-1 protein characterized as a multidrug transporter. We looked for the capacity of Hco-Pgp-13 to bind various substrates, including AH drugs. This was first performed *in silico* by docking of several compounds on 3D models of Hco-Pgp-13 built by homology with Cel-Pgp-1. The interaction of the same compounds was tested when possible *in vitro* by transport and ATPase activity measurements of Hco-Pgp-13 in two heterologous expression systems. We finally localized its expression in L3 and adult *H. contortus* to better characterize its importance in the living parasite.

Literature Review

Literature Review

I. *HAEMONCHUS CONTORTUS* AND *CAENORHABDITIS ELEGANS*

1. *Haemonchus contortus*, a parasitic gastro-intestinal nematode

According to the most recent phylogenetic classification (Blaxter and Koutsovoulos, 2015), *Haemonchus contortus* belongs to the rhabditomorpha family of nematodes of the clade V: Rhabditina, together with *Caenorhabditis elegans* (**Fig. 1**). This clade belongs to the sub-class C of nematodes: Chromadoria, where the clade III Spirurina and clade IV: Tylenchina are also found. Two other sub-classes exist in the phylum Nematoda: Sub-class I: Dorylaima and sub-class II: Enoplia. All these families contain free-living or parasitic, marine or terrestrial moulting animals, ranging in size from 0.2mm to over 6m (Blaxter and Denver, 2012). Although about 23000 species have been described in this phylum, the real species-level diversity is estimated around 1 million (Lambshead, 1993).

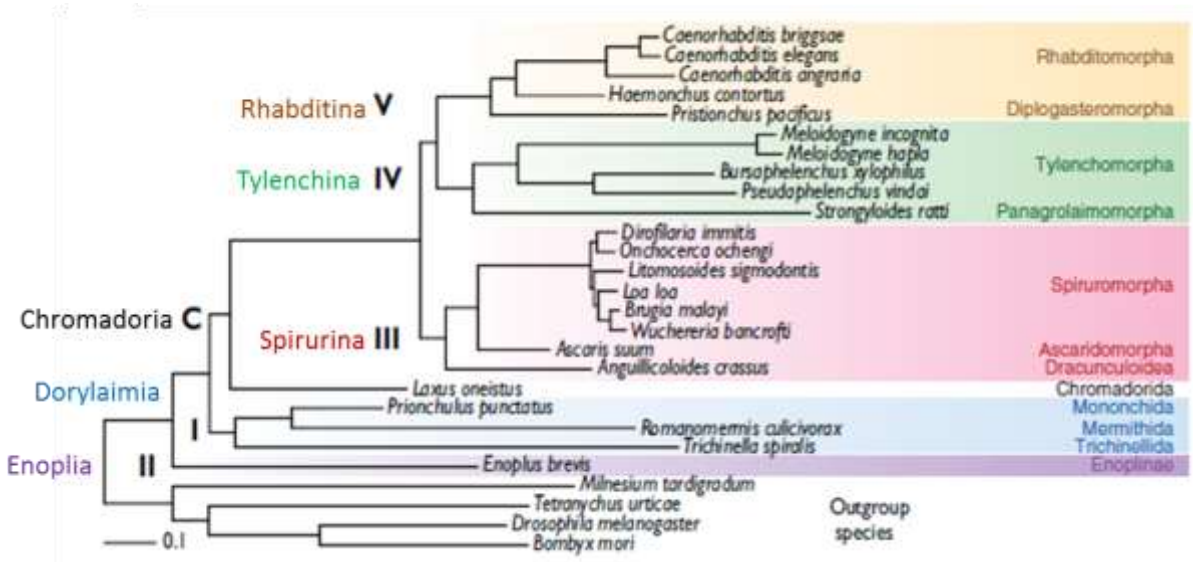


Figure 1. The phylogenetic structure of the Nematoda. A phylogeny derived from 181 protein coding genes from 23 nematode species, and four ecdysozoan taxa as outgroup. The alignment was subjected to analysis with PhyloBayes (Lartillot et al., 2009), and all nodes had posterior probability of 1·00. The major clades in Rhabditida are resolved, and Enoplia is recovered at the base of Nematoda. Clades I, II, C, III, IV and V were first defined in (Blaxter et al., 1998), and are indicated by their number. The figure is adapted from (Blaxter and Koutsovoulos, 2015).

a) Life cycle of Haemonchus contortus

The life cycle of *H. contortus* is direct, as it only requires one host **(Fig. 2A)**. It comprises two phases:

- The free phase starts when the eggs are laid by the female worm in the host gastro-intestinal tract and are subsequently eliminated in the feces (**Fig. 2; 1**). Eggs are approximately 45x80µm long (**Fig. 2B**) and once scattered in the pasture field, they develop to the first larval stage (L1) (**Fig. 2C**) in seven days with an optimal temperature of 27°C, while hatching takes longer at lower temperatures (Chermette, 1982) (**Fig. 2; 2**). L1 need a high moisture level and feed on bacteria to become L2 (**Fig. 2; 3**) (**Fig 2D**), which continue to feed on bacteria and molt to become filariform L3 larvae (**Fig. 2; 4**). The latter keeps the cuticle of the L2 stage, allowing it to survive between 6 to 12 months on the pasture, and is non-feeding (**Fig 2E**).
- The parasitic phase starts when infective L3 larvae are ingested by small ruminants when they graze (**Fig. 2; 5**). Larvae quickly migrate to the fourth stomach to become L4 within 48 hours and then rapidly reach the juvenile adult stage. However, if L3 were ingested just before winter period, L4 can stay several weeks in a dormant state, i.e. hypobiosis, and live within the epithelium of the abomasum before continuing the life cycle, in order to avoid harsh climate conditions for future larvae that cannot survive winter. Once at the adult stage, the nematodes are about 15-35 mm long. After mating, each female worm releases about 5000 - 7000 eggs per day and the cycle restarts (Coyne et al., 1991) (**Fig. 2; 1**).

b) Anatomy of *Haemonchus contortus* male and female adult

The general organization of the male and female adult is shown in **Fig. 3**, with males (A) being overall shorter (10-20mm long) than females (18-30mm long) (B). As generally in roundworms, a flexible but tough cuticle covers the body of adults. More specifically in *H. contortus*, the cuticle is transparent and shows longitudinal striations (not visible in **Fig. 3**). The whole bodies of worms are occupied by a tubular digestive system that opens on the anterior and posterior endings: at the tip of the pharynx, the mouth presents a unique dorsal lancet for perforating the abomasum tissue of the host to feed from its blood, and at the posterior end, the anus ends the long intestine. However, they are not composed of a circulatory system, i.e. neither a heart nor blood vessels. The females have two large gonads

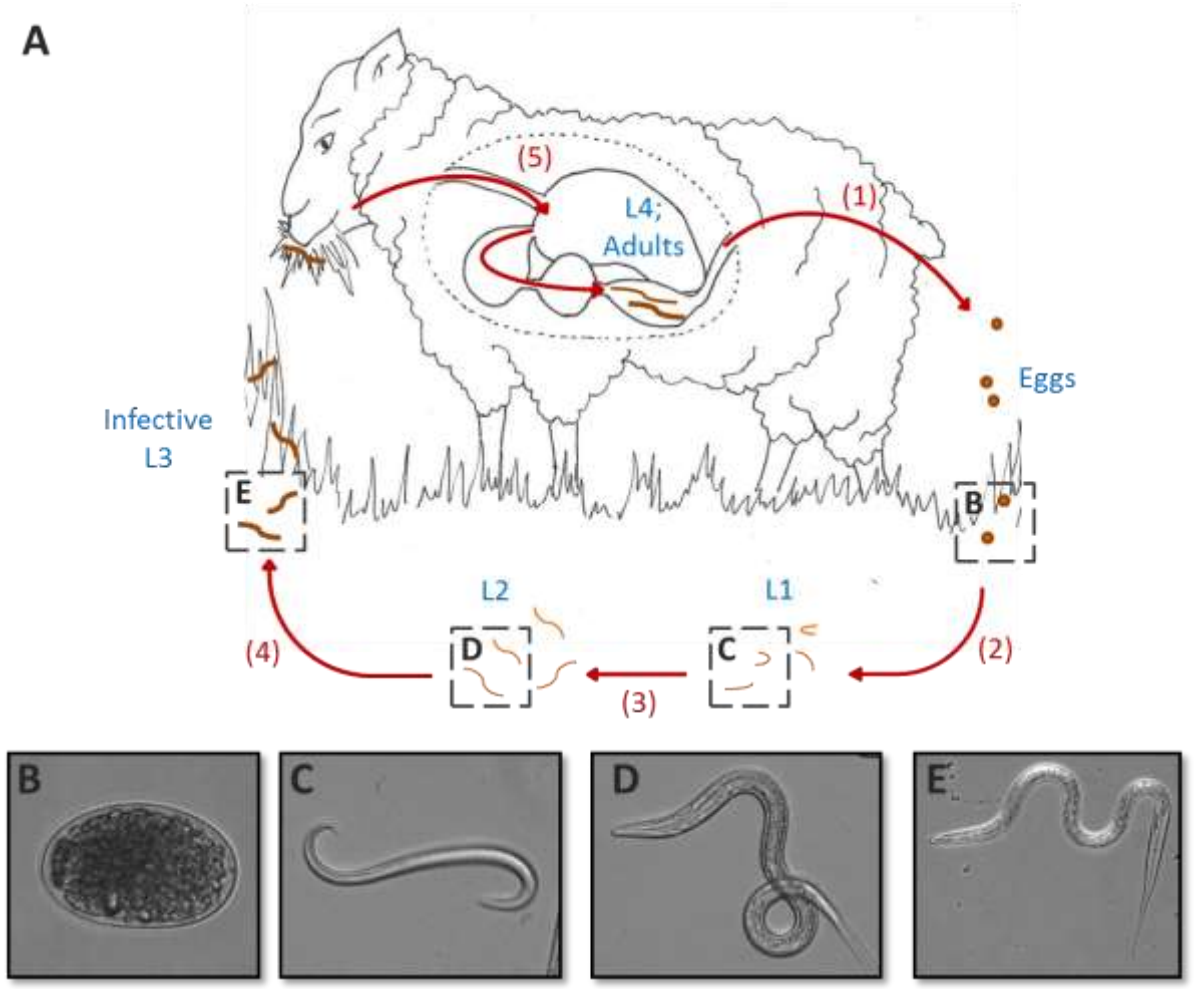


Figure 2. *Haemonchus contortus* life cycle. A. Different steps of the life cycle are represented as numbered dark red arrows: adults in the stomach of a host mate and females lay eggs that are released with faeces on the pasture (1); eggs hatch in L1 larvae (2); L1 develop to L2 (3); L2 become L3 larvae protected by a cuticle (4); a host ingests L3 that migrate to the abomasum and rapidly develop to L4 and adults (5). **B-E.** Zoom on various stages: eggs x200 (B), L1 x 100 (C) L2 x100 (D), L3 x100 (E).

starting with ovaries at around 1/3 of the intestine, and uteri that end in an opening called the vulva, which has a characteristic flap in *H. contortus*. In this parasite, female gonads characteristically wind around the intestine. Since the gut is filled with blood after a meal, it takes a reddish color that is easily visible around the white uteri, so that female worms look like small barber's poles. On the other hand, males have a single gonad that ends with a copulatory bursa presenting two barbed spicules, which allow its attachment to the female during copulation.

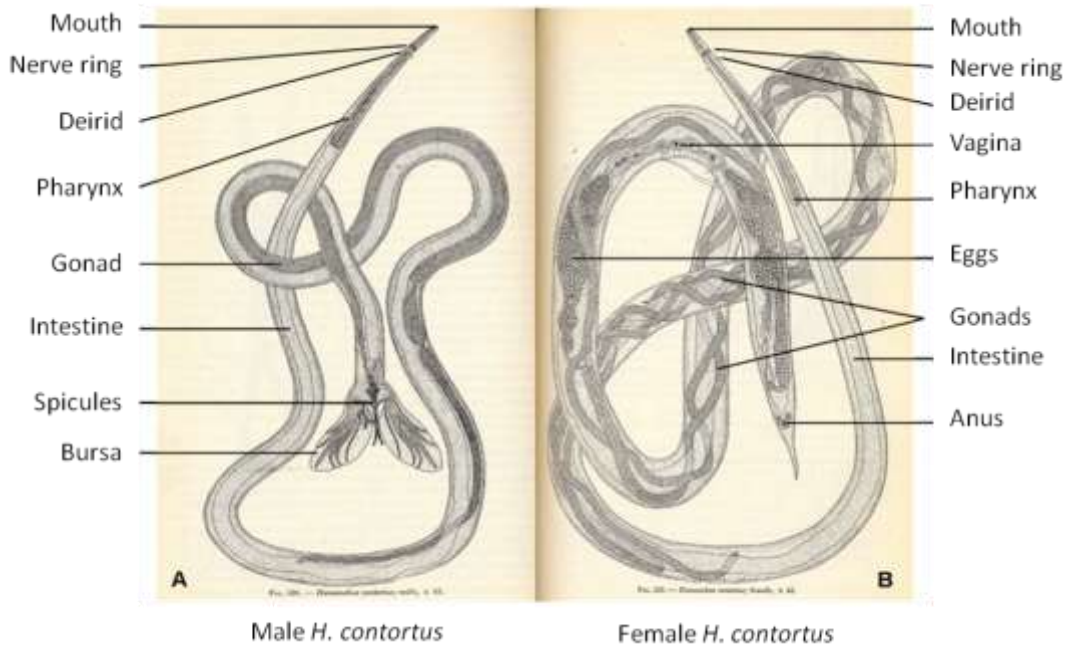


Figure 3. Anatomy of adult *Haemonchus contortus*. Schematic representation of a male adult. (A) and a female adult (B) from (Neveu-Lemaire, 1936).

The neuroanatomy of *H. contortus* is poorly described, except for amphids that have been studied in detail in order to better understand the role of these neurons in thermotaxis. Indeed, amphids are known in *C. elegans* to be thermo- and chemo-sensory neurons that are connected to the whole nervous system of the worm and reach the external environment on each side of the mouth (**Fig. 4**). In the L3 larval stage, this organ detects environmental signals such as light and warmth to find the best position to be ingested by its host (Ashton et al., 1999). Plus, once ingested by a host, the environmental conditions of *H. contortus* change greatly in the rumen, which stimulate development to the L4 stage. However, in some instances, development can be arrested in the L4 stage, rather than the L4 developing within a few days to the young adult stage. As these processes have been hypothesized to be controlled by stimuli detected by the amphids at the L3 stage, subcellular studies of their structure in various stages have been performed (Li et al., 2000a; Li et al., 2001; Li et al., 2000b).

Electron microscopy and three-dimensional reconstructions from electron micrographs allowed the identification of twelve amphidial neurons in L1 stage larvae of *H. contortus* (Li et al., 2000a). The same number of neurons was found in the amphids of infective L3, with slight differences in the pattern of some sensory cilia (Li et al., 2001). These neurons were named according to *C. elegans* nomenclature using three letters: A for amphids, followed by

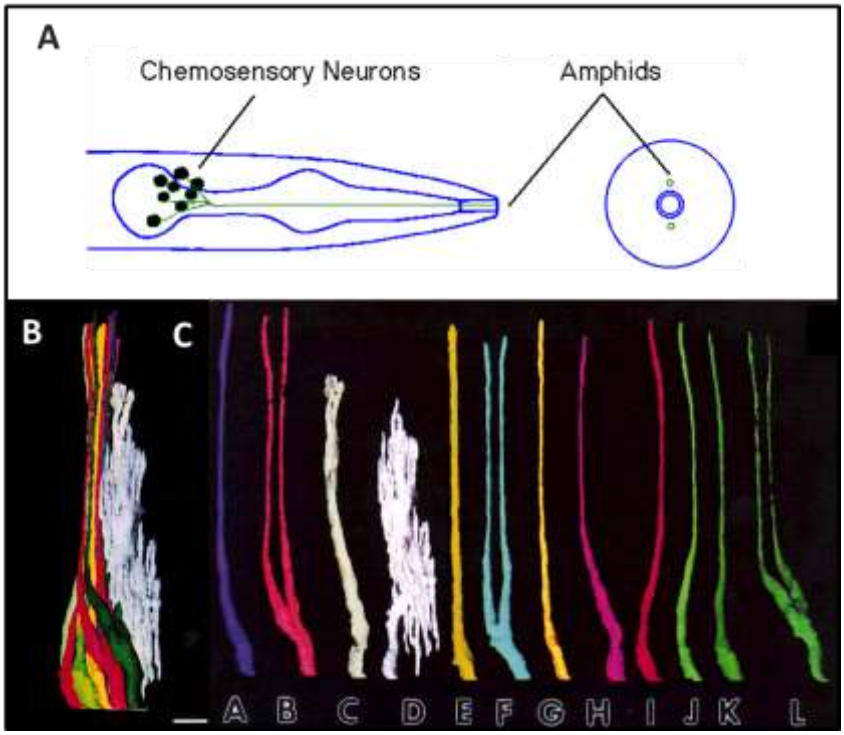


Figure 4. Amphidial neurons of *Haemonchus contortus*. A. Schematic representation of the location of amphids in the head of the worm (http://www.csi.uoregon.edu/ projects/celegans/talks/ nips1996poster.html). **B.** Stereo image of a reconstruction made with the SYNU surface-rendering program of the bundle of right-side amphidial dendrites before tight junction in longitudinal view. Anterior is up, dorsal is to the right, and the lateral side faces the viewer. At the base of the amphidial channel, the cilia extend posteriorly as dendrites and become thick and fan out. **C.** Gallery of the individual amphidial dendritic processes from the SYNU reconstruction shown in B. A = ASA, B = ADB, C = AWC, D = AFD, E = ASE, F = ADF, G = ASG, H = ASH, I = ASI, J = ASJ, K = ASK, L = ADL. Scale bar = 1 µm. Adapted from (Li et al., 2000a).

S or D for single or double dendritic processes in the channel, W for wing cell or F for finger cell (White et al., 1986). The third letter is assigned from A to L (Ward et al., 1975). As the arrangement of the cell bodies of the amphidial neurons in *H. contortus* was found to be similar to that of their counterparts in *C. elegans*, they were assigned letter names according to the positions of their apparent homologs in the free-living nematode: ASA, ADB, AWC, AFD, ASE, ADF, ASG, ASH, ASI, ASJ, ASK and ADL. Different cell bodies of neurons were ablated at the L1 stage and L3 stage and behavior was studied to understand the importance of these neurons in thermosensory control (Li et al., 2000b). This study showed that amphidial neurons AFD, and interneurons RIA, were the thermoreceptor and thermosensory integrative neurons. These neurons, as well as others composing the amphids, are critical in the pathogenicity of *H. contortus*.

c) Veterinary importance of the parasite

H. contortus was first identified as a tropical pathogen, but it has since spread worldwide (Jabbar et al., 2006) due to livestock movement and its high adaptive capacities. The L3 larvae are in fact covered by a protective sheath and the L3 stage can survive for many months provided they do not dessicate. In winter they can survive under snow, but freezing and thawing conditions damage the larvae. In addition, hypobiosis allows L4 to survive cold temperatures, within the host, during the hot dry summer season in tropical and Mediterranean climates, and the winter period in cool-temperate countries. Moreover, the vast mortality of L1 and L2 larvae is largely compensated by the mass production of eggs by females (Jacquiet et al., 1995). This is permitted by the life span of adult worms in the abomasum of their host for many months. However, infected animals can develop immunity and spontaneously eliminate the parasitic population. Some level of age-dependent immunity develops, but it is not a strong immunity and animals can become reinfected.

These gastrointestinal nematodes infect goats and sheep, and feed on the blood of their host through the stomach wall at the adult stage. In addition, the lacerations the worm causes to blood capillaries results in micro-hemorrhages. Haemonchosis thus causes anaemia, mostly characterized by pale mucous membrane visible in the lower eye lid, which is the basis of the Famacha anemia chart to determine the level of treatment required (**Fig. 5A**). Other clinical signs of the infection are bottle jaw due to swelling of fluid (**Fig. 5B**), metabolic troubles preventing weight gain, wool loss and decreased milk production. In the extreme form of anaemia, red blood cells are depleted to such a rate that oxygen cannot be transported around the body. This leads to collapse and death of the animal and, thus, can result in dramatic loss of livestock production (Saddiqi et al., 2011). This, added to the cost of drugs to fight the parasite, makes it one of the most expensive diseases for the livestock industry (Miller et al., 1998).

However, although the search for haemonchosis treatment requires the study of *H. contortus* response to drugs, this nematode is very hard to study directly in the laboratory, mainly due to its parasitic life cycle, which requires conditions that cannot be reconstructed, so far, *in vitro*. Molecular and biochemical techniques routinely used to study various processes occurring at the cellular level are thus mainly not applicable to this parasite. *C. elegans*, having a free-living life cycle, has then been studied for decades as a model organism of this parasite.

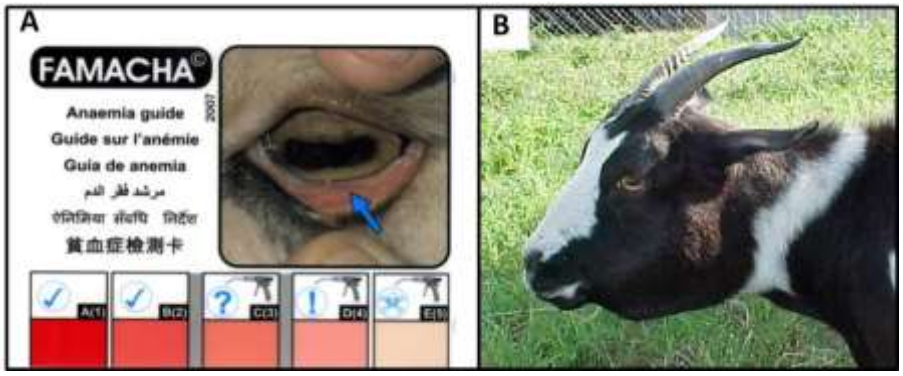


Figure 5. Main characteristics of haemonchosis. A. Famacha Anaemia chart showing a pale mucous membrane in the eye lid. (http://hoeggerfarmyard.com/wp-content/uploads/2012/08/famacha-21.jpg). **B.** Bottle jaw in a sheep with a severe barber's pole worm infection (http://www.floridameatgoats.com/Anemia.htm).

2. Caenorhabditis elegans, a model nematode

C. elegans is a free-living soil roundworm found worldwide that feeds on microbes, mainly bacteria. It displays a very close genome to *H. contortus* in the nematode phylum (**Fig. 1**). Indeed, Laing et al. (2011) found homologues and orthologues in *C. elegans* for 97.5% and 60% of *H. contortus* genes, respectively. This results in high similarities for the phenotypes of these two species in terms of morphology, physiology, but also molecular processes and biochemistry.

a) Life cycle of Caenorhabditis elegans

One of the main advantages of *C. elegans* over *H. contortus* for research studies is that this nematode has a free-living life cycle (**Fig. 6**) (www.wormatlas.org). The adult *C. elegans* exists as two sexual forms: hermaphrodites and males, which result from spontaneous non-disjunction occurring at very low frequency: 0.5% in the hermaphrodite germ line, and up to 50% through mating. One adult hermaphrodite produces about 300 – 350 eggs that develop in utero until the stage of the second cleavage, before eggs are laid and continue developing during about 9 hours. Eggs then hatch to free the L1 stage larvae, which develop to L2 in 12 hours, L3 in 8 hours, L4 in 8 more hours and become adult again in 10 hours. Eight hours later, adult are able to lay eggs. The total duration of its life cycle is thus short, about 3.5 days. A latent larval stage, the dauer stage, exists after L1 stage under stress conditions, and allows *C. elegans* to wait for more favorable conditions to restart its normal life cycle from the L4 stage.

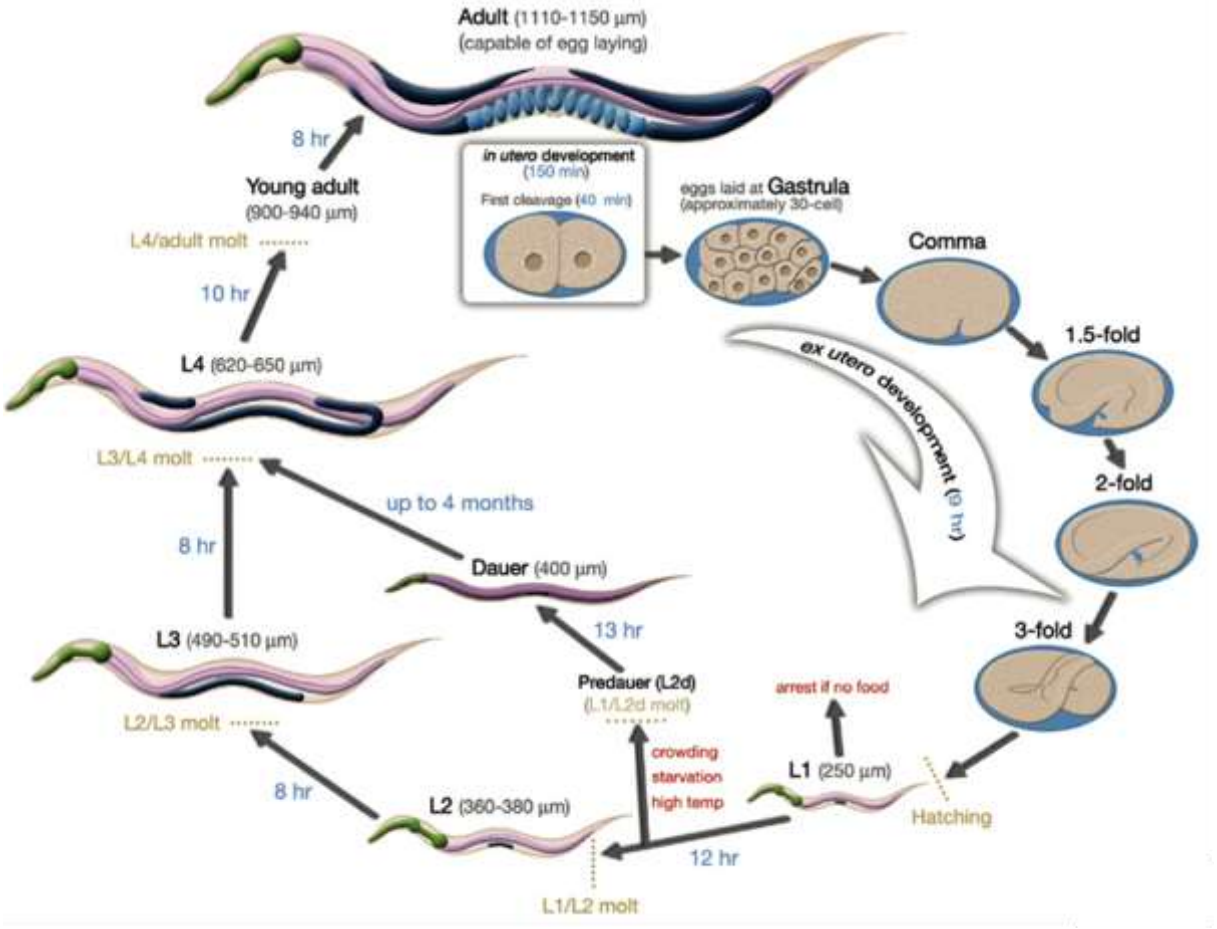


Figure 6. Life cycle of *Caenorhabditis elegans* at 22°C. 0 min is fertilization. Numbers in blue along the arrows indicate the length of time the animal spends at a certain stage. First cleavage occurs at about 40 min post-fertilization. Eggs are laid outside at about 150 min post-fertilization and during the gastrula stage. The length of the animal at each stage is marked next to the stage name in micrometers (µm). Image from http://www.wormatlas.org/ver1/handbook/anatomyintro/anatomyintro.html

b) Anatomy of *Caenorhabditis elegans* hermaphrodite adult

The anatomy of *C. elegans* has been extensively studied at the electron microscopy level, and its complete cell lineage has been made available (www.wormatlas.org). The hermaphrodite adult *C. elegans* is about 1mm long and presents an unsegmented, cylindrical shape that becomes narrower on the extremities (**Fig. 7A**). Similarly to *H. contortus* presented before, its whole body is covered by a striated cuticle secreted by epithelial cells such as hypodermis and seam cells on the two lateral regions of the worm, over which the alae forms at all stages except L2 – L3. A new cuticle is secreted at each stage with a molt at the end of the previous stage.

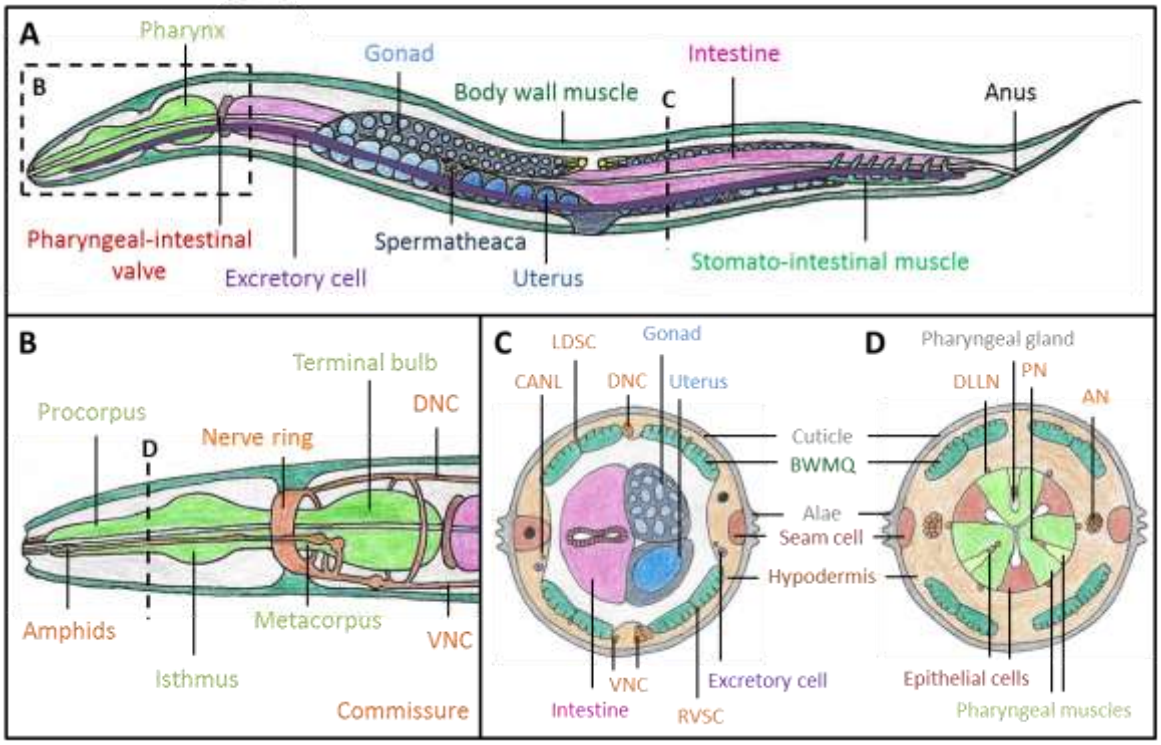


Figure 7. ***Caenorhabditis elegans*** **hermaphrodite adult anatomy. A**. Major anatomical features. The cuticle and nervous systems are not shown for better visibility of other organs. Two of the four quadrants of body wall muscles are represented. **B.** Zoom on the head of the nematode for simplified representation of the nervous system: the nerve ring and only the amphidial neurons of the head are shown for better visibility, as well as the start of the dorsal and ventral nerve cords that run along the entire body. **C.** Cross-section through the mid-body region of the *C. elegans* hermaphrodite (location marked with a dotted line in B.) showing the intestine and gonad at the centre of the pseudocoelomic cavity. **D.** Cross-section through the pharyngeal region of the *C. elegans* hermaphrodite (location marked with a dotted line in A.) showing the four muscle quadrants surrounded by the epidermis and cuticle. Only the nerve processes of the amphids on the two sides of the pharynx and the six labial nerves are shown for simplification. DC = Dorsal cord, VNC = ventral nerve cord, CANL = left CAN neuron with unknown function, LDSC = left dorsal sublateral cord, RVSC = right ventral sublateral cord, DLLN = dorsal lateral labial nerve, PN = pharyngeal neuron, AN = amphidial neuron. Schematic representation constructed from images found at www.wormatlas.org.

The nervous system organization has been described at the individual neuron level in *C. elegans* adults. The somatic nervous system is structured by 282 neurons with cell bodies clustered in ganglia in the tail, and the head where they form the nerve ring around the metacorpus part of the pharynx (**Fig. 7B**). The 4 main classes of neurons are motorneurons connecting muscle cells, sensory neurons that sense many stimuli such as temperature, chemicals, ambient osmolarity, oxygen level, pH, light and mechanical stimuli, interneurons connecting them and polymodal neurons performing at least two of these three possible functions. Most of these neurons travel longitudinally along the worm from the nerve ring,

either towards the tip of the head, or throughout the entire body. These are essentially located within the ventral and dorsal nerve cords, with processes located between the hypodermis and body wall muscles (**Fig. 7C**) and linked by commissures (**Fig. 7B**). In contrast, the 20 pharyngeal neurons are located directly among pharyngeal muscles (**Fig. 7D**).

The digestive system is similar to that of *H. contortus*, starting with the mouth that opens on the anterior end, followed by the pharynx and intestine structures linked by the pharyngeal-intestinal valve (**Fig. 7A**). These structures occupy the centre of the worm and open through the anus on the ventral side, just before the tail whip. The intestine is found either on the left or on the right side of the reproductive system of the hermaphrodite, with a switch in the middle. Indeed, the self-fertilization is made possible by the presence of two, bilaterally symmetric, U-shaped somatic gonad arms (oviducts), each followed by a spermatheaca, ending with a central uterus that opens on the ventral side of the midbody with the vulva (**Fig. 7A**).

The muscle system contains two types of muscles. Body wall muscles, that run along the body (**Fig. 7A**), are obliquely striated and arranged into four quadrants, two dorsal and two ventral (**Fig. 7C and 7D**). They receive neuronal input from motor neuron processes located in nerve cords or in the nerve ring to permit locomotion (**Fig. 7B, 7C and 7D**). Other muscles, found in the pharynx and around the intestine, rectum and vulva, are nonstriated and allow the functions of feeding, defecation and egg laying, respectively.

The excretory system allows osmoregulation and waste disposal. It consists of two canals running along the body on the two lateral sides of the worms (**Fig. 7A**), linked on the ventral side of the posterior head with an opening close to the nerve ring on this side. Finally, the coelomocyte system is composed of the pseudocoelomic cavity and three pairs of coelomocytes that endocytose fluid from the pseudocoelom, which probably plays the role of a primitive immune system (**Fig. 7C and 7D**).

Thus, this very precise knowledge of the anatomy of *C. elegans* is a major asset in the understanding of the biology of nematodes.

c) A model nematode in the laboratory

Many other advantages are found in *C. elegans* for *in vitro* and *in vivo* studies. Its short life cycle and high number of eggs produced by a single adult makes it possible to quickly generate genetically identical progeny by self-fertilization of the hermaphrodite, which can, for example, be used for generation of resistance lineages to increasing doses of a drug. On the

other hand, isolation, maintenance or spreading of mutations across strains can be done with male mating. Mutations can also be easily obtained by homology recombination, with many deletion strains for various genes made available on the *Caenorhabditis* Genetics Center (http://cbs.umn.edu/cgc/home), and gene rescue with the homolog from another species is routinely performed by microinjection.

Plus, *C. elegans* is transparent throughout its life cycle, which simplifies the record of visible phenotypic evolutions. Moreover, it easily feeds on *Escherichia coli* bacteria on agar plates or in liquid cultures. Its genome being short, about 100 millions of base pairs found across 5 autosomes and 1 X chromosome, *C. elegans* was the first animal to have its entire genome totally sequenced, in 1998. This, added to its stereotypical development and simple body plan composed of about 1000 somatic cells, made it a model of choice in many fields of life sciences such as genomics, embryogenesis, cell biology, neurosciences and aging. Its large palette of behavior also allows the study of many complex processes such as locomotion, feeding, mating, egg laying, memory, and sensory responses to various stimuli like touch, temperature and chemicals.

However, the difference in the free-living versus parasitic life cycles of *C. elegans* and *H. contortus*, respectively, must require the expression of different genes between these two species. Especially from L3 to adult stage, the *H.contortus* transcription changes for life within a host, e.g. by increasing peptidases production for blood-feeding activity (Schwarz et al., 2013), while *C.elegans* stays free and feeds from its environment. Nevertheless, the common pharmacology found across species of anti-parasitic drugs commonly used has allowed, over many years, the successful study and application of findings in *C. elegans* to various parasites (for a review: Holden-Dye and Walker, 2007).

II. ANTHELMINTICS AND MECHANISMS OF RESISTANCE

In order to reduce haemonchosis and its deleterious impact on livestock, efficient drugs against *H. contortus* have been sought. However, due to the selection of resistant parasites by each specific anthelmintic (AH) class, there is an ongoing demand for new AHs that overcome existing resistance (Holden-Dye and Walker, 2007). The main classes of anthelmintics efficient against roundworms, including *H. contortus,* are presented below, ordered by date of their discovery. Some anthelmintics initially used against other pathogens are also listed, because they later became a possibility for combination with commonly used parasiticides to regain control after resistance had arisen.

1. Main classes of anthelmintic drugs

a) Imidazothiazoles

Imidazothiazoles, also called tetrahydropyrimidines, are the oldest anti-parasitic agents used to treat cattle. They are agonists of the acetylcholine receptors (AChR) present at the surface of muscle cells. Tetramisole and pyrantel were the first two AHs of this family to be described in 1970 (Aceves et al., 1970; Aubry et al., 1970). These drugs cause prolonged activation of the AChR at neuromuscular junctions leading to sustained contraction of the somatic muscle, which results in paralysis of nematodes without causing their death. Pharmacological studies in *Ascaris suum* have shown that different AChR subtypes exist: N-AChR, B-AChR and L-AChR, each having various subunit compositions and being preferentially activated by nicotine, bephenium and levamisole respectively (Qian et al., 2006). *C. elegans* L-AChR is composed of five subunits and three proteins are essential to its function, allowing its assembly and targeting to the membrane (Boulin et al., 2008). It is also activated by pyrantel, though to a lower extent than levamisole, and while nicotine cannot activate this receptor it is a potent allosteric inhibitor.

b) Benzimidazoles

Benzimidazoles (BZ) were the first class of broad-spectrum anthelmintics established with the discovery of thiabendazole in 1961 (Gordon, 1961). They remained the mainly used anti-parasitic agents until the 1980s. The success of these AHs, that also comprise albendazole, febendazole, mebendazole and oxfendazole, is mainly due to their selective toxicity for helminths (Lacey, 1990). These drugs prevent microtubule polymerization by binding to β-tubulin (Lacey, 1988). Capping of the associating end of the microtubule, which constantly dissociates on the other extremity, results in its depolymerization (Lacey, 1990). The disintegration of the microtubule matrix, first observed in *Ascaris suum*, impairs many critical cellular processes such as cell division and transport, causing cell death, and eventually leading to the death of the parasite (Borgers and De Nollin, 1975). The variability in efficacy of each drug *in vivo* was correlated to their affinity for β-tubulin, except for oxfendazole and albendazole sulfone (Lubega and Prichard, 1991). Triclabendazole, on the other hand, is not effective against nematodes and cestodes, but controls all larval and adult stages of the parasitic trematode *Fasciola hepatica* (Boray et al., 1983).

c) Salicylanilides and Cyclodepsi-peptides

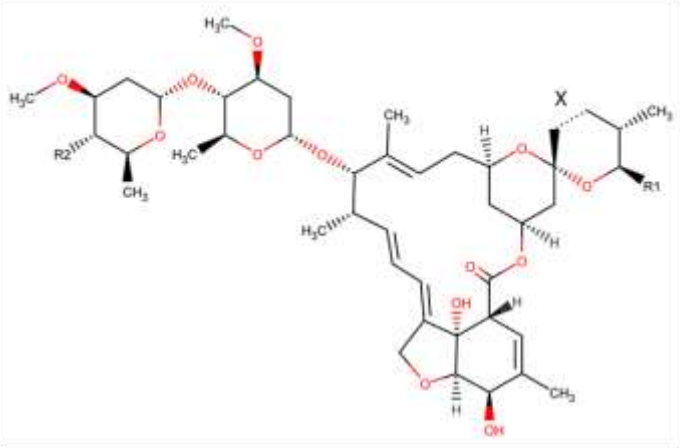
Salicylanilides represent a wide range of compounds initially developed as antifungal agents (Kraushaar, 1954). Among this class of anthelmintics, closantel and rafoxanide are mostly used, which are highly efficient against the adult stage of the trematode *F. hepatica* and blood sucking nematodes such as *H. contortus* (Swan et al., 1999; Van Den Bossche et al., 1979). Their molecular mode of action, however, is not completely elucidated. By uncoupling oxidative phosphorylation in the cell mitochondria, they disturb ATP production critical for energy metabolism, but this could also be due to initial impairment of glycolysis (Fairweather and Boray, 1999). The consequence is spastic paralysis of the parasite that dies from starvation after detachment.

The first compound of the cyclodepsi-peptides anthelmintic class, discovered in 1992, was PF1022A, a natural product of the fungus *Mycelia sterilia* that grows on the leaves of *Camellia japonica* (Sasaki et al., 1992). Emodepside (EMD) is a derivative of this compound and is licensed for treating roundworms and hookworms in cats. PF1022A and EMD are also known to be efficient against *H. contortus* strains resistant to IVM, BZ and LEV (Harder et al., 2005). The mode of action of EMD has been extensively studied and starts with the activation of a presynaptic latrotophilin receptor. This induces a complex signaling cascade that leads to a flaccid paralysis of pharyngeal and somatic muscles in nematodes.

d) Macrocyclic lactones

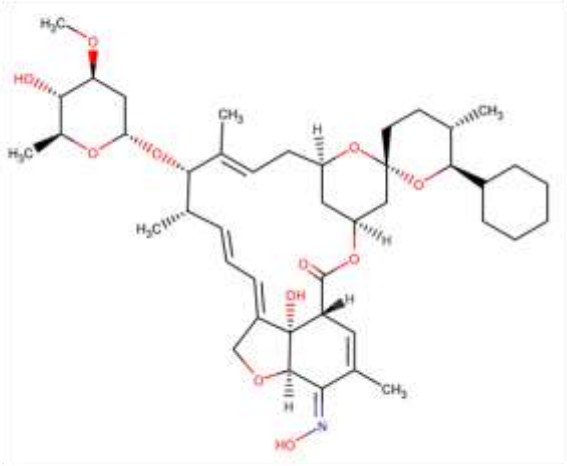
Since Ivermectin (IVM), the first registered macrocyclic lactone (ML) anthelmintic, was introduced on the market in 1980, other ML showing the same type of activity have been extensively developed. Abamectin, eprinomectin, doramectin and selamectin belong to the class of avermectins and moxidectin (MOX) and mylbemycin oxime are of the mylbemycin class (Haber et al., 1991). These two sub-families all share a macrocyclic lactone nucleus, but the mylbemycins lack the sugar group(s) present at the C13 of the macrocyclic lactone ring in avermectins, thus being more lipophilic (**Fig. 8**).

A



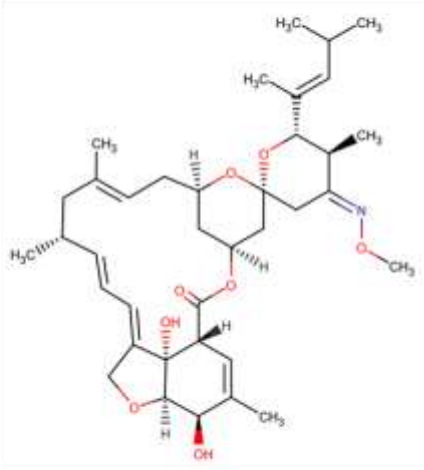
	X	R1	R2
Ivermectin B1a	CH_2-CH_2	$CH(CH_3)CH_2CH_3$	OH
Abamectin B1a	CH=CH	$CH(CH_3)CH_2CH_3$	OH
Doramectin	CH=CH	cyclohexyl	OH
Eprinomectin B1a	CH=CH	$CH(CH_3)CH_2CH_3$	$NHCOCH_3$

B



Selamectin

C



Moxidectin

Figure 8. Chemical structures of the mostly used macrocyclic lactone anthelmintics. A. Main avermectins. **B.** Selamectin. **C.** Moxidectin. (Structures found at: https://www.rsc.org/Merck-Index/).

IVM is the product of *Streptomyces avermitilis* fermentation. It binds to and activates a wide range of receptors, such as the gamma-aminobutyric acid (GABA)-gated chloride channel (Robertson, 1989). In nematodes, IVM shows the highest affinity for glutamate-gated chloride

channels (GluCls) whose irreversible activation leads to pharynx and body muscle paralysis (Cully et al., 1994; Forrester et al., 2003). These receptors can be formed by two distinct types of subunits: α subunits respond to IVM when expressed in *Xenopus* oocytes, whereas β subunits only respond to glutamate. Association of five subunits composed of the two types α and β produces a glutamate-gated channel that can be activated by IVM. Several subtypes for each subunit have been identified: GluClα1, 2A, 2B, 3A, 3B, 4 and GluClβ in *C. elegans*, HcGluClα, α3A, α3B and β in *H. contortus* (Yates et al., 2003). Each subunit is composed of a four α helices transmembrane domain where IVM binds, whereas Glutamate binds to the extracellular β leaflet domain (**Fig. 9**) (Hibbs and Gouaux, 2011).

Even though GluCl channels are not expressed in muscle cells of the body, paralysis appears to be due to the presence of these receptors on the motoneurons that synapse onto muscles (Portillo et al., 2003). Similarly, GluClα subunits in *C. elegans*, and HcGluClα3B subunits in *H. contortus* have been found to be expressed in the motor neurons that innervate pharyngeal muscle cells (Yates et al., 2003). IVM is thought to cause disruption of neuronal signaling to this organ, leading to pharyngeal pumping inhibition that provokes death of the worms by starvation (Geary et al., 1993).

ML are used to treat gastro-intestinal nematodes and ectoparasitic infections in the veterinary industry as well as filarial worms such as *Onchocerca volvulus* in humans (Wolstenholme and Rogers, 2005). The modes of action of avermectins in this last parasite are killing of the microfilarial stage and inhibition of reproduction in the adult worms. Ivermectin (Mectizan®) has been the drug of choice to treat onchocerciasis for more than 30 years and has been supplied free of charge by Merck & CO, Inc. in developing countries with the Community-directed treatment with ivermectin (CDTI) program. For this reason, the discoverers of Ivermectin, William C. Campbell and Satoshi Omura, have received the Nobel Prize in Physiology or Medicine in 2015, together with Youyou Tu who discovered a novel therapy against Malaria.

The success of ML in veterinary medicine can be attributed to their high potency at low dose, as compared to other anthelmintics previously used. Indeed, about 0.2 mg of ivermectin/kg of body weight are required to treat animals instead of about 7.5 mg/kg for imidazothiazoles and up to 15 mg/kg for benzimidazoles and the salicilanilide closantel (Thomaz-Soccol et al., 2004). ML present other interesting properties such as a broad-spectrum activity and ease of administration, orally, sub-cutaneously or topically. Even though the latter method leads to decreased bioavailability in cattle (Lespine et al., 2009), it

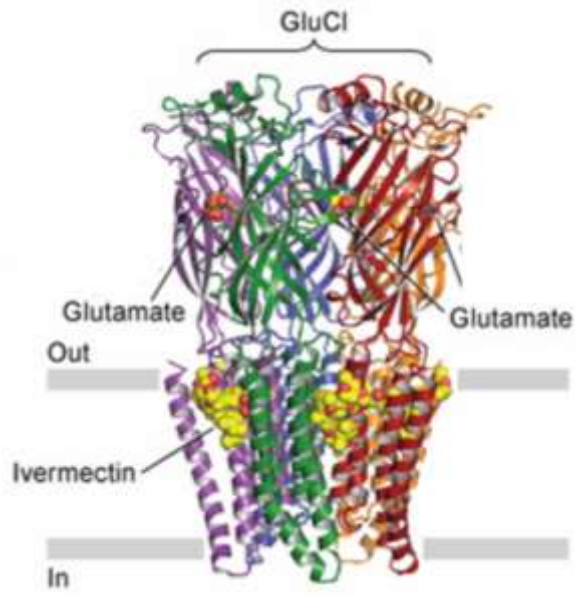


Figure 9. Interaction of Glutamate and Ivermectin with GluCl channel. Each of the five subunits of the GluCl channel is represented as its secondary structure with a single color, purple, green, red, orange or blue. The ligands ivermectin and glutamate are represented as spheres with carbon atoms in yellow, oxygen in red and nitrogen in blue. Adapted from (Hibbs and Gouaux, 2011).

allows the compound to be distributed throughout de body via the blood and lymph (Lespine et al., 2006a). Entering the cells is then facilitated by the high hydrophobicity of these molecules which can readily go through the cell membrane by passive transport. ML then persist inside the hosts over a moderately long period of time, thus protecting the ruminants against re-infection (Kerboeuf et al., 1995).

e) Amino-acetonitrile derivatives

Amino-Acetonitrile Derivatives (AADs) are the most recent chemically synthesized compounds used to fight various parasites. They were found, in 2008, to have a higher efficiency against most parasitic nematodes resistant to imidazothiazoles and ML (Kaminsky et al., 2008a). Due to the nematode-specificity of their target, the ACR23 subunit of the N-AchR, they show low toxicity to the host. AADs cause the paralysis of both the body wall and pharynx muscles leading to the death of *H. contortus* and other parasites at the L4 and adult stages. Monepantel was the first drug of this class described as a candidate for development in 2008 (Kaminsky et al., 2008b). Various mutations were found in two genes coding for N-AChR subunits in a population of *H. contortus* selected for reduced sensitivity to monepantel (Rufener

et al., 2009). The direct binding of monepantel to the DEG-3 subfamily of N-AChR was then demonstrated, and this AAD was found to be an allosteric modulator of the DEG-3 channels (Rufener et al., 2010). However, resistance to these compounds has already been described and compromises parasite control in farm animals (Van den Brom et al., 2015).

2. Mechanisms of resistance to anthelmintics

The emergence of resistance to anthelmintics over the last 50 years has been due to different factors: frequent use of a single drug, under-dosing and mass treatment that contributed to the selection of resistant strains (Jabbar et al., 2006). To reduce this, several measures have been taken, such as the combination of different drugs. Despite this, resistance to nearly all classes of AHs previously described is now widespread worldwide in sheep and goats, but also in cattle and horses (Kaplan, 2004). Various mechanisms are observed in parasites that allow them to overcome the negative effects of different drugs.

a) Modifying the drug target

Changes in the gene sequence of the targeted proteins can prevent its interaction with the drug. For BZ, several mutations on β-tubulin genes have been observed in resistant strains of *H. contortus* which lead to a decrease of the high-affinity binding of BZ to tubulin (Lubega and Prichard, 1991). To identify such a mechanism in BZ resistance, a candidate gene approach was set up on the isotype 1 β-tubulin gene. A single nucleotide mutation was identified in resistant *C. elegans* and *H. contortus* (Roos et al., 1993). Transgenic analyses were performed in order to study the modulation of the susceptibility of *C. elegans* strains by the identified mutation. The Phe-Tyr substitution at codon 200 was thus confirmed to be responsible for resistance to BZ (Kwa et al., 1995). Further studies showed that this polymorphism is associated with a resistant phenotype to this class of AHs in a wide number of species (Gilleard, 2006). More recently, two other point mutations, F167Y and E198A, were detected in resistant strains of several species (Ghisi et al., 2007; Silvestre and Cabaret, 2002), F167Y and F200Y appear to be mutually exclusive (Beech et al., 2011; Mottier and Prichard, 2008).

Levamisole resistance is not yet fully elucidated and knowledge regarding it depends on the species studied. In *C. elegans*, a loss of susceptibility has been linked with the expression of 21 genes. Some of them code for L-AChR subunits, others for proteins involved in assembly and processing of these subunits or regulation of the channel, and some for proteins involved in the calcium signaling cascade (Martin et al., 2012). In *H. contortus*, a truncated form of Hco-

UNC-63, a subunit of L-AChR, was detected only in resistant strains and inhibited the function of the receptor when co-expressed with the wild-type in *Xenopus* oocytes (Boulin et al., 2011). A novel approach, the cDNA-AFLP (Amplified Fragment Length Polymorphism) allowed the comparison of transcript profiles between susceptible and resistant strains of the parasite (Neveu et al., 2007). This led to the identification of the *Hco- acr-8b* gene, a truncated form of the *Hco-acr-8* gene encoding an N-AchR subunit, that was found by RT-qPCR (Reverse Transcriptase – semi-quantitative PCR) to be expressed only in resistant strains, and thus potentially involved in levamisole resistance (Fauvin et al., 2010). This was later confirmed and an insertion/deletion of 63bp was identified in the Hco-acr-8 gene, just downstream from the splice acceptor site for the alternative third exon (Barrere et al., 2014). The absence of the 63bp indel is indeed responsible for a difference in the open reading frame, with the presence of an early stop codon that leads to the expression of Hco-acr-8b transcript, which was linked with resistance status. The identification of this indel was then developed as a tool for levamisole resistance detection.

Changes in the sequence of one of the avermectin and mylbemycin target receptors, the GluCl receptor, may also modify the susceptibility of worms towards this drug. A functional analysis in *Cooperia oncophora* suggested that a L256F substitution in this subunit was responsible for only a slight decrease in IVM sensitivity, but the quantitative variation relative to the susceptible strain was low (Njue et al., 2004). A similar L256F mutation also caused a loss of sensitivity in the *H. contortus* GluClα3B receptor (McCavera et al., 2009). Overall, the sensitivity to IVM was dramatically compromised in a synthetic triple GluCl knockout strain of *C. elegans* (DA1316), which caused loss of receptors for IVM. IVM-sensitivity could be restored to susceptible parental strain level by expressing subunits of *H. contortus* or *C. elegans*, Hco-AVR-14B or Cel-AVR-14B respectively, under the control of the *avr-14* promoter, thus indicating a theoretical role of this region in resistance to IVM (Glendinning et al., 2011). Mutations in the GABA type-A receptor, subunit HG1 of *H. contortus* was also shown to have a potent involvement in IVM resistance (Feng et al., 2002). Indeed, two different alleles of the gene encoding this subunit, linked either with sensitivity or resistance of the parasites to IVM, showed different IVM modulatory responses to GABA when expressed in *Xenopus* oocytes. The link between this gene and another one, *glc-5* or *HcGluClα*, and ML resistance has been also confirmed *in vivo*. In fact, these alleles have been found to protect the worms against the inhibition of adult feeding and larval movement caused by drugs (Beech et al., 2010). In *C. elegans*, a link was found between a four amino-acid deletion in the ligand binding domain of

the GluClα subunit GLC-1 and resistance to avermectins (Ghosh et al., 2012). But one of the first mutagenesis studies, conducted on this drug target showed that a single nucleotide modification does not produce an important change in IVM sensitivity (Starich et al., 1995). Moreover, it was found that simultaneous mutations in three GluCl receptor subunits (GluClα1, 2 and 3) were associated with a high level synthetic resistance whereas mutations in only two of these three genes were not (Dent et al., 2000). Thus, the various subunits of this receptor seem to have a redundant function. In addition, changes in GluCl genes do not appear to be the main cause for ML resistance in parasitic nematode (Gilleard, 2006), and it is not found in the field.

b) Modifying drug biotransformation

Increasing the rate of modification of the drug to a non-toxic compound, or reducing the activation of pro-drugs can also alter their efficacy, so that such mechanisms may spread in the population by selection pressure (Cvilink et al., 2009). A biotransformation study performed *ex vivo* on microsomal fractions of *Fasciola hepatica* showed that the rate of triclabendazole metabolism into triclabendazole sulphoxide was significantly higher in triclabendazole-resistant flukes compared to susceptible ones (Alvarez et al., 2005). Metabolism of BZs was shown to possibly play a role in their resistance to this class of anthelmintics in various organisms. For example, *H. contortus* was shown to be more resistant to thiabendazole after glutathione S-transferase (GST) expression was induced by a cambendazole treatment (Kawalek et al., 1984). Plus, another team later found that inhibiting glutathione synthesis led to an increase in thiabendazole sensitivity in *H. contortus* resistant strains (Kerboeuf and Aycardi, 1999). However, these studies are indirect and the effect observed could be due to other induced mechanisms such as modification of transport.

On the other hand, ML resistance does not appear to be linked to the biotransformation of these molecules (Lespine, 2013). In fact, IVM and MOX were shown to have high chemical stability in sheep ruminal and abomasal content as they are only poorly metabolized by biotransformation enzymes usually detoxifiying xenobiotics (Lifschitz et al., 2005). Plus, the enzymes involved in this process, mainly cytochromes, differ between host species and drugs (Zeng et al., 1998; Zeng et al., 1996; Zeng et al., 1997). 60 to 80% of macrocyclic lactones are then found as the parental form in the plasma of the host (Gonzalez-Canga et al., 2009). Elimination of these compounds is thus thought to be mainly due to their transport.

c) Modifying drug transport

An increased efflux or decreased influx of the drug can also reduce its action inside the cell. This can, for example, be due to the overexpression of ATP-binding cassette (ABC) transporters from the multidrug resistance (MDR) family. As these transporters are able to expel ML out of mammalian cells, their overexpression in parasites could also be a mechanism of resistance to these drugs (Pouliot et al., 1997).

III. ABC MDR TRANSPORTERS AND ANTHELMINTICS RESISTANCE

1. ABC transporters: structures and functions

a) ABC transporters structures and classification

Proteins of the ABC transporters family all share a common structure: they contain at least one nucleotide binding domain (NBD) also called ATP binding cassette (ABC), which gives its name to the protein family (Hyde et al., 1990). To be active, ABC transporters need two such domains to bind to and hydrolyze ATP, thus supplying the protein with energy necessary for transport. ABC transporters are mostly membrane transporters, and contain transmembrane domains (TMDs) most often composed of 6 transmembrane helices (TMs) each. These TMDs are formed by a majority of hydrophobic amino-acids that allow the anchoring of the protein within the leaflets of the plasma membrane. Substrates bind within the funnel-shaped intertwining TMDs. Functional ABC transporters are either full-transporters that contain 2 NBDs and 2 TMDs, forming a typical "tandem" structure, or half-transporters that only contain one TMD and one NBD, and need to homo- or hetero-dimerize to be active (**Table 1**). Exceptions are soluble ABC transporters that are mainly expressed in the nucleus to act as gene regulators.

In all living kingdoms, ABC transporters contribute to cell homeostasis. In bacteria, ABC transporters can import compounds essential for cell viability and pathogenicity, or export endogenous molecules out of the cell (**Table 1**) (Davidson et al., 2008; Sarkadi et al., 2006). They can also promote the translocation of lipids from the inner to the outer leaflet of the cell membrane. In eukaryotes, most of ABC pumps extrude molecules from the plasma membrane, leading them to the extracellular compartment, but some ABC transporters can also be found in organelles such as the mitochondria, the endoplasmic reticulum, the peroxisomes, or

vacuoles in plant cells. In mammals, seven sub-families of ABC transporters are expressed in various tissues, and are responsible for the transport of a wide variety of compounds, as indicated in **Table 1**. They are composed of different combinations of NBD and TMD domains (Sarkadi et al., 2006; Szakacs et al., 2006):

- The ABCA sub-family contains 13 full transporters, ABCA1 to A13, with the same domain arrangement shown in **Table 1**. They are all mainly involved in lipid transport in different tissues.
- The ABCB sub-family contains 11 proteins. Three of them are full transporters (ABCB1, B4 and B11) and display remarkable multispecific properties. The eight other ABCB transporters are half-transporters and are expressed in internal membranes where they handle specific endogenous substrates.
- The ABCC family is composed of 13 full transporters. Seven of them, ABCC1 to C3 and C6 to C9, contain an additional TMD of 5 helices at the N-term of the protein ("TMD0" in **Table 1**), linked to the first TMD1 by the loop "L0" (Bryan et al., 2004; Deeley et al., 2006). ABCC7 harbors a supplementary cytosolic regulatory domain (R) that plays a critical role in the regulation of its function. Indeed, it is a specific ion channel that passively conducts chloride ions in epithelial cells. The R subunit must be phosphorylated in order to facilitate the channel gating (Gadsby and Nairn, 1999). During the transport of Cl^-, ATP binding on the NBDs only has a regulatory effect on the ionic conduction and ATP hydrolysis remains very slow. Various mutations in this protein, also called cystic fibrosis conductance regulator (CFTR), make it not functional, so that the ensuing default of Cl^- transport causes damages in various tissues of patients suffering from cystic fibrosis. The sulfonylurea receptors SUR1/ABCC8 and SUR2-ABCC9 are not transporters but form the ATP-binding subunit regulating the ATP-dependent potassium channels in pancreatic and heart cells respectively (Bryan et al., 2004). However, they also form the receptors of various compounds acting as blockers and openers of the K^+ channel, thereby presenting an unusual capacity of multispecific recognition of various drugs (Bessadok et al., 2011).
- All four ABCD members are half-transporters and transport various fatty acids.
- The unique ABCE, a regulator of protein synthesis, and the three ABCF proteins believed to play a role in inflammatory processes, lack TMD domains.
- All five ABCG transporters are half-proteins (ABCG1 to G4 and G8) with an inverted NBD-TMD arrangement. They form homodimers except for ABCG5 and G8.

Table 1. Localisation and function of different human ABC transporters of each family.

Protein	Tissue distribution - polarized cell localization	Function / disease when mutated	Topology
ABCA1	Ubiquitous	Cholesterol and lipid transfer onto HDL / Tangier disease	TMD1-NBD1-TMD2-NBD2
ABCA4	Rod photoreceptors	Transport of N-retinylidiene - phosphatidylethanolamine / various retinal diseases	TMD1-NBD1-TMD2-NBD2
ABCB1 / MDR1 / Pgp	Blood-brain barrier, liver, intestine, kidney, placenta, stem cells - apical	cholesterol and various phospholipids , Multi-drug resistance	TMD1-NBD1-TMD2-NBD2
ABCB2-3 / TAP1-TAP2	All cells	Transport of peptides into the endoplasmic reticulum for antigen presentation during IR	TMD-NBD (het)
ABCB4 ; B11/ BSEP	Liver - apical	Transport of phosphatidyl-choline ; bile salt / progressive familial intrahepatic cholestasis	TMD1-NBD1-TMD2-NBD2
ABCB5-10	mitochondria	Iron, heme export from mito-chondria / X-linked sideroblastic anemia and cerebellar ataxia (XLSA/A) (ABCB7)	TMD-NBD
ABCC1 / MRP1	Lung, testes, kidney, peripheral blood mononuclear cells, cardiac and skeletal muscle, placenta - basolateral	Drug resistance	TMD0-L0-TMD1-NBD1-TMD2-NBD2
ABCC2 / MRP2	Blood-brain barrier, liver, intestine, kidney, placenta, lung - apical	Efflux of organic anion	TMD0-L0-TMD1-NBD1-TMD2-NBD2
ABCC4 / MRP4	Ovary, testes, kidney, lung, prostate - apical & basolateral	Nucleoside transport	TMD1-NBD1-TMD2-NBD2
ABCC7 / CFTR	Exocrine tissues	Cl- ion channel / cystic fibrosis	TMD1-NBD1-R-TMD2-NBD2
ABCC8 / SUR1	Pancreas	Regulatory subunit of K^+(ATP) channel, insulin secretion / Neonatal diabetes mellitus	TMD-NBD
ABCC9 / SUR2	Heart, muscle	Regulatory subunit of cardiac and vascular K^+(ATP) channel	TMD-NBD
ABCD1-4	Peroxisomes	Long chain fatty acid transport / X-linked Adrenoleukodystrophy (ABCD2)	TMD-NBD (hom)
ABCE1 ; F1-3	Ovary, testes, spleen ; ubiquitous	Oligoadenylate binding protein. Regulation of protein synthesis	NBD
ABCG2 / BCRP/MXR	Blood-brain barrier, placenta, liver, intestine, breast, stem cells - Apical	Multi-drug resistance	NBD-TMD (hom)
ABCG1-4	Macrophages	Cholesterol transfer onto HDL	NBD-TMD (hom)
ABCG5 ; G8	Intestine, liver - Apical	Transport of sterols/ sitosterolaemia	NBD-TMD (het)

TAP = transporter associated with antigen processing; IR = immune response; CFTR = cystic fibrosis transmembrane conductance regulator; SUR = sulfonylurea receptor; MRP = multidrug resistance protein ; BCRP = breast cancer resistance protein; HDL = high density lipoprotein ; (hom) = homodimer ; (het) = heterodimer. Adapted from (Dean et al., 2001; Szakacs et al., 2008; Vasiliou et al., 2009).

b) MDR ABC transporters in mammals

Among ABC transporters, those called "MDR" are functionally similar to Pgp, that was first identified in the 1970's on cultured tumor cells to be responsible for cross-resistance to various cytotoxic compounds, often used in anticancer chemotherapy, and thus coined "multidrug resistance" (Dano, 1973; Juliano and Ling, 1976). Subsequently, Pgp has been shown to be involved in clinical multidrug resistance in patients bearing tumors, leading to chemotherapy failures. Although this MDR phenotype was found to be also due to either alterations of drug metabolism or apoptosis mechanisms, the overexpression of some ABC transporters at the cell surface remains the major cause, and have thus been widely studied (Gottesman et al., 1996).

These transporters are expressed on apical cell membranes in various tissues (**Table 1**), such as intestinal cell brush border limiting digestive absorption, hepatocytes biliary pole and kidney tubular cells mediating excretion. They then prevent transported drugs and potent toxic compounds from entering the systemic circulation and propagating throughout the body (Sarkadi et al., 2006). Among the 48 mammalian ABC transporters, three main members have been extensively studied as they were proven to belong to the MDR class: ABCB1, C1 and G2 (**Fig. 10**). They show various physiological functions depending on their site of expression (Lespine et al., 2009):

- The P-glycoprotein (Pgp/MDR1/ABCB1) is the most extensively characterized MDR ABC transporter. It was named after its identification in chinese hamster ovary (CHO) MDR cells that showed a decreased drug permeability linked with the overexpression of this 170 kDa glycoprotein (Juliano and Ling, 1976). Its role is critical at the blood-brain barrier where it prevents neurotoxicity in mammals (Roulet et al., 2003; Schinkel et al., 1994). It is also present at the apical membrane of enterocytes, where it expels compounds into the intestinal lumen, and in hepatocytes, where various drugs and toxic xenobiotics are eliminated via biliary excretion (Thiebaut et al., 1987). This wide tissue distribution prevents many exogenous compounds from remaining in the systemic circulation and causing toxicity to cells in the whole body. Its substrates are chemically diverse, ranging from anticancer drugs (paclitaxel, actinomycin D, doxorubicin, vincristine, etc) to HIV protease inhibitors such as ritonavir, and also include other cytotoxic agents, like colchicine, and cyclic or linear peptides such as valinomycin (Ambudkar et al., 1999). All these compounds are hydrophobic or amphipatic, so that they enter cells by passive diffusion (Gottesman and Pastan, 1993).

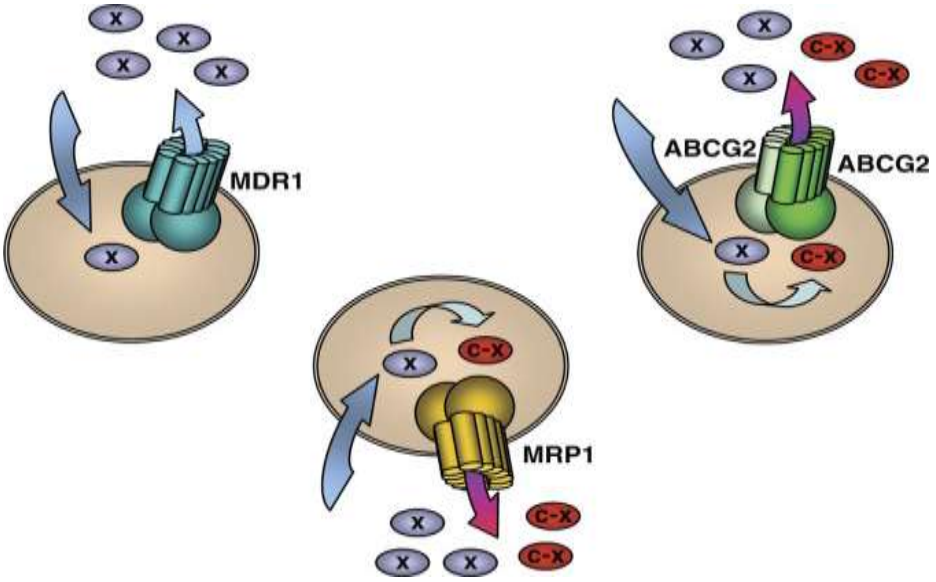


Figure 10. Function of the multidrug/xenobiotic ABC transporters. Multidrug/xenobiotic ABC transporters reside in the plasma membrane and extrude various hydrophobic and/or amphipatic xenobiotics and metabolic products. MDR1/Pgp transports hydrophobic compounds (X), while MRP1 and ABCG2 can extrude both hydrophobic drugs and intracellularly formed metabolites, e.g., glutathione or glucuronide conjugates (C-X). Figure from (Sarkadi et al., 2006).

- The multidrug resistance proteins (MRPs/ABCC1 to C6 and C10 to C12) are efflux pumps that have a great pharmacological importance because they display multi-specificity against many drugs. Plus, they are present at either apical or basolateral membranes of many epithelia (Borst et al., 1999). However, only ABCC1 is unambigously involved in MDR phenotype in tumor cells.

- The breast cancer resistance protein (BCRP/MXR/ABCP/ABCG2) is a half-transporter described as having a multidrug resistance capacity (Doyle and Ross, 2003). It is expressed on the apical membrane of many tissues, such as the intestine and the brain, but also in many stem cells, constituting the "side population". Plus, it is found on the placenta membrane where it protects the fetus from xenobiotics by expelling them (Chen et al., 2003), and in mammary glands where exogenous compounds are eliminated in the milk (Jonker et al., 2005).

c) Molecular structure and mechanism of ABC transporters

Structural studies have been critical for the elucidation of the 3D conformations of ABC transporters. *In silico* ligand-based modelling, as well as crystallography studies that have

allowed protein-based modelling of ligand binding via docking, together with homology modelling of unresolved ABC transporters have been powerful tools when combined to *in vitro* studies. They have allowed gaining insight into the processes of NBD dimerization and thus nucleotide hydrolysis, as well as substrates binding and transport, leading to a more and more precise understanding of the whole translocation mechanism of ABC transporters.

In first approaches, several isolated NBDs have been crystallized, belonging to ABC transporter of various bacteria. The crystal of the *Salmonella typhimurium* histidine permease HisP (Hung et al., 1998), followed by many more (Chen et al., 2003; Hopfner et al., 2000; Lewis et al., 2004; Schmitt et al., 2003; Smith et al., 2002; Verdon et al., 2003a; Verdon et al., 2003b; Zaitseva et al., 2005), brought insight into the mechanism of binding and hydrolysis of ATP by these domains.

NBDs are highly conserved and contain hallmark sequences of the ABC family across species (Higgins, 1992). The Walker A / P-loop motif binds to the γ-phosphate of ATP and the Walker B motif binds Mg^{2+} to hydrolyze ATP in all ATP hydrolyzing proteins. The A-loop consists of an aromatic residue located 25 amino acids upstream of the Walker A, and stabilizes bound ATP, as well as the H-loop / His-switch that interacts with the γ-phosphate of ATP (Zaitseva et al., 2005). The D-loop is an aspartate residue allowing cross-communication between the two ATP binding sites once NBDs have dimerized. The Q-loop is a γ-phosphate linker glutamine responsible for molecular communication with TMDs. They are all shared by many ATP-binding proteins, whereas the C-loop is an ABC signature LSGGQ motif, also called linker peptide, which sandwiches the ATP molecule bound to the opposite NBD (Ambudkar et al., 2003; Sauna and Ambudkar, 2007; Walker et al., 1982).

A functional ATP site is formed by the interaction of residues of both halves in a head-to-tail arrangement, thus sandwiching 2 ATP molecules. Specific residues critical for NBDs function were identified by co-crystallization with non-hydrolyzable ATP analogue and complementary *in vitro* experiments, mainly mutagenesis followed by ATP binding and hydrolysis assays. For example, the crystal of HlyB (Zaitseva et al., 2005), an element of the secretion machinery of *E.coli*, helped identifying the H662 residue of this protein as essential for its ATPase activity, by forming a catalytic dyad with E631. Many more were then mapped and found to be highly conserved among species, those involved in critical functions in mammalian Pgps were reviewed by (Ambudkar et al., 2006). Overall, these residues were found for each ATP-binding site at the interface of the two NBDs, on the Walker A and B motifs, H and Q-loop of one subunit and the C- and D-loop of the facing one (Hopfner et al., 2000; Sauna

and Ambudkar, 2007; Smith et al., 2002). ATP was also found to interact with the A-loop (**Fig. 11**) (Hung et al., 1998; Smith et al., 2002; Zaitseva et al., 2005). Dimerization of NBDs is thought to be mediated by ATP binding itself, as NBDs connect each other through ATP (Jones and George, 2004). The C-loop appeared critical for ATP hydrolysis and communication with the drug-substrate sites, probably by facilitating the formation of the nucleotide sandwich dimer (Tombline et al., 2004a; Tombline et al., 2004b). The latter process became increasingly considered over the years as the core of the catalytic cycle of ABC transporters (Higgins and Linton, 2004; Smith et al., 2002; Tombline et al., 2004a; Tombline et al., 2004b; Tombline et al., 2005).

Across the lipid bilayer, TMDs formed of 12 helices in total combine to form a funnel-like shape and delineate an inner chamber. The latter is either open to the cytoplasmic side of the membrane, in which case substrates can bind with high affinity to the transmembrane domains, or to the extracellular medium, with lower affinity of the TMDs to the substrates that allow their release (Ambudkar et al., 2006). The combination of substrate binding and export on the one hand with ATP binding and hydrolysis on the other hand is thought to be made possible by the presence of critical sequences such as the Q-loop that link the NBDs and TMDs in the 3D conformation of exporters.

The molecular 3D structure of the drug-binding site of ABC transporters started to be elucidated with the first successful crystallizations of entire ABC proteins, either half-transporters in homo- or hetero-dimers or full transporters. The first crystal of an ABC importer to be resolved was that of the bacterial importer of vitamin B_{12}, BtuCD (Locher et al., 2002). Its structure contains two times 10 TM helices, in a configuration perpendicular to the membrane plane, and thus structurally diverges from ABC exporters. Another bacterial ABC importer was later crystallized, ModBC, which harbors a more classical configuration with 2 times 6 TM helices forming an open inward-facing conformation (Hollenstein et al., 2007). These two examples illustrate the classification of ABC importers in two main types reviewed in (Rice et al., 2014).

In mammals, Pgp has become a paradigm for ABC exporters due to its discovery a long time ago and its involvement in resistance to anticancer chemotherapy. Garrigues et al., (2002) combined ligand-based modelling approach on Chinese hamster (*Cricetulus griseus*) Pgp (Cgr-Pgp) with *in vitro* ATPase measurements. They evidenced mutual relationships between a set of selected transported drugs, and revealed a bi-pharmacophoric binding pocket. On the one hand, actinomycin D (ACD), cyclosporin A (CSA) and verapamil (VRP) superimposed on the so-called pharmacophore 1 that accommodated the largest molecules.

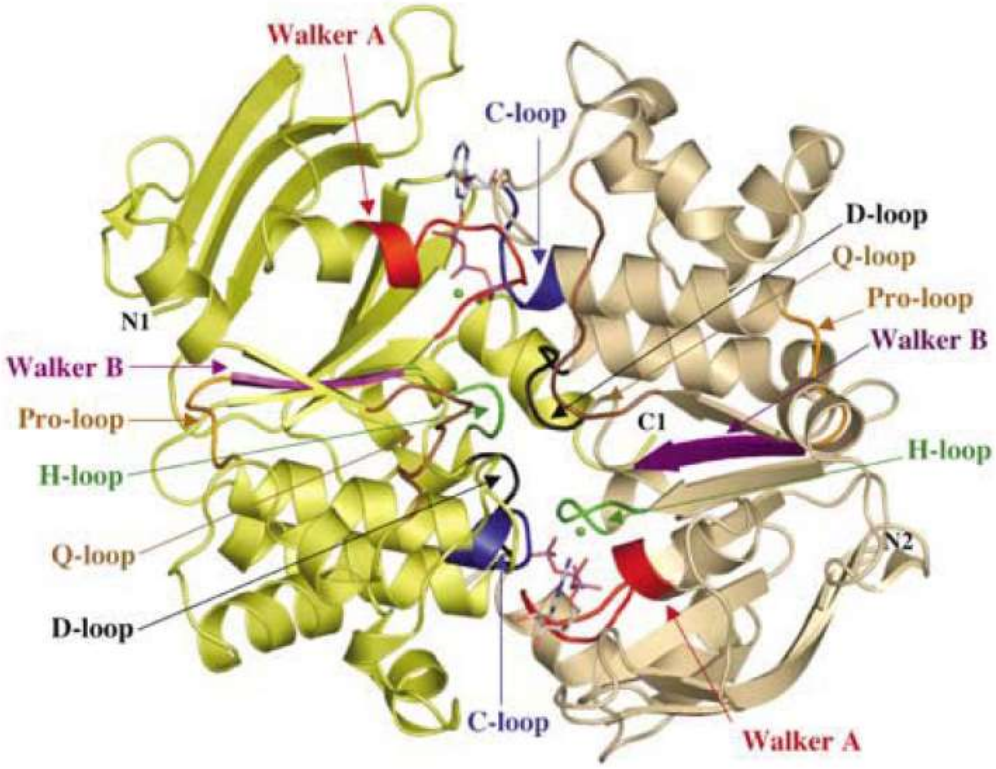


Figure 11. Crystal structure of the HlyB-NBD H662A dimer with bound ATP/Mg^{2+}. ATP in stick representation and Mg^{2+} (green spheres) are sandwiched at the interface of the two HlyB-NBD monomers (shown in light tan and light yellow). N- and C-termini of the individual monomers are labeled. Conserved motifs are colored in red (Walker A motif; Walker et al., 1982), brown (Qloop), blue (C-loop or ABC signature motif), magenta (Walker B), black (D-loop), and green (H-loop) and labeled accordingly. The figure was prepared using PyMol (www.pymol.org). Figure from Zaitseva et al. (2005).

On the other hand, tentoxin (TTX) and vinblastine (VBL) bound to pharmacophore 2 that bound smaller molecules (Garrigues et al., 2002). These two pharmacophores were found to present an interacting point, with VBL partly overlapping on pharmacophore 1. This is the first published work that simultaneously presents a multisite model and a pharmacophoric analysis of the drug binding site of Pgp. However, such a model is not sufficient alone to understand the precise molecular mechanisms involved in multidrug recognition and transport capacities, as this requires additional structural data.

Historically, the first image of Pgp has been revealed in 1997 from single particles electron microscopy imaging and analysis, with a very low resolution of 25 Å (Rosenberg et al., 1997). This showed a lateral communication between the interior of the protein and the surrounding membrane phase. Then, from 2D crystals imaging with a medium resolution of 8 Å, it became possible to discern a pseudo-symmetrical structure for the twelve helices of the whole protein (Rosenberg et al., 2005). Later, the first entire ABC exporter 3D-crystallized was

Staphylococcus aureus vancomycin intermediate resistance 1866 transporter (Sav 1866). This protein was crystallized in an outward-facing conformation, either ADP bound (Protein data bank code: 2HYD) (**Fig. 12**) (Dawson and Locher, 2006), or AMP-PNP (a non- hydrolysable ATP analogue) bound (PDB: 2ONJ) (Dawson and Locher, 2007). Its structure was resolved by X-ray diffraction at fair resolutions: 3.0 Å and 3.4Å, respectively. This allowed the description of the conformation of this exporter after release of a substrate. At this stage, the cavity was lined with mainly polar and charged amino acids. It was accessible from the outer leaflet of the membrane but not from the inner leaflet, although the bottom of the cavity was found to reach beyond the intracellular membrane boundary. Substrates were deduced to be able to escape either in the outer leaflet of the membrane, if hydrophobic, or in the aqueous extracellular compartment, for polar substrates. The NBDs, on the other hand, were shown to be interacting with each other during the ATP-bound state coupled to the outward-facing conformation. The major novelty was the description of intricately associated TM domains of an ABC transporter. Indeed, the TMDs were shown to swap 2 helices between each other, rather than being aligned side by side, as previously thought. Also, intracellular loops of one TMD were shown to interact with the opposite NBD, again showing higher interaction than expected. These features explained the limited maximum separation of the TMDs during the outward-facing stage, which flexibility is facilitated by their interaction with NBDs that get closer to each other when binding ATP. Thus, these data were in favor of a transconformation caused by the binding of ATP rather than the binding of substrates, which allows ABC exporters to alternate chelation and release of ligands.

Later, the multicopy suppressor of htrB, MsbA, an ABC lipid flippase (for LPS in particular) of *Salmonella typhimurium*, was also crystallized as an outward-facing structure, in complex with AMPPNP, at a resolution of 3.7 Å (**Fig. 12**) (PDB: 3B60) (Ward et al., 2007). This structure was found to be very close to Sav1866, confirming the 3D conformation of an ABC transporter after ATP hydrolysis. Further, the researchers also showed the correction of two earlier structures of orthologs of the same protein, *E. coli* MsbA in an open-inward conformation (PDB: 3B5W), and *Vibrio cholerae* MsbA in a closed-inward conformation (PDB: 3B5X), both without bound nucleotide. These data showed different distances between the NBDs of the two inward-facing MsbA proteins, and thus helped understanding the flexibility of various domains involved in the translocation process **(Fig. 13)**. Indeed, these three structures are proposed to be part of the successive intermediates of the transconformation steps required for the active transport cycle. ABC exporters have been since then envisioned as alternating cytosolic accessibility combined with high affinity for the

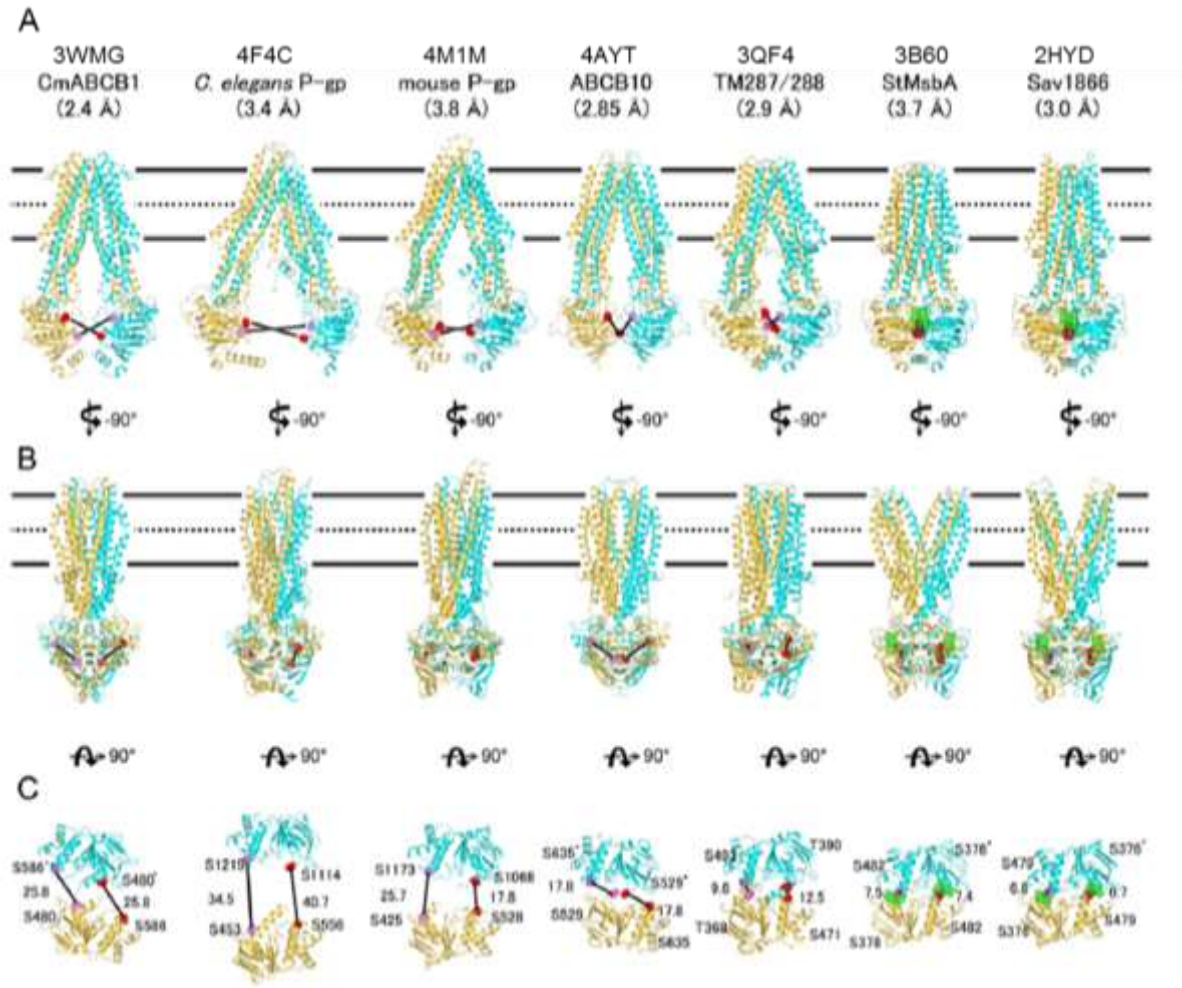


Figure 12. Comparison of the crystal structures of ABC transporters. Shown is alignment of overall structures (A and B) and NBDs (C) of CmABCB1 (Kodan et al., 2014), *C. elegans* P-gp (Jin et al., 2012), mouse P-gp (Li et al., 2013), ABCB10 (Shintre et al., 2013), TM287/288 (Hohl et al., 2012), StMsbA (Ward et al., 2007), and Sav1866 (Dawson and Locher, 2007), viewed parallel to the plane of the membrane (A and B) and from the extracellular side (C). The PDB code is indicated above the name of the protein and resolutions of each crystal structure are indicated in parentheses. The distances (Å) between the Cα atom of the first Ser (Ser480 in CmABCB1; Thr368/ Thr390 in TM287/288) in the P-loop/WalkerA motif and that of the Ser (Ser586 in CmABCB1) in the ABC signature motif of the adjacent NBD are indicated by bold numbers with the black lines. These Ser or Thr residues are shown as red spheres. The nucleotides bound to StMsbA and Sav1866 are shown as green spheres. Horizontal black bars represent the expected positions of the hydrophilic surfaces of the lipid membrane; gray bars represent the expected positions of the hydrophobic surfaces. Thick dashed lines represent the middle of the membrane bilayer. Adapted from (Kodan et al., 2014).

transported substrate, and exoplasmic accessibility during which a decreased affinity for the substrate allows its release.

Thus, the large cavity formed by TMDs during the open inward-facing conformation correlates well with the known LPS transport from the inner to the outer leaflet of the

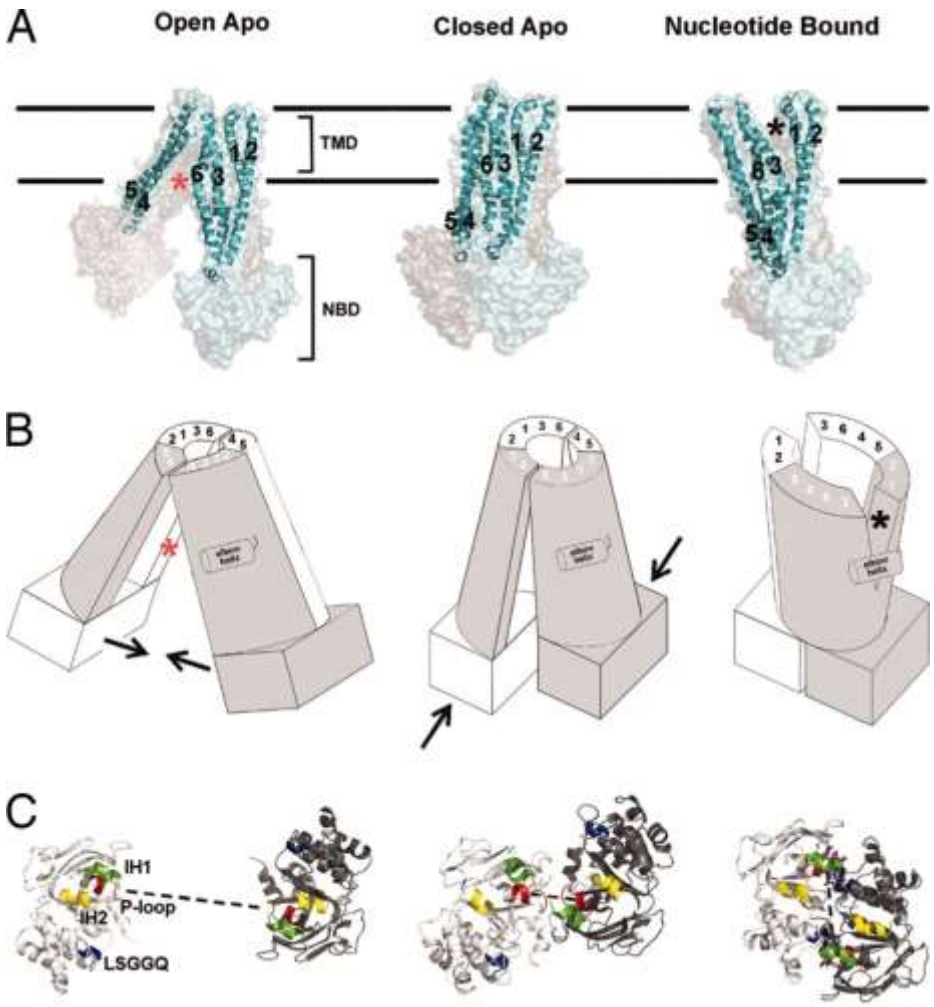


Figure 13. Summary of conformational changes in MsbA. (A) Conformational changes within the MsbA dimer alter the accessibility to the internal chamber from inward to outward facing. For clarity, only TM helices (labeled 1–6) of one monomer (cyan) are shown inside a surface rendering of the dimer. The open and closed apo conformations form an inward-facing V between TM4/TM5 and TM3/TM6 (red asterisk). The nucleotide-bound conformation (MsbA-AMPPNP) forms an outward-facing V between TM3/TM6 and TM1/TM2, just above the elbow helix (black asterisk). Upon nucleotide binding, TM4/TM5/IH2 moves, causing TM3/TM6 to split away from TM1/TM2, which results in an outward-facing conformation. Both inward- and outward-facing conformations are mediated by intramolecular interactions within a single monomer, but by different sets of helices. (B) Simplified cartoon model illustrating the points above. The relative position of each TM helix is labeled with a number (one monomer in white and the other in gray). The arrows illustrate the motions required to go to the next state. (C) Top-down view of NBDs (one monomer shown in white and the other in gray). IH1 (green) and IH2 (yellow) from both monomers are shown. In the absence of nucleotide (apo), the NBDs are in similar orientations with the ATP-binding half-sites (LSGGQ and P-loop) facing each other; the P-loops (red) are roughly aligned (dashed lines) with one another across the dimer interface. Upon nucleotide binding (AMPPNP - magenta), the canonical ATP sandwich is formed, aligning the nucleotide between the LSGGQ and P-loop. IH1 tracks with the cis-monomer, whereas IH2 tracks with the trans-monomer. The motion of the NBDs from closed-apo- to nucleotide-bound transmits a structural change (described above) to the TMs via IH1 and IH2, resulting in an outward-facing conformation. From (Ward et al., 2007).

bacterial membrane, which could not be explained by the outward-facing conformation alone. In addition, the crossover of TM4-TM5-IL2 and TM10-TM11-IL4 interacting with the opposite TMD, also found in these two inward-facing conformations, reinforces the stability and the symmetry of the dimerized protein all along the transconformation process.

The mouse ABCB1a structure crystallized by Aller and colleagues (Aller et al., 2009) was the first structure of a mammalian ABC transporter released in the PDB (3G5U), with a resolution of 3.8 Å (Table 2). Its conformation was open inward-facing and without nucleotide bound, thus most probably corresponding to the step before ligand binding. At this stage, NBDs of Mmu-ABCB1a were found to be about 20 Å apart, against about 50 Å for the open-inward structure of *Escherichia coli* (Eco-)MsbA. This revealed differences in inward-facing conformations, either between species or over time if the opening found for Mmu-ABCB1a was an intermediate of a likely dynamic process of "opening-closing of NBDs". In the inward-facing conformation, the cavity was found to open both to the cytoplasm for polar/charged ligands and to the inner leaflet of the membrane on two sides: between TM4-TM6 and TM10-TM12 for hydrophobic ligands. Residues lining the cavities were found to be mainly hydrophobic and aromatic, contrary to what was found for the outward-facing conformation of Sav1866. The very large volume, about 6000 $Å^3$, found for the cavity revealed a possibility of binding either a (very) large substrate or two smaller substrates simultaneously.

This drug binding site location was confirmed by the co-crystallisation of this protein with two stereoisomers of cyclic hexapeptide inhibitors: cyclic-tris-(R)-valineselenazole (QZ59-RRR) that bound to the middle of the inner cavity (PDB: 3G60), and cyclic-tris-(S)-valineselenazole (QZ59-SSS), which showed two binding sites across the previous one (PDB: 3G61). Several of the residues of interaction of verapamil previously identified on human (Hsa-) Pgp (Loo et al., 2006a, b; Loo and Clarke, 1997), correlated with the residues lining the two binding sites identified on Mmu-ABCB1a. They also showed a high degree of conservation between mammalian species, suggesting a common mechanism of multi-specific drug recognition.

Later, the first crystal structure of a nematode Pgp, *C. elegans* Pgp-1 (Cel-Pgp-1), was released with a resolution of 3.4 Å (**Fig. 12**) (PDB: 4F4C) (Jin et al., 2012). *In vitro* measurements of ATPase activity stimulation of Cel-Pgp-1 by various compounds known to be substrates of Hsa-Pgp showed a multispecific recognition capacity also for the nematode

Table 2. Overview of crystal structures of ABC exporters from various species.

Protein	PDB code	Resolution (Á)	Full / Half (hom/het)	Lgd / Nt / Unbound /	Conformation	Reference
Sav1866	2HYD/ 20NJ	3.0/3.4	Half (hom)	Nt bound	Outwards open	Dawson and Locher, 2007
Sty-MsbA	3B60	3.7	Half (hom)	Nt bound	Outwards open	Ward et al., 2007
Eco/Vch-MsbA	3B5W/X	5.3 / 5.5	Half (hom)	Unbound	Inwards open / closed	Ward et al., 2007
Mmu-ABCB-1	3G5U	3.8	Full	Unbound	Inwards open	Aller et al., 2009
Mmu-ABCB-1	3G60/1	4.4/4.35	Full	Lgd bound	Inwards open	Aller et al., 2009
Cel-ABCB-1	4F4C	3.4	Full	Unbound	Inwards open	Jin et al., 2012
TM287/288	3QF4 4Q4A/H/J	2.9	Half (het)	Un/Nt bound	Inwards open	Hohl et al., 2012 ; 2014
Mmu- ABCB-1	4KSB/C	3.8/4.0	Full	Unbound	Inwards open	Ward et al., 2013
Hsa- ABCB-10	3ZDQ	2.85	Half (hom)	Unbound	Inwards open	Shintre et al., 2013
Hsa- ABCB-10	4AYX/ T/W	3.3	Half (hom)	Nt bound	Inwards open	Shintre et al., 2013
CmABCB1	3WME/F	2.75/2.6	Half (hom)	Unbound	Inwards open	Kodan et al., 2014
CmABCB1	3WMG	2.4	Half (hom)	Lgd bound	Inwards open	Kodan et al., 2014
Mmu- ABCB-1	4M1M	3.8	Full	Unbound	Inwards open	Li et al., 2013
Eco-McjD	4PL0	2.7	Half (hom)	Nt bound	Outwards closed	Choudhury et al., 2014
NaAtm1	4MRN	2.5	Half (hom)	Unbound	Inwards open	Lee et al., 2014
NaAtm1	4MRP/R/S/V	2.35-2.97	Half (hom)	Lgd bound	Inwards open	Lee et al., 2014
Sce-Atm1	4MYC	3.06	Half (hom)	Unbound	Inwards open	Srinivasan et al, 2014
Sce-Atm1	4MYH	3.38	Half (hom)	Lgd bound	Inwards open	Srinivasan et al, 2014
Mmu- ABCB-1	4Q9H/ I/J/K/L	4.0-3.6	Full	Un/Lgd bound	Inwards open	Szewczyk et al., 2015
Cth- PCAT1	4RY2	3.6	Half (hom)	Unbound	Inwards closed	Lin et al., 2015
Cth- PCAT1	4S0F	4.1	Half (hom)	Nt bound	Outwards closed	Lin et al., 2015
Cje-PglK	5C78/6	2.9/3.94	Half (hom)	Unbound	Inwards open	Perez et al., 2015
Cje-PglK	5C73	5.9	Half (hom)	Nt bound	Outwards semi-closed	Perez et al., 2015
Hsa-ABCG5/8	5DO7	3.93	Half (het)	Unbound	Inwards open	Lee et al., 2016

The protein names are indicated in the first column with their previous nomenclature or with the code: 3 letters for the name of the species - Name of the protein - number. The PDB entry code in the second column is then followed by the resolution (Resol) in the third column, and the indication of full or half-transporter and homodimer (hom) or heterodimer (het) in the fourth column. The fifth column indicates if the protein was found unbound, ligand (Lgd) bound or nucleotide (Nt) bound. The sixth column indicates in which conformation it was found, and the references are listed in the last column.

transporter. However, several differencies were found between the structure of Cel-Pgp-1 and the unique mammalian structure previously released 3G5U of Mmu-ABCB1a. The NBDs of 4F4C were separated by about 10 more Å than in 3G5U. Two semi-helices, named TMa and TMb, were found on the N-term extremity previously found to be cytosolic in other structures, closing the space between TM4 and TM6 of Cel-Pgp-1. On the opposite side, TM10 and TM12 were found to be discontinuous, either forming a bigger entrance on this side for substrates to reach the inner pocket through the inner leaflet of the membrane, or facilitating the flexibility necessary for the transconformation of the protein. Finally, residues attribution showed a frameshift of 1 residue on TM3 and TM5, and 4 residues on TM4, as compared to 3G5U.

The 3G5U structure was then corrected several times, by different authors. The first correction in 2013 shifted 1 residue on each of the TM5, 6 and 8, and 4 residues on TM4, to obtain the new PDB structure 4LSG released without publication. In addition, 2 novel structures of mouse Pgp were later published: 4KSB and 4KSC released from new crystals, at resolutions of 3.8 Å and 4.0 Å respectively (Ward et al., 2013). On these structures, the distance between NBDs was wider than previously, about 30 Å, the TM12 was continuous, and the attribution of residues on TM3-4-5 was modified.

Then, Li et al. (2013) published another Mmu-ABCB1 crystal structure with a new irradiation technique, released at 3.8 Å resolution (**Fig. 12**) (PDB: 4M1M). Many helices were shifted for all (TM3-4-5-12) or part of (TM8-9) their residues when compared to 3G5U, with new residues lining the inner pocket better matching with experimentally identified residues involved in ligand binding in Hsa-Pgp (Loo et al., 2003, 2006a, b; Loo and Clarke, 1997, 2001).

Finally, another unbound crystal (4Q9H), together with 4 structures of co-crystals of Pgp, in open inward-facing conformation, were released at resolutions of 3.4 Å to 3.8 Å. The latter bound to 4 different rationally-designed homotrimeric ligands: QZ-Ala (4Q9I), QZ-Val (= QZ59-SSS from Aller et al. (2009)) (4Q9J), QZ-Leu (4Q9K) and QZ-Phe (4Q9L) (Szewczyk et al., 2015). Because the previous co-crystals released were of lower resolutions: 4.4 Å for 3G60 and 4.35 Å for 3G61 (Aller et al., 2009), these new structures allowed more precise interpretations of ligand-Pgp interactions. Interestingly, two subsets of ligands were distinguished, differing by size and hydrophobicity. The subset A contained the small ligands QZ-Ala and QZ-Val that shared an upper and lower binding site, while the subset B containing the bigger hydrophobic ligands QZ-Leu and QZ-Phe essentially bound to a second upper site. The most striking feature of these co-crystals was the kinking of TM4 in response to the subset A of bound ligands, compared to the rather straight structure of this TM found in co-crystals

with the subset B of ligands. This helix appears as critical for ligand entry and binding, as it forms an access portal form the inner leaflet of the membrane, and mutations in this region of a Pgp of a different specie disrupted substrate transport (Kodan et al., 2014). It has been suggested that ligand binding could induce closure of NBDs, in the presence of ATP, to increase catalysis (Doshi and van Veen, 2013; Szabo et al., 1998), so that TM4 could be involved in TMD-NBD coupling. This is consistent with ATPase good activators (subset A) kinking TM4 while ATPase poor activators (subset B) maintained TM4 straight, as observed in Pgp structure in the absence of ligands.

Several ABC exporters from other species were crystallized in the meantime, confirming rather similar 3D conformations for these proteins across species (**Table 2**). Some interesting variations of structures and mechanisms between species were nevertheless revealed (**Fig. 12**). The heterodimeric transporter *Thermotoga maritima* TM287-TM288 (TM287/288) was crystallized with a resolution of 2.9 Å in its open inward-facing conformation (**Fig. 12**) (PDB: 3QF4) (Hohl et al., 2012). In contrast to other crystallized ABC transporters, NBDs were still in contact even without nucleotide bound. Furthermore, the inner cavity appeared accessible from the cytoplasm, but not from the inner leaflet of the membrane. Finally, only one ATP needs to be hydrolyzed by NBDs to allow transport by this protein. At variance, the human mitochondrial homodimer ABCB10 showed around 20 Å separating its two NBDs in the open inward-facing conformation, which was found in the absence (PDB: 3ZDQ) as well as in the presence of 2 nucleotide analogs (AMPPCP and AMPPNP) bound (PDB: 4AYX/T/W) (**Fig. 12**) (Shintre et al., 2013). Otherwise, the homodimer *Cyanidioschyzon merolae* ABCB1 was crystallized both unbound at 2.6 Å resolution (PDB: 3WMF) and bound to a unique allosteric inhibitor at 2.4 Å resolution (PDB: 3WMG), revealing an open-inward conformation in both cases, with a disordered TM4 in each subunit (**Fig. 12**) (Kodan et al., 2014). NBDs were apart from 25 Å in this inward-facing conformation of the protein. The transporter of antibacterial peptide microcin McjD of *E. coli* was crystallized at 2.7 Å resolution (4PL0) in a new nucleotide-bound closed outward-facing state (Choudhury et al., 2014). Its structure highly resembles that of outward-open MsbA and outward-open Sav1866, but without spanning of two TM helices on the opposite side of TMDs. This structure, closed to both sides of the membrane, was thus proposed as a transition state between the previous outward-facing state, after release of the substrate outside of the cell, and before release of ADP + Pi and subsequent return to the inward-facing conformation, that will allow binding of new substrates.

The *Saccharomyces cerevisiae* Atm1 mitochondrial transporter is the ortholog of human ABCB7 causing sideroblastic anemia XLSA/A when mutated. It was crystallized in inward open conformation either unbound (PDB: 4MYC) or glutathione-bound (PDB: 4MYH), at resolutions of 3.1 and 3.4 Å, respectively (Srinivasan et al., 2014). The two conformations were found to be almost identical, and showed a 6900 $Å^3$ positively charged internal cavity close to the interface between the inner membrane and the cytosol. GSH bound within this cavity, similar to the 2 binding sites just previously found for GSSG, in the bacterial homolog *Novosphingobium aromaticivorans* NaAtm1 (Lee et al., 2014). Indeed, this protein had been crystallized either unbound (PDB: 4MRN) or bound to various ligands (PDB: 4MRP-V). Thus, the binding site of GSH and GSSG appears close to the expected cytosolic entry of polar ligands, in contrast with the hydrophobic Pgp inhibitors QZ59-RRR and QZ59-SSS that bound much deeply in the inner pocket of Mmu-ABCB1 in the 3G5U crystal (Aller et al., 2009) (**Fig. 14**). Although this structure had been later corrected, the binding sites of such ligands showed to be similar in the co-crystals of Szewczyk et al. and match well with their supposed access through the opening of ABCB transporters to the inner leaflet of the plasma membrane (Szewczyk et al., 2015). Interestingly, one of the corresponding residues to those mutated in the human ABCB7 causing XLSA/A disease was found on the GSH binding site. Furthermore, the structure of the C-termini of each of the 2 monomers of this proteins were resolved for the first time, revealing a 24 amino-acid long alpha helix that tightly interact with one another, and appear to block the NBDs in open-inward conformations by connecting them at the level of the Walker A motif. The start of the unique C-terminal end of the full-transporter Cel-Pgp-1 was interestingly superimposed to one of these helices, in contrast to C-term ends of other proteins (**Fig. 15**).

Clostridium thermocellum PCAT1, a polypetide processing and secretion transporter, consists of a homodimer containing two peptidase domains, and was crystallized in two conformations (Lin et al., 2015). One crystal, without nucleotide, was found with a closed inward-facing conformation with the two accesses towards the inner leaflet of the membrane and the cytoplasm open, but NBDs in contact with each other (PDB: 4RY2). The other crystal bound to non-hydrolyzing ATP-gammaS was found in a closed outward-facing conformation, with the accesses towards the membrane and cytoplasm closed, as well as the access to the extracellular compartment, similar to *E. coli* McjD (PDB: 4S0F).

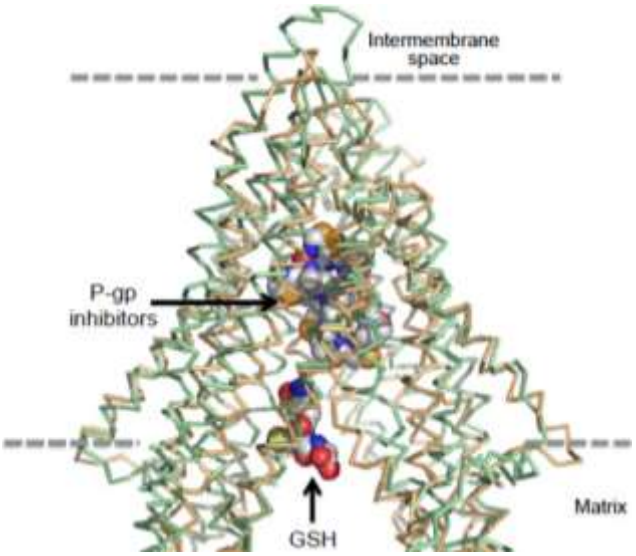


Figure 14. Different binding sites for GSH to Atm1 and for inhibitors to mouse Pgp. The superimposed structures of Sce-Atm1 with bound GSH (light orange ribbon) and mouse Pgp with its inhibitors (light green ribbon; (Aller et al., 2009)) show the strikingly different binding sites of the ligands. The Pgp inhibitors are located in the plane of the hydrophobic phase of the membrane, whereas GSH is close to the hydrophilic phase. From (Srinivasan et al., 2014).

Campylobacter jejuni PglK is a homodimeric active lipid-linked oligosaccharide (LLO) flippase, and was crystallized in three different conformations (Perez et al., 2015). Two unbound structures were found in an open inward-facing conformation. The space between NBDs was 44 Å (PDB: 5C78) or 30 Å (PDB: 5C76), similar to previous findings for Cel-Pgp-1 (Jin et al., 2012), and might be due to their obtention using two different detergents, rather than intrinsic properties of the protein. Furthermore, one nucleotide-bound conformation was found to be outward-facing (PDB: 5C73), in an intermediate occluded state between the totally closed *E. coli* McjD and *Clostridium thermocellum* PCAT1, and the fully open Sav1866. A similar inner cavity to what was previously shown for other ABC exporters was described.

The very last crystal of ABC protein released is that of human ABCG5/G8, a heterodimer that exports sterols (cholesterol and phytosterols) in enterocytes and hepatocytes (Lee et al., 2016). It was found as an unbound open-inward facing conformation without nucleotide bound, but probably cholesterol bound between TMDs, however in a too low resolution to be resolved. In this model, remarkably, no swapping of TM helices was found, and NBDs were in contact with each other at their extremities, thus forming a general

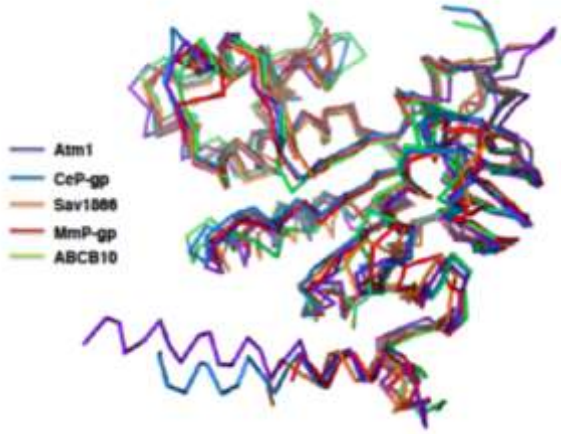


Figure 15. Superposition of the nucleotide binding domains of known bacterial and eukaryotic ABC exporters. The NBDs of the crystal structures of Atm1, *C. elegans* Pgp (Ce; PDB ID code 4F4C), mouse Pgp (Mm; 3G5U), bacterial Sav1866 (2HYD) and human mitochondrial ABCB10 (3ZDQ) were superimposed. The extra-long, C-terminal helix that confers stability to the Atm1 dimer in the open conformation can be clearly seen. Note that the C-terminal helix of *C. elegans* Pgp is resolved in only one half of this ABC transporter. From (Srinivasan et al., 2014).

structure of inward-facing conformation differing from the majority of exporters previously described. This appears consistent with a bioinformatic analysis based on homologies between the TMDs that led to a classification of the ABC exporters among two classes (Wang et al., 2009).

In addition to provide insights into different steps of the translocation of substrates across the membrane, these crystals have contributed to the better description of precise individual amino acids critical for ligand binding in the TMDs (review : (Shilling et al., 2006)). Indeed, *in vitro* data such as mutagenesis, ATP hydrolysis, ligand binding, or cysteine cross-linking studies, were done either in parallel to the crystal structure released (Aller et al., 2009; Choudhury et al., 2014; Li et al., 2013) or in separate studies of ABC exporters of the same or different species (Bessadok et al., 2011; Loo et al., 2006a, b; Loo and Clarke, 2001, 2002).

Plus, *in silico* docking calculations helped gaining insight into the binding properties of ligands on released crystal structures. The latter were also very useful to model other ABC transporters not crystallized so far. Then, the different types of *in vitro* or *in silico* studies previously mentioned could be perform in parallel to or based on the model to better characterize the properties of still poorly described transporters. To date, the structures of highest homology with human Pgp are human ABCB10, mouse ABCB1 and Cel-Pgp-1.

However, ABCB10 is homodimeric, and only mouse Pgp and Cel-Pgp-1, as well as sometimes the bacterial homologs, were used as templates for various homology modelling approaches of the Hsa-Pgp transporter.

It is important to note that *in silico* docking calculations performed on such structural bases were highly dependent on the reliability of the considered structural models. Chufan et al. (2013) based their flexible *in silico* docking calculations on the structure of mouse Pgp 4KSB (Ward et al., 2013). In their strategy, they left flexible the residues of this protein previously found to interact with QZ59 peptides. By docking tariquidar, cyclosporin A, valinomycin and FSBA, they confirmed that a majority of flexible residues showed an involvement rate > 50% for binding to these ligands. Plus, they showed that wide molecules could only bind close to the cytoplasmic opening of Hsa-Pgp whereas smaller ligands could fit in the inner cavity. *In vitro* transport and ATP hydrolysis studies showed that these ligands can bind to Hsa-Pgp on different sites, as mutations perturbing their binding to one site did not totally stop the transport of each molecule (Chufan et al., 2013). This was consistent with previous *in vitro* studies that had suggested the presence of several, at least two overlapping binding sites in Hsa-Pgp (review: (Ambudkar et al., 2006)).

In the study of Jin et al. (2012), the structure of Cel-Pgp-1, 4F4C, was used for modelling of Hsa-Pgp, and residues previously identified as involved in substrate binding in Hsa-Pgp were successfully checked to be facing the inner pocket of the obtained model. This indicated a high conservation of 3D conformation of ABC transporters across eukaryotic species. Three more studies were recently based on Cel-Pgp-1 structure of Jin et al. (2012) for modelling Hsa-Pgp (Jara et al., 2013; Prajapati and Sangamwar, 2014; Prajapati et al., 2013), and various compounds were tested for binding to Hsa-Pgp using docking calculations on these models (for a review : (Domicevica and Biggin, 2015)). This template could also be very useful and probably more accurate to model nematode full ABCB transporters showing closer homology to Cel-Pgp-1 than what Hsa-Pgp does.

Molecular dynamics were used as well to better characterize the function of these transporters, taking into account the large flexibility of this protein (review: (Domicevica and Biggin, 2015)). Some of them integrated the lipid bilayer in their simulations to obtain dynamics that reflect the functional motion of Pgp. However, these types of techniques require taking into account a time-scale larger than the currently simulated hundreds of nanoseconds for being representative of the time needed for the ligand translocation. Especially, the investigation of Hsa-Pgp transporter transconformation mechanism is critical for rationally overcoming

resistance to chemotherapy. Several models have been proposed along the years for the transport cycle of Pgps, as illustrated in Ambudkar et al., (2006). Despite recent advancements that were made using more and more precise techniques such as single particle electron microscopy (Moeller et al., 2015) or single-molecule Förster resonance energy transfer (FRET) (Husada et al., 2015), the successive steps of this complex mechanism still lack a fine description. Plus, the full transconformation mechanism might be unique to each Pgp given the differences already found for each of the crystals in their structure (Beis, 2015).

In spite of this lack of understanding of the entire mechanism of ligand translocation by Pgps, *in vitro* experiments have allowed for many decades the characterization of a wide variety of compounds as being able to interact with Pgp, and possibly reverse MDR resistance in mammals (review: (Ambudkar et al., 1999)). These were mainly identified in transport studies where Pgp was expressed in heterologous systems. In the presence of Pgp fluorescent substrates, such as calcein-AM, Hoechst 33342 or rhodamine 123, the incubation of compounds that led to accumulation of their fluorescence within cells gives indications on the nature of Pgp-mediated MDR-reversing agents. For a long time, the most potent competitor for Pgp binding used for these assays was the calcium channel blocker verapamil (VRP) (Tsuruo et al., 1981). But cyclosporine A was further found (Eneroth et al., 2001) to be 10 times more potent than VRP, i.e. in the same range as IVM. Loperamide and ketoconazole were also identified as 4 and 2 times more potent than verapamil, respectively. Dozens of other compounds have then been added to the extensive list of Pgp-interacting agents, thus describing the multispecific binding properties of this transporter. Jin et al., (2012) showed that several of the multiple substrates of Hsa-Pgp were able to stimulate the ATPase activity of Cel-Pgp-1, thus indicating that at least one nematode Pgp also shows the capacity to recognize multiple substrates.

2. ABC MDR transporters and resistance to anthelmintics

a) Anthelmintics transport by ABC MDR transporters in mammals

Pgp was the first MDR ABC transporter described as being involved in IVM detoxification in mice in 1994 (Schinkel et al., 1994). Indeed, after IVM treatment, knock-out homozygous mice for the Pgp genes died from an increased level of toxicity mainly in the brain, demonstrating the important role of this protein as a barrier against neurotoxic compounds. A similar hypersensitivity towards IVM was reported in a population of Collie dogs naturally lacking Pgp protein translation due to gene mutation (Roulet et al., 2003). *In vivo* disposition

not only of IVM but also of EPR were later shown to mainly depend on Pgp, particularly in the brain and intestine, when mdr1ab(-/-) mice were compared to WT mice (Kiki-Mvouaka et al., 2010). In contrast, moxidectin concentration in these tissues appeared mostly Pgp independent.

In vitro experiments have also permitted to gain insight into the transport of anthelmintics by Pgps. IVM was first characterized as a Pgp inhibitor and MDR-reversing agent that can directly bind to Pgp (Didier and Loor, 1996; Pouliot et al., 1997). *In vitro* transport experiments of various anthelmintics across the plasma membrane of cells expressing mdr1 was also monitored by following the intracellular accumulation of fluorescent substrates of Pgp (Griffin et al., 2005). IVM and SEL thus appeared as potent Pgp inhibitors whereas MOX was 100 times weaker. DOR was similarly shown to reverse human Pgp-mediated MDR (Gao et al., 2010). The modulation of verapamil-stimulated murine Pgp ATPase activity by ML revealed a similar potency for most avermectins: IVM, DOR, EPR and ABA. SEL and MOX were about five and ten times less efficient than others, respectively (Lespine et al., 2007). This suggested that the presence of one or more sugar moieties in the structure of ML can influence the affinity to Pgp (see Suppl. Fig. S1, Manuscript n°2: David et al., 2016).

Among other anthelmintics, some were also shown to interact with Pgp, but with less potency than IVM. Amino-BZ was found to interact with Hsa-Pgp (Nare et al., 1994), and TCZ was widely proven to be transported by *F. hepatica* Pgps, in addition to *Ovis aries* (Oar)-Pgps (Lifschitz et al., 2009; Meaney et al., 2013; Mottier et al., 2006; Savage et al., 2013a, b, 2014; Wilkinson et al., 2012). TCZ and CLO also decreased RHO transport activity of Mmu-MDR1a with an E_{max} 8 and 4 times higher than IVM respectively (Dupuy et al., 2010). In this study however, TBZ, LEV and most of other anthemintics tested did not modify RHO transport. LEV however increased the bioavailibilty of IVM, without changing its efficacy, in Humans treated against *O. volvulus* (Awadzi et al., 2004). In contrast, CLO used in combination with IVM did not modify its pharmacokinetics compared to its administration alone in cattle (Cromie et al., 2006). Finally, EMD was shown during its synthesis to interact with Pgp, and its penetration in brain was recently shown to be increased in ABCB1a deficient mice compared to control mice (Elmshauser et al., 2014).

Interestingly, IVM can also interact with MRPs and BCRP in human, although with lower affinities than with Pgp (Jani et al., 2010; Lespine et al., 2006b). IVM and MOX were also described as substrates of sheep and mice BCRPs, respectively (Perez et al., 2009; Real et al., 2011). Although BCRP is critical at the breast level in elimination of ML through milk, it cannot compensate for Pgp deficiency at the blood-brain barrier, and MRP function appears

limited in the brain (Borst et al., 1999; Cisternino et al., 2004). Thus, Pgp plays a major role at this level in mammals and its overall presence in various tissues is critical for detoxification, due to the higher affinity of ML for this protein compared to other MDR transporters.

b) Anthelmintics transport by *C. elegans* ABC MDR transporters

The first Pgp identified in *C. elegans* was Pgp2 in 1992 (Lincke et al., 1992). 15 homologues of the single Pgp protein expressed in humans have been found in this nematode, whose encoding genes appear tandemly duplicated in clusters of two or four closely related genes (Zhao et al., 2004). Including MRPs and BCRPs, 60 ABC transporters have now been described in *C. elegans*. As for the human ABC transporters, they have been classified into different subfamilies: ABCA to ABCH, depending on the structure of the protein (Sheps et al., 2004). Each ABCB/Pgp protein is expressed in *C. elegans* at different developmental stages and in specific tissues in the adult worm (**Fig. 16**) (Zhao et al., 2004). By analyzing their expression pattern, the authors showed that throughout the development cycle of *C. elegans*, Pgps are found almost in every tissue of the worm. For instance, excretory cells express Pgp4 and Pgp12, whereas Pgp2 and Pgp14 are mainly found in the pharynx. At the adult stage, each transporter protects *C. elegans* against specific toxic compounds with redundant functions for some proteins (Lespine et al., 2012). For example, a deletion experiment showed that both MDR1 and Pgp1, expressed in different tissues in *C.elegans*, are involved in resistance to heavy metals (Broeks et al., 1996). Indeed, nematodes who had both of these genes deleted were found to be hypersensitive to cadmium and arsenite.

The overexpression of several ABC transporters has been found to be linked to resistance to ML in this model organism. In a study, IVM resistant worms were selected by stepwise exposure to this drug, leading to cross- resistance with MOX, LEV and pyrantel, and the overexpression of the *mrp-1* and *pgp-1* genes was detected in the resistant strains (James and Davey, 2009). The effect of the inhibition of transcription of various ABC transporter genes by RNAi was studied in IVM susceptible and resistant *C. elegans* (Yan et al., 2012). They found that transcription inhibition of no single gene could efficiently increase IVM

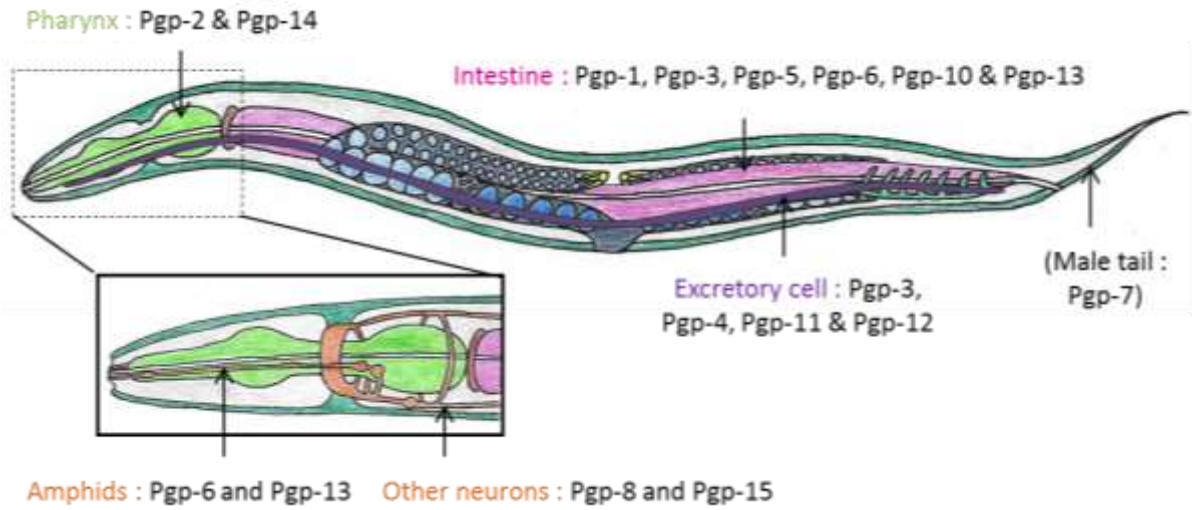


Figure 16. Expression pattern of *C. elegans* Pgps in the hermaphrodite adult. The only exception expressed in the male is indicated in parentheses. Schematic representation constructed from (Zhao et al., 2004).

susceptibility in these worms. However, depending on which genes were down-regulated simultaneously, motility, pharyngeal pumping and/or egg production were differently impaired, *mrp-1* and *haf-2* having the most important effects.

Lespine et al. (2007) also showed that the transporter activity of Pgps is inhibited by IVM and ML by competition, thus confirming their substrate status that explains the importance of Pgps in ML resistance. Very recently, both the expression of various *pgp* genes and phenotypes (motility and pharyngeal pumping) were compared between IVM or MOX sensitive and resistant strains of *C. elegans* (Ardelli and Prichard, 2013; Bygarski et al., 2014). This assessment has highlighted the importance of several Pgps in the protection of this organism against IVM and MOX, each of them expressed in different tissues. The loss of function of different Pgps in *C. elegans* was also shown to result in an increasing impairment of development of these nematodes exposed to rising concentrations of IVM, in comparison to wild-type worms (Janssen et al., 2013). Finally, IVM or MOX-selected strains of *C. elegans* showed an increased susceptibility to the three ML IVM, MOX and EPR after exposure with the ABC transporter substrate verapamil (Menez et al., 2016). This accounted for the critical role of Pgps for detoxification of ML in *C. elegans*.

c) Anthelmintics transport by Pgps of *H.contortus*

The first gene coding for *H. contortus* Pgps, called *hcpgp-1* on the one hand and *pgp-a* on the other hand was simultaneously identified in 1998 by two teams (Kwa et al., 1998; Xu et al., 1998). The latter showed that this gene was expressed more in IVM resistant *H. contortus* than in the susceptible worms. The same year, the selection of specific Pgp alleles was suggested in strains showing IVM and MOX resistance (Blackhall et al., 1998).

Then, other studies indicated that resistance to ML is tightly linked to the expression of P-glycoproteins in this organism using various techniques such as southern blots (Sangster et al., 1999) or tests of MDR-reversing agents (Molento and Prichard, 1999). These results were then confirmed by detecting overexpression of several Pgps *in vivo* after exposure of *H. contortus* to IVM or MOX (Prichard and Roulet, 2007). More recently, the expression of different genes was assessed in an *H. contortus* strain resistant to BZs, levamisole and IVM. Pgp-2 and Pgp-9 mRNA expression was increased compared to susceptible worms as previously shown; however, Pgp1 was expressed to a lesser extent than in the control (Williamson et al., 2011). The interaction of various anthelmintics with Hco-Pgps was also tested by monitoring Rho123 efflux from eggs, previously to the identification of each of the homologs expressed in the parasite. In this study, all ML except IVM were shown to increase Rho efflux from eggs, as well as LEV (Kerboeuf and Guegnard, 2011). TBZ however failed to have any effect.

In vivo, a pharmacokinetic study in lambs infected with resistant *H. contortus* has allowed the determination of the effect of three ML (IVM, MOX, abamectin) several days post-administration (Lloberas et al., 2013). This study showed that two days after exposure, all three drugs are found at the same concentration in plasma, but MOX persists longer (up to 13 days) than the two other drugs (8 days). The abomasal concentration of IVM was found to be greater if administered by the enteral route than subcutaneously. Moreover, three days post-treatment, IVM concentration was higher in *H. contortus* recovered from infected lambs compared to the other drugs. However, the mass of *H. contortus* recovered was the lowest, and the efficacy of treatment determined by fecal egg count reduction test (FECRT) was the highest after MOX administration. Pgp2 expression levels were higher after IVM treatment in the resistant strain compared to the wild-type. In conclusion, this study found that MOX was the most efficient drug (86%) against IVM resistant *H. contortus*, compared to abamectin (39.7%) and IVM (20.1%), showing the important role of Pgps *in vivo* against the effects of ML in this worm.

The localization of the Pgps in *H.contortus* was unknown until 2002, when researchers analyzed the RNA distribution of the *pgp* genes previously sequenced by *in situ* hybridization (Smith and Prichard, 2002). They could detect a high RNA level in the pharynx of the worm, in the intestine region close to the pharynx and along its luminal edge. There was no staining in the muscles, nor in the hypodermis or the cuticle. However, another study detected the presence of Pgp by using an antibody in the cuticle of the adult worm as well as in the larvae and in the egg shells (Riou et al., 2005). Today, the exact localization of P-pgs in *H. contortus* is still unclear, except for Hco-Pgp-2 expressed in the pharynx and intestine, and Hco-Pgp9.1 found in the uterus (Godoy et al., 2016; Godoy et al., 2015b).

Indeed, until recently, the genome of *H. contortus* was not completely assembled and annotated, so that the exact number and sequence of *pgp* genes in this organism was not known. However, identifying homologies between the genome of *H. contortus* and the sequence of *C. elegans* mRNA transcripts can give indications about the *pgp*s of the parasite as these worms are closely related. Such analyses have indicated that some of the *pgp* genes present in *C. elegans* have duplicated, while others have been lost in *H. contortus* (Laing et al., 2013). A recent study had thus identified nine *pgp*s by bioinformatic analysis, but the level of expression of their transcripts did not show any significant differences between IVM sensitive and resistant strains, possibly because their selection had been rapid, happening only over three generations in this study (Williamson and Wolstenholme, 2012). Finally, the completely assembled genome of *H. contortus*, published in 2013, shows that *pgp-5, 6, 7* and *8* of *C. elegans* are not found in the parasitic species (Laing et al., 2013). On the other hand, *C. elegans pgp-3* and *4* have arisen after duplication events and correspond to the single gene *Hco-pgp-3*. Similarly, *Cel-pgp-12, 13* and *14* correspond to *Hco-pgp-13* only. On the opposite, the single copy *pgp-9* in *C. elegans* matches two copies in *H. contortus*.

These data then allowed further identification of the mechanisms of resistance associated to specific Pgp expression in this nematode. Indeed, *in vitro* techniques implemented on mammalian Pgps have allowed in the last two years the functional characterization of specific Pgps of *H. contortus*: Pgp2, Pgp9.1 and Pgp16 (Godoy et al., 2015a, 2016; Godoy et al., 2015b). All were shown to transport the anthelmintic drugs IVM and MOX.

These new types of studies made possible by the recent annotations of *H. contortus* genome will pave the way for new treatments and allow overcoming the issue of the decreasing response to currently used ML.

3. Overcoming ML resistance in parasitic nematodes

a) Different possible strategies

Glutamate was found to be an allosteric modulator of IVM and moxidectin that increases their activity by potentiating their binding to HcGluClα receptors (Forrester et al., 2002). A GluCl agonist, ibotenate, was then shown to increase the efficacy of IVM by measuring the worm number in infected gerbils (Forrester et al., 2004). Thus, it may be possible to increase an anthelmintic treatment efficacy by co-injecting an allosteric modulator of ML.

Several studies have shown that the overexpression of ABC MDR transporters in parasites seem to be inducible by ML (Lespine et al., 2012). This effect might be decreased by down-regulating the transporters by gene knock-out or siRNAs (**Fig. 16**). But, since various Pgps and other ABC transporters have been found to be involved in ML resistance, it would be difficult to target them all simultaneously. The task would be easier if a single regulator was activated by the ML and was responsible for the upregulation of different Pgps, but such molecule has not yet been identified. It is known, however, that Pgp expression can be induced in mammals by nuclear receptors when activated by their ligand (Staudinger et al., 2003). If similar mechanisms exist in nematodes, as suggested by the discovery of xenosensors in *C. elegans* that are involved in resistance to colchicine and chloroquine, antagonists of these nuclear hormone receptors could be useful tools to fight AH resistance. Targeting this regulatory pathway however, requires the identification of such regulators in parasitic species, as they are still unknown in *H. contortus*.

Non-specific inhibitors have also been proposed as a method to reduce ML resistance. For instance, disturbing the membrane integrity around MDR transporters with lipid excipients or modulating the amount of cholesterol present in the membrane and forming rafts have been investigated (**Fig. 17**) (Riou et al., 2003; Seelig and Gerebtzoff, 2006). These different possibilities have proven to modify the activity of ABC transporters in nematodes, but in the case of cholesterol, its overloading decreased resistance, contrary to what was expected. Multidrug resistant tumor cells have also been shown to be more sensitive to oxidative stress than drug-sensitive cells (**Fig. 17**) (Karwatsky et al., 2003). The higher ATP requirement of overexpressed Pgps is responsible for drug resistance in these cells, which leads to apoptosis more easily than in control cells. Similarly, resistant strains of parasites can

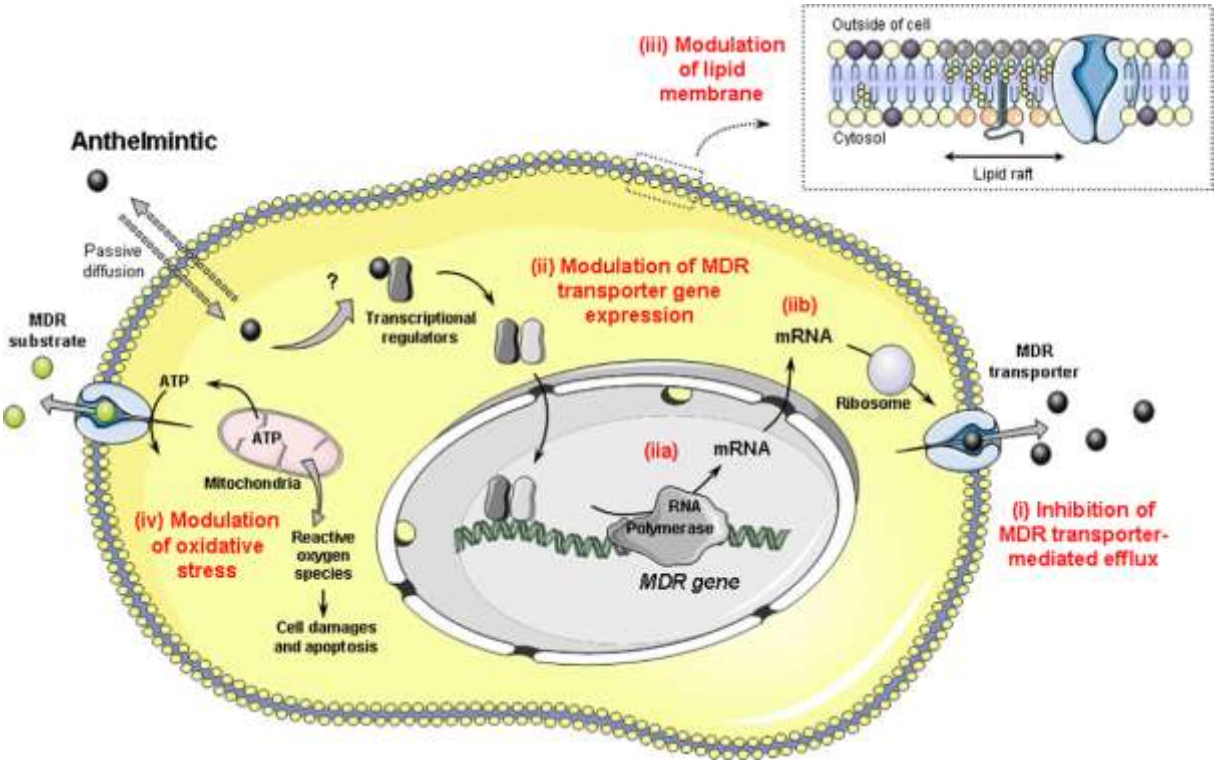


Figure 17. Putative targets for reversing multidrug transporter-mediated resistance. Different strategies could be considered to increase drug concentration in the target cell in the case of MDR transporter-mediated resistance and to overcome this resistance. (i) Inhibition of MDR transporter-mediated efflux by using MDR inhibitors. (ii) Modulation of MDR transporter gene expression at transcriptional level through an inhibition of transcriptional regulator activity (iia); or at post-transcriptional level through a decrease of the MDR mRNA stability in the cell (iib). (iii) Modulation of lipid composition of the cell membrane harbouring the MDR transporters with surfactants or cholesterol depletion to decrease the basal activity of MDR transporters. (iv) Another side mechanism associated with the presence of MDR substrates is the modulation of oxidative stress with an increase of reactive oxygen species (ROS) production and cell damage resulting in apoptosis induction. From (Lespine et al., 2012).

be more easily affected by apoptosis-inducing agents than sensitive parasites, but the compound would have to be specific enough to nematodes, not to be deleterious for the host.

A particularly promising strategy to overcome ML resistance is the use of ABC transporter inhibitors (**Fig. 17**). Reducing the drug efflux by these transporters would lengthen the exposure time of the targeted cells to toxic compounds, thus increasing their efficacy (Lespine et al., 2008). However, the inhibitors need, again, to be as specific as possible to the parasite transporters in order to avoid cytotoxicity in mammalian cells. Firstly, avermectins themselves have been investigated as MDR-reversing agents, due to their ability to competitively inhibit the transport of other compounds by Pgps, but their possible neurotoxic effect in the host prevents their extensive use (Pouliot et al., 1997). Since then, a wide number

of compounds have been tested, such as verapamil or quercetin, first *in vitro*, to assess their ability to decrease Pgp transport in cultured cells (Dupuy et al., 2001), and then, *in vivo*, by determining the bio availability of co-administered ML (Dupuy et al., 2003; Molento et al., 2004). Loperamide and ketoconazole were thus shown to impact the pharmacokinetics of IVM or MOX in different animal species (Alvinerie et al., 2008; Hugnet et al., 2007; Lifschitz et al., 2002; Lifschitz et al., 2004). Verapamil was also found to increase the efficiency of ML in jirds infected with resistant *H. contortus* (Molento and Prichard, 1999). Other studies have shown the possibility of restoring drug efficacy in *H. contortus* using verapamil or other inhibitors of Pgp function (Ardelli and Prichard, 2013). For example, *H. contortus* selected with IVM in jirds showed an increased susceptibility to IVM and MOX after verapamil treatment. The FECRT result of lambs infected with a strain of *H. contortus* resistant to IVM was higher after treatment with IVM and loperamide than after administration of IVM alone (Lifschitz et al., 2010). Another team also demonstrated the increased susceptibility to IVM of both sensitive and resistant L1 larvae of *H. contortus* when co-administered with the previously cited inhibitors, valspodar, ketoconazole or pluronic P85 (Bartley et al., 2009); however, this capacity has not yet been investigated for all the inhibitors that have been successfully tested so far, like ketoconazole, itraconazole, or other structurally unrelated compounds (Lespine et al., 2008). Moreover, the first clinical trials of such MDR inhibitors in cancer therapy remain unsuccessful, mainly due to their secondary effects (Nobili et al., 2006). Indeed, if the reversing agents also show affinity for host transporters, they will inhibit the export of endogenous compounds and impair cell homeostasis causing toxicity in the host. As a result, the critical requirement for an inhibitor is to target the parasite transporter as strongly as possible without blocking the host transporter in order to avoid toxic effects in mammals.

b) Specific resistance reversal in *C. elegans* and *H. contortus*

To reach high-specificity with a given compound, it is necessary to study the affinity of interaction between this drug and the transporter at a molecular level, as already done with mammalian Pgps, MRPs and BCRPs (Lespine et al., 2009).

However, these studies, which usually assess the level of transport, the ATPase activity of the targeted protein, and cellular cytotoxicity, are only beginning on Pgps of nematodes. Indeed, because of the difficult constitutive expression of worm transporters in cells endogenously expressing only a few of these proteins, this type of experiment is still at its nascent stage with *C. elegans* Pgps. Moreover, due to its life cycle requiring a host to develop

to the adult stage, proteins of parasitic species are even harder to study. The genome of *H. contortus* being only recently fully annotated (Laing et al., 2013), no drug-target interaction research has yet been published on the proteins of this worm.

As many more Pgps are expressed in nematodes compared to humans, it is only possible to identify the molecular interactions and pharmacological properties of a compound relative to each transporter individually. Each protein being expressed in different tissues of the nematodes, some of them may have more important roles in resistance to AHs than others. Thus, it would be interesting to identify Pgps having the most critical roles in resistance and study these one in priority. A recent study analyzed the physiological effect of IVM on *C. elegans* strains knocked out for different *pgp* genes. By following the velocity of these strains with or without exposure to IVM, the authors showed that *pgp-2, 5, 6, 13, 7* and *12* knock outs were the most sensitive to IVM. Indeed, treatment with this drug led to an increase of velocity followed by an inhibition of the movement of these strains more rapidly than in the sensitive wild-type (Ardelli and Prichard, 2013). Interestingly, two of them, *pgp-6* and *pgp-13*, are expressed in the amphids (Zhao et al., 2004), where HcGluClα3A subunit of the GluCl receptor might be expressed (Portillo et al., 2003). HcGluClα3B subunit, on the other hand, has been detected in pharyngeal neurons, and the lateral and ventral nerve cords. Both subunits where also found expressed in motor neurons commissures (Portillo et al., 2003). A regulation of the exposure of GluCl receptors to ML in the neighboring neuronal and pharyngeal regions might thus be important to avoid paralysis and death of the parasite by starvation.

Freeman et al. (2003) have suggested the importance of amphids in IVM sensitivity. They investigated the hypothesis that these neurons might be another point of entry of the drug, as *C. elegans* larvae had been found to be paralyzed after IVM treatment even when their pharyngeal pumping function was inhibited (Smith and Campbell, 1996). They found that in a resistant worm, the neurons were shorter and did not approach the external environment as closely as they did in two different susceptible strains, indicating the possible importance of the amphids in the susceptibility or resistance to ML. Furthermore, Dent et al., (2000) showed that *C. elegans* worms with mutation of the Dyf (dye filling defective) gene *osm-1* were more resistant to IVM than wild-type *C. elegans*. Dyf mutants are phenotypically deficient in the uptake of fluorescent dyes from the environment by the amphids (Dent et al., 2000). In their paper, the authors hypothesize that the resistance to IVM conferred by mutations of Dyf genes is explained by a reduced permeability to the drug in *C. elegans*. There are more than 20 genes that give a dye filling defect when mutated in *C. elegans*. DYF-7 is a factor secreted by

amphidial neurons to anchor dendritic tips during cell migration, thus ensuring that these sensory neurons reach the external environment (Heiman and Shaham, 2009). Urdaneta-Marquez et al. (2014) showed that a loss-of-function allele of *Cel-dyf-7* causes amphid neuron morphology and dye filling defects together with increased IVM resistance. Moreover, in their study, *Hco-dyf-7* was also genetically linked to ML resistance and dye-filling defects in strains of *H. contortus* resistant to IVM and MOX. However, in *C. elegans*, not all Dyf mutants that are resistant to IVM show an impaired dye filling phenotype, for example in the case of *osm-5* mutant (Dent et al., 2000). This indicates that another mechanism, such as the transport of IVM out of amphidial neurons through Pgps, could play a role in the importance of this organ in ML resistance in nematodes.

Hypothesis and objectives

Very little is known on the role of individual Pgps in the parasitic nematode. Gaining insights into the function of Pgps in *H. contortus* relies essentially on the determination of their substrate profile. As several studies indicated the importance of amphidial neurons in protection of the nervous system of *H. contortus* from AH drugs (Dent et al., 2000; Freeman et al., 2003; Menez et al., 2016; Urdaneta-Marquez et al., 2014), a particular interest will be given to Pgps potentially expressed in these structures. In *C. elegans*, two *pgp* genes were found expressed in the amphids, *pgp-6* and *pgp-13* (Zhao et al., 2004). As *pgp-6* has no ortholog in *H. contortus*, we focus on the *pgp-13* ortholog in this parasite and we study the function of its product Hco-Pgp-13.

- Our first objective was to study the binding properties of various substrates onto a nematode Pgp, Cel-Pgp-1, as a model. Cel-Pgp-1 crystal (PDB : 4F4C) was indeed the first structure of a nematode Pgp available (Jin et al., 2012) and can be used as a template to model by homology the 3D structure of Hco-Pgp-13, in order to study the substrate profile of this transporter, using *in silico* docking calculations. In a first step, several compounds with the capacity to stimulate the basal ATPase activity of Cel-Pgp-1 *in vitro* were studied *in silico*, and results of docking calculations were compared to these experimental data. This validation of the strength of the *in silico* approach on a 3D structure found experimentally is described in Part I, A. Then, based on the high prediction capacity for drug affinity of such docking model, we investigated for the first time the direct interaction of ML with a nematode Pgp, which is exposed in Part I, B.

- The second objective was to determine the nature of Hco-Pgp-13 substrates, in order to gain insight to its role *in vivo* and investigate the possibility of its interaction with ML, as the first step required for their transport. First, the cDNA sequence of Hco-Pgp-13 was experimentally validated from reverse transcribed *H. contortus* mRNA, based on the sequence predicted from large-scale sequencing (Laing et al., 2013), and phylogenetic analyses were performed to better characterize this sequence, in Part II, A. Because the sequence of Cel-Pgp-1 shows close homology to Hco-Pgp-13, 3D models were constructed based on the PDB template 4F4C. The computational analysis of drug docking was then performed on Hco-Pgp-13 models. This was combined with *in vitro* assays, mainly ATPase activity assays in a *Pichia*

pastoris heterologous expression system, to further characterize the function of Hco-Pgp-13. This data is reported in the Part II, A and C. We have finally localized the protein product of closest ortholog of *C. elegans pgp-13* in various stages of *H. contortus* life cycle, in order to better characterize its importance in the living parasite in Part II, A and B.

This work provides new information on the function of Cel-Pgp-1 as a mutlidrug transporter and a relevant and predictive tool to study the function of Pgp orthologs in parasitic nematode species. In addition, specific information on Hco-pgp-13 on localization in *H. contortus* adults and larvae are presented together with the transport function of Pgp-13 that will contribute to unravel the potential involvement of Hco-Pgps in drug transport.

Experimental Work

PART I

In silico functional characterization of Cel-Pgp-1

The first main objective was to check that *in silico* docking calculations performed with Cel-Pgp-1 crystal (PDB: 4F4C) would be a reliable strategy for the study of the interaction of a nematode ABC transporter with putative ligands. As this 3.4 Å resolution crystal was the first of a nematode Pgp to be released, it could then be used as a template to construct a 3D homology model of Hco-Pgp-13 and perform similar docking calculations on this structure.

We thus started our *in sillico* docking approach with the docking of compounds characterized as substrates of Cel-Pgp-1 by their activation of basal ATPase activity *in vitro* by Jin et al. (2012). This allowed correlating our *in silico* prediction of energy binding, representative for the affinity of each tested compound to Cel-Pgp-1 binding pocket, with experimental data reavealing or not their interaction with Cel-Pgp-1, and thus giving indications about their possible affinity *in vitro*. This revealed a very good correlation between the two approaches and help shedding light on the molecular properties leading to the substrate multispecificity of such transporter. It also allowed the description for the first time of the wide binding pocket of a nematode Pgp, divided in two sub-domains linked by a constriction zone.

We then used Cel-Pgp-1 crystal structure to test the interaction *by in silico* docking of various anthelmintics, including a series of seven ML and seven AH of different classes. Strikingly, all ML drugs bound to the same sub-domain of Cel-Pgp-1, with similar high affinities. All ML presenting a disaccharide bound with higher affinity than those that harbor only one saccharide or none. Other AH bound to various locations, revealing several of them as being potential competitors of ML for binding to Cel-Pgp-1.

In addition, this data showed that the docking experiments on the template 4F4C gave precise and reproducible results. Thus, a model of Hco-Pgp-13 based on the crystal of *C. elegans* could be used for docking calculations in order to investigate its substrate profile.

A. Manuscript n°1 in preparation:

Modeling multispecific drug recognition by *Caenorhabditis elegans* P-glycoprotein 1

Marion David[a,b], Stéphane Orlowski[c], Shaima Hashem[c], Roger K. Prichard[b], François André[c*] and Anne Lespine[a*] .

[a] Toxalim (Research Centre in Food Toxicology), Université de Toulouse, INRA, ENVT, INP-Purpan, UPS, Toulouse, France

[b] Institute of Parasitology, McGill University, Sainte-Anne-De-Bellevue, Canada;

[c] CEA, IBiTec-Saclay, SB2SM and UMR9198 CNRS, I2BC, Université Paris-Saclay; Gif-sur-Yvette, France.

[*]Corresponding authors:

François André: francois.andre@cea.fr

Tel: +33 (0)1 69 08 44 32

Fax: +33 (0)1 69 08 87 17

Address: CEA, IBiTec-Saclay, SB2SM and UMR9198 CNRS, I2BC

91191 Gif-sur-Yvette cedex, France.

Anne Lespine: anne.lespine@toulouse.inra.fr

Tel: +33 (0)5 82 06 63 52

Fax: +33 (0)5 61 28 53 10

Address: INRA, UMR 1331 TOXALIM - Equipe E6-TMR: Transporteurs Membranaires & Résistance ; 180, chemin de Tournefeuille - BP 93173 ; F-31027 TOULOUSE Cedex 3, France

Abstract

Some ABC transporters present multispecific recognition capacities, which allows them to be involved in multidrug resistance in cancer cells and in phase III cell detoxication in mammals. In other organisms, these proteins also participate in pleiotropic antibiotic resistances in bacteria, and in multiple resistance, against anthelmintics in nematodes, and against other antiparasitic drugs in various pathogenic parasites. In spite of this pharmacological importance, the molecular mechanisms of such multispecific drug recognition are still poorly understood. In this study, we took benefit from the crystal structure of the membrane ABC protein, P-glycoprotein 1 of *Caenorhabditis elegans* (Cel-Pgp-1, PDB 4F4C), recently resolved to a good resolution (3.4 Å). We undertook an *in silico* study for modelling multiple drug binding to this full-size ABCB sub-family member of a model nematode that presents fair homology to human Pgp (ABC B1). Using a semi-flexible docking strategy, we first determined the binding modes of a series of molecules that have been previously assayed for their effects on Cel-Pgp-1 ATPase activity. We found an excellent correlation between these *in vitro* enzymological data and our *in silico* calculated binding energies (Eb), which fully validated our *in silico* approach. These first data collected on the validation set composed of 7 drugs allowed pointing out the central role of the large "inner chamber" of the protein, in its open inward-facing conformation, for ligand recognition. Further analysis of the *in vitro*/*in silico* correlation showed the necessity to consider the initial step of membrane partition of the tested ligands before they directly interact with the protein. Finally, for a total of 18 *in silico* tested molecules presenting broad chemical structures, we found that 14 were "good" ligands (Eb in the range -7 to -17 kcal/mol), 1 was a "weak" ligand and 3 were non-ligands, demonstrating the multidrug property of this ABC transporter. Its multispecific capacity relies on a rich palette of putative interacting residues. By lining the inner chamber, these amino-acids offer a large number of possible combinations for the binding of various structurally unrelated molecules (of size ranging from 300 to 1250 Da) to Cel-Pgp-1. Such a continuum of specific sub-sites within this large multispecific binding domain gives the opportunity to the smaller ligands to display two relevant binding sites, either partially overlapping and hence "dual", or separated and hence leading to a stoichiometry of 2. The binding sites of small ligands then show a propensity to be altered when submitted to slight chemical modifications. Alternatively, the larger ligands are more constrained by the shape of the inner chamber walls: the more flexible ones can well accommodate the largest part of the chamber; whereas the more rigid ones can fit the chamber walls either very well or not at all. Finally, our multidrug binding model provides a new vision of ligand/protein interactions that

is different from the classical "key-in-the-lock" dogma prevailing for specific ligand-receptor interactions. This model is expected to improve strategies for rational drug design aimed at overcoming various multidrug resistance phenotypes in which Pgp homologs are involved.

Keywords

ABC transporters; P-glycoprotein; *in silico* docking; multispecific recognition; multidrug resistance; *Caenorhabditis elegans.*

Abbreviations

ABC = ATP-binding cassette; ACD = actinomycin D; CAM = calcein-AM; CCH = colchicine; Cel = *Caenorhabditis elegans*; Cgr = *Cicretulusgriseus;* CLS = cholesterol ; CSP = cyclosporin A; DNR = daunorubicin; DPM = dipyridamole; DXR = doxorubicin; Hco = *Haemonchus contortus*; Hsa = *Homo sapiens*; HST = hoechst 33342; KTC = ketoconazole; MDR = multidrug resistance; ML = macrocyclic lactone(s); Mmu = *Mus musculus*; NBD = nucleotide binding domain; PCT = paclitaxel; PDB = Protein Data Bank; Pgp = P-glycoprotein; PRG = progesterone; RHO = rhodamine 123; RMSD = root mean square deviation; TMD = transmembrane domain; VBL = vinblastine; VCR = vincristine; VLN = valinomycin; VRP = verapamil; VSP = valspodar.

I. INTRODUCTION

The superfamily of ATP-binding cassette (ABC) proteins comprises a huge number of members, which are expressed in prokaryotes as well as eukaryotes, in all living kingdoms from plants to yeasts, parasites and animals, including mammals (Higgins, 1992). Most of them are membrane proteins working as active transporters, with few exceptions, which use ATP hydrolysis at their nucleotide binding domains (NBD) to energize the transmembrane translocation of various substrates. Their respective physiological roles are thus dependent both on the nature of their handled substrates: peptides, sugars, ions, lipids, etc, and on the transport direction, importers being found only in bacteria. Among all of them, a rather small class is particularly interesting due to its remarkable property of multispecific substrate recognition, and eventual transport. This is indeed a very rare capability exhibited by enzymes or transporters that contradicts the classical dogma of ligand-receptor recognition ("the key in the lock"), and which makes them especially adapted for the efflux of various exogenous

compounds. In humans and mammals, they belong essentially to the sub-families B and C, and the three ones that are currently acknowledged to play a crucial role in cellular detoxification processes, constituting to the so-called phase III system, are ABCB1, C1 and G2 (Eckford and Sharom, 2009). Due to its oldest discovery and the mass of experimental data accumulated about it, ABCB1, or P-glycoprotein (P-gp), appears paradigmatic for multispecific ABC drug transporters (Ambudkar et al., 1999). Actually, these three multidrug ABC transporters have an important pharmacological impact because they are responsible for transmembrane transport of a very broad spectrum of drugs of various pharmacological interests, contributing to their pharmacokinetic characteristics: intestinal absorption, biliary and urinary elimination, and disposition at the level of blood-tissue barriers, typically leading to medicine interactions (Sarkadi et al., 2006). Due to their (over)expression, either innate or acquired, by some tumor cells, they are also often responsible for multidrug resistance (MDR) phenotypes against anticancer chemotherapy by a number of cytotoxic drugs, leading to treatment failures (Szakacs et al., 2006).

Such an obvious toxico-pharmacological importance requires a fine understanding of the underlying molecular mechanisms harbored by these multidrug ABC transporters, in order to be in position to rationally control their activity during drug (or toxic) expositions of patients, for example by predicting, short-circuiting, and/or inhibiting their transport function. However, these precise molecular mechanisms still remain rather obscure, ill-described and sometimes under debate. A number of enzymological studies have nevertheless addressed for Pgp this unique property of multispecific recognition, and they essentially led to three alternative functional models: (i) a unique but fully adaptable binding site; (ii) several binding sites, selective for each of the various chemical families of handled substrates; (iii) a large binding pocket or domain comprising various pluri-potent sub-sites that more or less overlap (Sharom, 2006). Some reports, supported by various experimental approaches, have indicated the fairly likelihood of multiple binding sites, but they remain in discordance about their number and selectivity (Boer et al., 1996; Shapiro and Ling, 1997; Orlowski and Garrigos, 1999; Martin et al., 2000; Wang et al., 2000). With this background, our team has established arguments favoring the third possibility by building a double pharmacophoric model of chinese hamster, *Cicretulusgriseus* (Cgr-) Pgp based on the analysis of the mutual relationships between a set of Pgp substrates modulating its ATPase activity, and correlating them with the 3D alignments of the hydrophilic and hydrophobic recognitions elements of these substrates. Furthermore, a size effect for large ligands recognizing one pharmacophore and competing with ligands of the other

one strongly indicated the closeness of the two corresponding binding sites (Garrigues et al., 2002). However, it is difficult to further progress toward a refined binding model of Hsa-Pgp for several reasons. In one hand, a ligand-based approach is hampered by the fact that a classical quantitative structure-activity relationships (QSAR) analysis is not applicable since there can be different binding sites available for ligands of related chemical structures. In the other hand, a structure-based approach suffers from the paucity of reliable high resolution crystal structures since the few presently available structures of murine Pgp reveal some discrepancies (Aller et al., 2009; Li et al., 2013; Ward et al., 2013).

Recently, Cel-Pgp-1 crystal structure has been released at the resolution of 3.4 Å (PDB 4F4C), in the absence of nucleotides (Jin et al., 2012) and hence in the open inward-facing conformation supposed to be capable of chelating substrates with high affinity in the donor compartment (i.e. cytosol or cytosolic membrane leaflet) for their translocation. This nematode ABC protein presents a high sequence similarity with Human (Hsa-) Pgp, of about 63 % (58% if restricted to TMDs) and an identity of 46% (37% TMD-restricted), according to BlastP. Plus, Cel-Pgp-1 has previously been reported for being involved in anthelmintics resistance, mainly regarding macrocyclic lactones (Janssen et al., 2013), and the functional characterization of the purified protein, based on ATPase activity modulation assays accompanying its structural study, further showed that it appeared to display some traits of multispecific recognition for a set of molecules of diverse chemical structures (Jin et al., 2012). This similarity with Hsa-Pgp was then interesting for gaining insight into the molecular mechanisms of multidrug recognition by Pgp in general. This crystal structure gave the opportunity to launch an *in silico* study of docking of various potential ligands, chosen among the known mammalian Pgp transport substrates, in order to test their respective binding modes. This also allowed comparing the residues of Cel-Pgp-1 found to interact with substrates to what is known for mammalian Pgp, as some residues have already been reported to be involved in (multi)drug recognition, in particular considering the so-called “hotspot residues for drug binding” (Shilling et al., 2006). Finally, this study was aimed at shedding some light on the various molecular characteristics of multidrug recognition by a multispecific full-size ABC transporter of the B sub-family. We could thus investigate the binding site location(s) in the protein, multiplicity of these sites (indicating the binding stoichiometry) and the mutual relationships between them. We could also calculate the binding energies of ligands (taken as the evaluation of their apparent affinities, without taking into account the initial membrane partition step), evaluate the influence of the

ligand molecular size, and investigate the effects of slight chemical structure modifications on the binding.

As a first, obviously necessary step, we checked for docking calculations a "validation set" of 7 molecules tested by Jin et al., (2012) for their Cel-Pgp-1 ATPase response (6 positive ones and a negative one), in order to correlate our *in silico* data with *in vitro* enzymological published data. After full satisfaction of this validation step, we performed further *in silico* docking analyses of 11 other molecules, and found that they all bind within the "inner chamber" of the protein with a large range of calculated binding energies, encompassing various continuously connected sub-sites to take benefit of a large palette of anchoring residues offered by the protein. This allows a great number of molecules of large chemical diversity to find a solution for binding, hence explaining the capacity of multispecific recognition without requiring an inherent large flexibility of the receptor protein.

II. COMPUTATIONAL METHODS

1. Structure of Cel-Pgp-1

The Cel-Pgp-1 X-ray structure, determined at a resolution of 3.4 Å (PDB code 4F4C) (Jin et al., 2012), was used in all docking calculations. The whole chain A was taken into account, with the exception of two detergent molecules (undecyl 4-O-alpha-D-glucopyranosyl-1-thio-beta-D-glucopyranoside, PDB entry name 0SA), bound in the inner chamber, and that have been removed for grid maps calculations. The 4F4C structure includes the full glycosylated Cel-Pgp-1 sequence (1321 amino acids), but N-terminal (M1-R3) and C-terminal (E1307-K1321) segments are missing in the structure, as well as a short segment (A52-E54) located in an extended loop of the first TMD domain, and a 49-residues segment (K666-E715) belonging to the linker region. Interestingly, an additional helix-turn-helix motif (Q9-V32) is present in the N-terminal domain, a structural feature that has not been observed in other Pgp structures released in the PDB.

The atomic coordinates PDB file was then converted into a PDBQT file by AutoDock Tools 4 (Morris et al., 2009) for docking calculations. PyMOL (The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC) was used as visualization tools for various tasks (3D alignments, ligands and hotspots location, grid box positioning for AutoGrid 4) and for structure rendering in figures.

2. Preparation and conformational analysis of ligands

The molecular structures of ligands were extracted from chemspider, drugbank (DB) or pubchem (CID) (Suppl. Fig. S1, S4 and S8), depending on the availability of structures. Each molecular structure was carefully scrutinized for chirality, and sometimes corrected when inconsistencies were found in the literature.

In the semi-flexible mode, the ligand is handled as flexible around all the rotatable bonds. However, the conformational space of the ligand can be poorly explored when it contains ring structures as in avermectins, since AutoDock does not consider single bonds in non-aromatic cycles as rotatable bonds. To overcome this limitation, in order to better sample the initial conformational space accessible to the ligand, we generated for each ring-containing compound 10 different low energy conformations. For this, we used Marvin Sketch and the minimization under the MMFF94 force field provided in Marvin Suite (https://www.chemaxon.com/products/marvin/marvinsketch/). The diversity of the 10 lowest energy conformers was evaluated by their pairwise root mean square deviations (RMSD) after superimposition under PyMOL. For each obtained cluster of close conformers, a representative one, defined as the center of the group according to the calculated RMSDs, was selected. One to five different conformers were thus selected as starting points for further docking procedures. In all cases, the selected conformers presented rather close energies, corresponding to rapidly interconverting forms of the molecule, and the docking results were generally comparable between each conformer. Thus, the most representative conformer for docking results was chosen for data presentation.

3. Docking calculations

Molecular docking experiments were performed using AutoDock 4 (release 4.2.6) in the semi-flexible mode with the Cel-Pgp-1 4F4C PDB structure kept rigid, and prepared with AutoDock Tools (Morris et al., 2009). AutoDock, which is the most cited docking software (Sousa et al., 2006; Sousa et al., 2013), has a free-energy scoring function, based on AMBER force field and a large set of diverse protein-ligand complexes with known inhibition constants. Few residues of the protein could have been declared as flexible in the PDBQT file, but the program restrains the total number of torsional degrees of freedom to 32, shared between the ligand and the receptor. This is a drawback in the case of the Pgp structure, since the inner chamber is large, and in the absence of consistent indications about the exact location of the binding sites of the various drugs, different cavities have to be taken into account in the

calculation. Thus, we privileged an approach based on better coverage of ligand flexibility and a grid box extended to the whole membrane part of the receptor protein. Indeed, for all ligands tested, the docking box, in which grid maps were computed using program AutoGrid 4, encompassed all the TM helices and the whole internal cavity, including lateral access channels and protein surface, to allow a large sampling of potential poses. The grid built by AutoGrid included 100, 124, and 126 points in x, y, and z directions, with a grid spacing of 0.375 Å to allow a good compromise between resolution of the explored volume and the size of the binding area (box dimensions 37.5 x 46.5 x 47.3 Å, centered in the inner cavity of Pgp, at the point x=22.2 Å; y=77.6 Å; z=-1.4 Å). For each ligand conformer, 100 independent calculations were performed using the Lamarckian genetic algorithm. All the other parameters were set at the default value.

The 100 generated poses were assigned a score calculated by AutoDock that can be considered as an estimated free energy of ligand binding (indicative of binding affinity). They were then clustered as a function of the closeness of their positions and conformations with RMSD set at 2.0 Å, and finally ranked by their binding score (for the best pose in the cluster). The results are displayed in an energy scores histogram, which reproducibility could be assessed by comparing docking calculations performed on close ligand conformers, or on a truly duplicated calculation. As a result, binding energies (positions of the best pose in each cluster of the histogram) were found to fall within a range of 0.25 kcal/mol, and the number of poses in a cluster within 10%. This gives an indication of the accuracy of the histogram parameters in our series of runs, i.e. the binding energies and the overall distribution of clusters. In some cases, an isolated cluster of very good affinity (lowest binding energy) and containing a small number of conformations (typically 1 to 2 poses) was observed; this actually always corresponded to a physically relevant docking position.

4. Data analysis

Different parameters and observables can be used in the interpretation of docking data issued by AutoDock: binding energies (i.e. docking scores), histogram bars energy range, profile of the histogram, and location of calculated positions in the protein structure. The position of clusters in figures and tables corresponds to the binding energy value of the lowest energy pose in the cluster. The spreading of clusters in the energy scores histogram was considered as a general indication for ligand docking calculation reliability: the less scattered the energies, the higher specificity of binding can be expected. Conversely, a pseudo-Gaussian

profile for a group of histogram bars may suggest a non-specific docking. However, in some cases such as that of the complex molecule ACD, very negative binding energies (i.e. very good docking scores) were found in poorly populated clusters whereas the general aspect of the histogram was scattered. We considered these poses as relevant, despite the lack of sampling, in view of the gap energy with the next clusters, revealing high binding site specificity. In addition, we launched a 500 poses calculation on ACD for larger sampling and reproducibility testing, which confirmed the existence of the 2 lowest energy poses, and their overall ratio (15 conformations over 500, i.e. 3% of poses).

The question arose as to whether the lowest energy cluster or the most populated cluster (highest histogram bar) had to be considered. The AutoDock docking score is based on an empirical free-energy force field which has been parameterized using a large number of protein-inhibitor complexes for which both structure and inhibition constants were known, and thus should reliably reflect the affinity of the ligand for the receptor, and the stability of the ligand-protein complex. In contrast, the number of poses found in a cluster reflects the number of times that conformers are found in very close (within RMSD) binding sites, without any anticipation of their stability. Thus considering the highest cluster in the energy histogram may be not relevant *per se* for identifying the most probable docking site of the considered ligand. Practically, for each ligand, when the difference in the energies between the two best-scored clusters was more than 2 kcal/mol, the lowest energy one was considered as the most representative, since the other ones correspond to ligand-protein complexes associated to negligible lifetime. Alternatively, when the energy gap was narrower, we manually analyzed all the clusters within 2 kcal/mol below the lowest one for the localization of the included poses in the protein 3D structure. Finally, one or two main clusters were selected as representative, and they corresponded most often to the first and second minimum energy clusters, except when clusters corresponding to non-relevant positions in the protein were found interleaved in the ranking. In some cases, this protocol of validation led us to consider a double binding site on the protein, leading to a stoichiometry of one or two depending on whether these two docking positions were partially overlapping or not. Such a possibility simply reflects the large size of the multispecific binding domain, which likely encompasses the whole "inner chamber".

Lastly, criteria for discriminating relevant from non-relevant docking poses had to be defined. This delineation was made necessary by the fact that we enlarged the zone of docking search to the whole membrane part of the protein. The large size of the grid box ensured an exhaustive conformational exploration not biased by preliminary beliefs, although it

participated to the scattering of the docking results. Poses that were outside of the expected ligand binding pocket (see Discussion §4.1) were not considered for further analysis. These were essentially poses located either "outside" of the Pgp structure, i.e., at the protein/lipid interface, or lowermost in the transmembrane domain, i.e., at the level of interface with the cytosolic medium ("cytosolic antechamber"), where the phospholipids polar headgroups are likely to be invaginated into the cavity between the transmembrane helices in the inward-facing Pgp conformation (Haubertin et al., 2006). These poses can have a functional relevance, for example as allosteric modulator sites, or not, but likely not as binding sites for transport. Finally, a search of the possible access channels allowing ligands to reach the inner pocket was performed using Mole 2.0. (http://mole.upol.cz/). In addition to the wide opening of the protein towards the cytosolic interface, the result showed two lateral tunnels located between TM10 and TM12, communicating with the cytosolic leaflet of the membrane (data not shown).

For each lowest energy pose of selected clusters, the number and nature of interacting residues were analyzed within the protein. Among these, particular interest was given to residues belonging to the "hotspots for drug binding", described hereafter.

5. Determination of the residues constituting the "hotspots for drug binding"

A number of experimental works have been conducted for the purpose of determining the key residues responsible for multidrug recognition by mammalian Pgp (human and rodent isoforms). They initially included various directed-mutagenesis analyses that allowed identification of residues whose mutation led to alterations of the MDR profile, in contrast to numerous mutations that led to unspecific global decrease of Pgp function. The significance of these mutations was shown by cytotoxicity assays, which were a good indication of their involvement in the multi-specific drug recognition and binding. Furthermore, some additional data were collected with a chemical cross-linking approach, using a few drug derivatives bearing a moiety that could be activated, and ensuring specificity by testing protection by the native drug. All these data have been compiled in a review by Shilling et al., (2006), and were completed by including some references by Loo & Clarke (Loo and Clarke, 2001, 2002; Loo et al., 2006a, b; Bessadok et al., 2011). Finally, the release of the crystal structures of murine Pgp, co-crystallized with hydrophobic cyclic peptide inhibitors than can be considered as functionally relevant ligands, pointed to a set of contact residues; however, two versions of the interpretation of the experimental X-ray diffraction data have been published (Aller et al., 2009;

Li et al., 2013) that differed in the orientation of some transmembrane helices and side chains, which led us to consider both lists of contact residues equally. All these identified residues form a collection of 62 residues, as displayed in Suppl. Table S1, coming from 4 mammalian proteins (Hsa-ABCB1, Mmu-ABCB1a & b, Cgr-ABCB1). Moreover, 14 of these 62 residues are common between at least two different techniques. They are all situated in the transmembrane part of the protein. Altogether, they provide a frame in the inner chamber that offers a set of anchoring points for multispecific recognition and binding, and eventual translocation, of various transport ligands.

Multiple protein sequence alignments have been performed on Cel-Pgp-1, human ABCB1, murine ABCB1a and ABCB1b, and Chinese hamster ABCB1, using Muscle software (Edgar, 2004). We checked that the transmembrane segments were satisfactorily aligned. Among the 62 human ABCB1 hotspot residues, 16 (26%) were found identical and 12 homologous with the corresponding residues in Cel-Pgp-1, representing a global conservation ratio of 45%. As a comparison, human ABCB1 and Cel-Pgp-1 display a global similarity of 63% (BLASTP positive matches), and still 58% when considering the TMDs only. This indicates that hotspots are subjected to more genetic variation than the overall sequence. More precisely, this set of residues is mainly hydrophobic, but less markedly for Cel-Pgp-1: the hydrophobic (F-Y-A-L-I-V-M) / hydrophilic (S-T-N-Q-H) ratio is 44/14 for Hsa-Pgp and 37/20 for Cel-Pgp-1, respectively.

III. RESULTS

1. Docking experiments on compounds stimulating or not the ATPase activity of Cel-Pgp-1

To study drug interactions with Cel-Pgp-1, we have performed *in silico* docking calculations using AutoDock 4 in semi-flexible docking mode on the Cel-Pgp-1 crystal structure (PDB 4F4C) for a set of seven drugs belonging to a wide range of bioactive compounds and presenting various structures and molecular weights (MW) (Table 1 and Suppl. Fig. S1). Six of these drugs were chosen based on their previous biochemical characterization, showing ability to stimulate the ATPase activity of solubilized Cel-Pgp-1 from 3 to 10 times (Jin et al., 2012) (Table 1). Rhodamine 123 was chosen as a negative control, as it showed no stimulation of ATPase activity of Cel-Pgp-1.

a) Binding mode of known substrates of Cel-Pgp-1

For valinomycin (VLN), the energy scoring histogram showed a single lowest energy cluster at -10.3 kcal/mol (VLN1) with a significant number of poses (42%) (Table 1 and Fig. 1A). This strongly supports the good specificity and high affinity of VLN1 for one binding site. This docking position was similar to all other clusters, located in the "cytosolic part" of the inner chamber formed by the TMDs of Cel-Pgp-1, at the level of the interface between membrane and cytosol (Fig. 1B). It interacted with 14 residues of the transporter, including 3 hotspot residues, and formed 3 H-bonds (Table 1 and Fig. 3A).

For vinblastine (VBL), two lowest energy clusters, VBL1 and VBL2, were found at -10.5 and -9.9 kcal/mol, consistent with a good affinity, and contained 15 and 14 poses respectively (Table 1 and Fig. 2A). Their binding positions were overlapping, with different orientations, deep in the inner chamber formed by the TMDs of Cel-Pgp-1 (Fig. 2B and 2C). VBL1 and VBL2 interacted respectively with 16 and 12 amino acids residues, including 10 and 7 hotspot residues, and formed 2 and 1 H-bonds (Table 1). Their binding sites showed 9 common interacting residues, including 6 hotspots (Table 2). They were both totally apart from that of VLN1, but shared 3 common interacting residues with VLN1: Q913, R916 and G1032, separating the binding sites of VBL and VLN (Table 2 and Fig. 3A).

For actinomycin D (ACD), the clustering histogram showed a couple of lowest energy docking poses, ACD1 and ACD2, with binding energies of -17.0 kcal/mol and -14.9 kcal/mol, respectively, indicating a very high affinity of ACD for Cel-Pgp-1 (Table 1 and Suppl. Fig. S2A). The interaction of these two poses with Cel-Pgp-1 appeared in the inner chamber, from its cytosolic part to the middle of the TMDs (Suppl. Fig. S3A and S3B). ACD1 and ACD2 both interacted with 19 residues of Cel-Pgp-1, including 8 and 7 hotspot residues, and forming 2 and no H-bond(s), respectively (Table 1). They showed 15 common interacting residues, confirming the high specificity of the molecule to this binding site (Table 2). The two positions were indeed superimposed for most of the molecule, only slightly differing by the orientation of the peptidic moiety that was closer to the cytoplasmic opening of the inner chamber of Cel-Pgp-1 (Suppl. Fig. S3A and S3B). ACD binding site shared several interacting residues with each of the well separated binding locations of VLN and VBL (Table 2). ACD indeed overlapped with each of these moelcules with respectively one or the other pseudo-peptidic moiety (Fig 4B and 4C, left panel), which formed a H-bond with either Q913 or R916, thus stabilizing the binding of ACD around its central polycycle (Fig. 3B).

For dipyridamole (DPM), the energy clustering was more scattered than for previous molecules, with the highest number of poses in a cluster of only 4, and binding energy ranging from -7 to -1 kcal/mol (Suppl. Fig. S2B). This reflected some diversity of the docking results. However, 2 clusters were found 1 kcal/mol away from all the others: at -7.1 kcal/mol (DPM1, 3 poses) and at -7.0 kcal/mol (DPM2, 4 poses) (Table 1 and Suppl. Fig. S2B). For both clusters, binding to Cel-Pgp-1 appeared to be localized in the inner chamber, DPM1 being found more at its apex than DPM2 (Fig 4B and 4C, middle panel). DPM1 and DPM2 showed 13 and 12 interacting residues, including 11 and 7 hotspot residues and forming 2 and 4 H-bonds, respectively, and shared only 3 common interacting residues (Table 1 and 2). DPM thus appeared to present a "dual" binding site composed of two partially overlapping sub-sites, localized across the VBL1 and VBL2 binding sites (Fig. 4C, middle panel). DPM2 binding site also overlapped with that of ACD.

For progesterone (PRG), the energy histogram was very simple and showed two main clusters, very close in energy. These two minimum binding energy cluster PRG1 (13 poses) and PRG2 (85 poses) showed an energy of -7.23 kcal/mol and -7.17 kcal/mol, respectively (Table 1 and Suppl. Fig. S2C). PRG1 and PRG2 interacted with 10 and 9 residues, all or 6 of them identified as hotspot residues, and formed 2 or no H-bond(s) with Cel-Pgp-1 (Table 1). PRG1 was found more at the apex of the inner chamber in the transmembrane domains of Cel-Pgp-1, than PRG2 (Suppl. Fig. S3C and 3D), so that their interacting residues were totally different (Table 2). PRG thus presented two separated binding sites, consistent with a stoichiometry of 2, which showed similar affinities with Cel-Pgp-1. The PRG1 binding site was mainly superimposed with DPM1, whereas the binding site of PRG2 overlapped with that of VBL, DPM2 and ACD (Table 2 and Fig. 4C, middle panel).

For paclitaxel (PCT), the energy histogram presented a scattered clustering, with a minimum energy pose at -8.3 kcal/mol (PCT1) (Table 1 and Suppl. Fig. S2D). It was found at the cytosolic part of the inner chamber of Cel-Pgp-1 (Fig. 4B and 4C), similarly to the second and third lowest energy poses. Only the fourth lowest energy pose (PCT4) at -7.5 kcal/mol (Suppl. Fig. S2D), was more buried in the inner chamber of Cel-Pgp-1 (Fig. 4B and 4C, right panel). PCT1 and PCT4 contained 14 and 17 interacting residues, including 3 and 15 hotspot residues, and formed 2 or no H-bond(s), respectively (Table 2). They shared one common interacting hotspot residue, but their binding sites did not overlap, consistent with a stoichiometry of 2 (Table 2 and Fig. 4C, right panel). The binding site of PCT1, however,

overlapped with that of VLN and ACD, whereas the binding site of PCT4 superimposed those of VBL, DPM and PRG.

Altogether, the *in silico* docking data showed high affinity interactions with Cel-Pgp-1 for a set of 6 validation drugs. This is in full agreement with the report of their *in vitro* functional characterization on Cel-Pgp-1, which showed that these 6 drugs clearly increased Cel-Pgp-1 ATPase activity (Table 1) (Jin et al., 2012).

b) Binding mode of a compound that did not stimulate Cel-Pgp-1 ATPase activity

For rhodamine 123 (RHO), the energy histogram showed 3 lowest energy clusters ranging between -5.0 and -4.6 kcal/mol (Table 1 and Suppl. Fig. S2E). However, these 3 clusters corresponded to a binding position situated low on the TMD of Cel-Pgp-1, outside of the membrane plane (Suppl. Fig. S3E): such an "ectopic" location outside the inner chamber should make it difficult for the ligand to undergo a protein transconformation-induced transmembrane translocation (according to the molecular mechanism suggested by Aller et al. (2009)), and is actually consistent with the absence of Cel-Pgp-1 ATPase activation by RHO. Therefore, such a docking location of a ligand on the "cytosolic antechamber" should be considered as likely non-relevant for a transport function.

In contrast, the 4th and 5th lowest energy clusters for RHO corresponded to a location within the inner chamber of Cel-Pgp-1, but they displayed energies around -3.8 kcal/mol (Table 1 and Suppl. Fig. S3F). This small negative energy score clearly differed from those of the molecules activating Cel-Pgp-1 ATPase, which ranged between -7 to -17 kcal/mol (see above), and indicates a very low affinity of RHO for this binding site on Cel-Pgp-1. Hence, the whole validation set, including the 6 Cel-Pgp-1 ATPase activating molecules (VLN, VBL, ACD, DPM, PRG, PCT) and the non-activator RHO, showed an excellent correlation between *in vitro* data and *in silico* docking results. This provided a good support to the reliability of our computational modelling approach.

From a technical point of view, the semi-flexible strategy was chosen here rather than a fully flexible docking calculation. Indeed, under AutoDock 4.2, a flexible protocol is much more machine time-consuming and the number of residues to be considered as flexible in the huge inner chamber would be too high for the maximum number of torsional degrees of freedom allowed in the program (32 including those of the flexible ligand). Also, declaring a

limited set of flexible residues would imply knowing in advance where the drug binds in the 12 transmembrane helices chamber. This contradicts our strategy to exhaustively search all possible docking positions in the membrane part of the ABC protein. Thus, the semi-flexible searching mode is more suited for a first exploratory study when considering the validation ligand set composed of 7 compounds, under a total of 14 different conformers, initially considered. Including lateral chain flexibility of the interacting residues would *a priori* lead to improved binding energies, but owing to their large number (a mean of 13.7 interacting residues per ligand, only a fraction of them being merely flexible), it can be expected that the docking positions obtained are not significantly different. Thus, we consider that the semi-flexible docking strategy of AutoDock is suitable and sufficient for this screening study, aimed at describing the multispecific recognition capacity of Cel-Pgp-1.

Another theoretical point that could be addressed is the possible influence of surrounding membrane lipids on the docking calculations, since the inner chamber is localized at the level of the transmembrane helices, with a lateral aperture in direct communication with the surrounding membrane lipid phase. However, the presence of lipids in the potential access channels to the inner chamber cannot be taken into account by the docking calculations due to the fluid character of the lipid phase. Nevertheless, determining the direct ligand-protein interactions remains a key goal since these interactions are likely responsible for activating the further steps of ligand translocation.

c) Binding site characterization of known substrates of Cel-Pgp-1

The general characteristics of the binding domain of Cel-Pgp-1 could be described following the docking results of the 6 known substrates : VLN, VBL, ACD, PRG, DPM and PCT, whereas the RHO lowest-energy binding site was clearly out of this binding chamber (Suppl. Fig. S3E). The high-affinity binding site spread from the cytoplasmic opening of the transporter to the core of the protein (Fig. 4A), buried between, and interacting with, all the TMDs, except TM2, TM4, TM8 and TM9 that did not reach the inner chamber. Its general shape appeared to be composed of 2 sub-domains (Fig. 4B). Its cytosolic part, open to the inner leaflet of the membrane (Binding Domain 1, BD1) accommodated the largest ligands (ACD, VLN), while the other, smaller ligands (PRG, DPM and VBL) bound a second pocket buried within the helices of the transporter (Binding Domain 2, BD2). PCT, of intermediate size, can bind alternatively to BD1 and to BD2. The largest tested ligand, ACD, overlapped with BD1 and BD2, its polycyclic intermediate moiety interacting with the narrowing long chain residues

found between BD1 and BD2 (Fig. 3B). The interaction of ACD with the area separating BD1 and BD2 revealed 8 residues of interest, including Q913 and R916 that formed H-bonds and M367 also interacting with VBL1 (Fig. 3B and Table 2).

In the overall binding domain, 12 key residues involved in H-bonds with at least one of the ligands were identified (Table 2, Fig. 3A and 3B). Interestingly, 9 residues on TM12 were hotspot residues according to our alignment with mammalian Pgp (Suppl. Table S1). In addition, the two small half-helices, TMa and TMb at the N-terminal of Cel-Pgp-1, not described in mammalian Pgp, interacted with the 6 molecules tested. In particular, TMb contained 4 residues (E22, D23, K26 and K30) contributing to the formation of the area separating BD1 and BD2 (Fig. 3B). Two of them, as well as two residues of TMa, formed H-bonds with 3 different substrates (Table 2). Therefore, these small helices, specific for Cel-Pgp-1, participate in delineating the ligand binding sites, and may be critical for the affinity of many substrates to Cel-Pgp-1.

2. Binding mode of other substrates of mammalian Pgp previously tested on the ATPase activity of Cel-Pgp-1

We have docked *in silico* 5 more compounds of interest, for which the effect on the ATPase activity of Cel-Pgp-1 has been previously reported by Jin et al. (2012): colchicine (CCH), vincristine (VCR), verapamil-S (VRP), doxorubicin (DXR) and daunorubicin (DNR) (structures presented in Suppl. Fig. S4).

a) Docking of colchicine, vincristine and verapamil

For CCH, the two lowest energy clusters CCH1 and CCH2 were found at -8.9 and -8.5 kcal/mol with 12 and 70 poses, respectively (Fig. 5A). They were located at different positions of the inner chamber, both in BD2 (Fig. 5B&C). CCH1 and CCH2 both interacted with 14 residues, including 6 and 14 hotspot residues, and formed 1 H-bond on Q913 and N994, respectively (Table 3 and 4). They only shared 1 common interacting residues, L91, that they reach on its opposite sides, so as the two positions did not overlap, and were then consistent with a binding stoichiometry of 2 (Fig. 6B and 6C).

For VCR, the positions of the two lowest energy clusters, VCR1 (19 poses at -9.9 kcal/mol) and VCR2 (4 poses at -9.8 kcal/mol) (Table 3 and Suppl. Fig. S5A) were partially overlapping in BD2 (Fig. 6B&C), and representative of the few following clusters. VCR1 and

VCR2 interacted with 13 and 16 residues, respectively, including 10 hotspot residues and forming 3 H-bonds in both cases. They shared 4 common interacting residues, including 3 hotspots (Table 4 and Suppl. Table S2), hence forming a dual binding site. Globally, VCR2 was found superimposed to the closely chemically related ligand VBL, with 12 common interacting residues between VCR2 and at least VBL1 or VBL2, including 7 hotspots, whereas VCR1 and VBL were less overlapping, with overall 5 common interacting residues, including 4 hotspots.

For VRP, the clustering was more scattered than that of CCH or VCR (Suppl. Fig. S5B), but nevertheless two relevant lowest energy clusters were found at -6.5 kcal/mol (VRP1, 6 poses) and -6.1 kcal/mol (VRP2, 5 poses) (Table 3). They both positioned in the deepest part of BD2, with the two different possible conformations: VPM1 being folded and VPM2 stretched (Fig. 6B and 6C). The same situation was found whether starting from a folded or a stretched conformer for the docking calculations (not shown). The few following clusters were all similar to VRP1 or VRP2. VRP1 and VRP2 partly overlapped by sharing 6 interacting residues (including 5 hotspot residues) among their respective 9 and 13 interacting residues (including 8 and 10 hotspot residues, respectively (Table 2 and Suppl. Table S2). VRP docking thus gave a dual binding site, locally integrating the conformational dynamics of this flexible ligand.

As a first summary from these 3 tested molecules, these binding energies (Eb from -6.1 to -9.9 kcal/mol) (Table 3) were comparable to the Eb values of the ligands from the set of validation (from -6.5 to -10.5 kcal/mol, except the particular case of ACD) (Table 1), suggesting that CCH, VCR and VRP can specifically bind to Cel-Pgp-1. Interestingly, these molecules had low to medium MW (400 to 830 Da, Table 3) as compared to the first molecules tested (315 to 1250 Da, Table 1), and all were found to bind in the binding domain BD2.

b) Docking of the structurally very close molecules doxorubicin and daunorubicin

For DXR, the two lowest energy clusters were found at -9.3 kcal/mol (DXR1, 3 poses) and -8.3 kcal/mol (DXR2, 7 poses) (Table 3 and Suppl. Fig. S5C). They interacted with 14 and 11 residues, including 12 and 7 hotspot residues, respectively, and they both formed 4 H-bonds (Table 3). DXR1 bound slightly deeper in BD2 than DXR2, and they partly overlapped with 5 common residues, all hotspot residues (Table 4 and Fig. 6B and 6C, left panel), thus leading to a dual binding site.

For DNR, the lowest energy cluster (DNR1), found at -5.6 kcal/mol (2 poses), was followed by two clusters positioned out of the inner chamber (Table 3 and Suppl. Fig. S5D). DNR1 bound to the deepest part of BD2, with 12 interacting residues, including 10 hotspot residues and 4 H-bonds (Table 3 and Fig. 6B and 6C, middle panel). DNR4 and DNR6 bound around the same area but at slightly different positions and orientations. However, their rather high Eb values made them unlikely to be relevant specific binding sites. DNR1 and DXR1 binding sites highly superimposed, as they shared 11 common interacting residues with 9 hotspot residues, and formed 1 H-bond with the same residue, T1025 (Table 4 and Suppl. Fig. S7). Interestingly, the only structural difference between DXR and DNR is a supplementary OH group in DXR that brings the molecule from 528 to 544 Da. The comparison of the H-bond patterns established by DXR1 and DNR1 allowed pointing out a noticeable role for Q98.

3. Binding mode of other molecules of interest

We performed further docking experiments on 6 other mammalian Pgp substrates, comprising the two fluorescent dyes, Hoechst 33342 (HST) and calcein-AM (CAM), the three drugs, ketoconazole (KTC), cyclosporin A (CSP) and valspodar (VSP), and finally the endogenous membrane component, cholesterol (CLS) (structures shown in Suppl. Fig. S8).

For HST, the lowest energy cluster (HST1, 1 pose) appeared at -10.2 kcal/mol, and was positioned very similarly to HST2, while the only other main cluster, HST3 (45 poses), was found at -10.0 kcal/mol (Table 5 and Fig. 7A). Positions of HST1, with 15 interacting residues (8 hotspot residues, 0 H-bond), and HST3, with 9 interacting residues (8 hotspot residues, 1 H-bond) presented two quite perpendicular orientations that crossed in the lower part of BD2 (Fig. 7B&C). They shared 2 common interacting residues, with one hotspot (Table 6 and Fig. 9), hence showing a dual binding site.

For CAM, the lowest energy cluster, at -5.4 kcal/mol, was found out of the inner chamber, but the second cluster (CAM2, 2 poses), at -5.3 kcal/mol, was found in BD1, near its cytosolic aperture (Table 5 and Suppl. Fig. S9A). CAM2 had to 13 interacting residues (2 hotspot residues, 6 H-bonds). However, its rather high Eb values indicated very low affinity for this specific binding site, found similar for other clusters.

For KTC, we found that the three lowest energy clusters positioned at 3 different relevant sites in BD2. KTC1 (1 pose), at -9.98 kcal/mol, and KTC3 (15 poses), at -9.85 kcal/mol, both in a stretched conformation, displayed 17 interacting residues (15 hotspot

residues, 1 H-bond) and 13 interacting residues (10 hotspot residues, 1 H-bond), respectively (Table 5 and Suppl. Fig. S9B and 10B). They shared 7 interacting residues, all hotspots (Table 6), leading to a partial overlap in the deepest part of BD2 (Fig. 9), thus showing a dual binding site. In contrast, KTC2 (1 pose), at -9.96 kcal/mol, displaying 16 interacting residues (6 hotspot residues, 3 H-bonds) (Table 5 and Suppl. Fig. S9B and 10C), was found to be positioned, in a highly folded conformation, low in BD2, and without any overlapping with KTC1 or KTC3, hence finally consistent with a binding stoichiometry of 2.

For CSP, the lowest energy cluster CSP1 (2 poses), at -10.3 kcal/mol, was only considered as it was representative of the few following clusters in the energy histogram (Table 5 and Suppl. Fig. S9C). The position of this cluster was in the lowest part of the inner chamber, in BD1 (Fig. 9) and showed 17 interacting residues (2 hotspot residues, 2 H-bonds) (Table 5 and 6).

For the very closely related molecule VSP, the two lowest energy clusters, VSP1 (6 poses), at -9.5 kcal/mol, and VSP2 (25 poses), at -9.2 kcal/mol, were both positioned very close together, in BD1 (Table 5, Fig. 9 and Suppl. Fig. S9D). VSP1 and VSP2 interacted with 18 and 14 residues, including 5 and 3 hotspot residues and 2 and 1 H-bond, respectively (Table 5). They shared 8 common interacting residues, including 2 hotspot residues, and formed a common H-bond with K30. They partly overlapped with CSP binding sites, with which they shared 11 and 9 residues, respectively, including 1 hotspot residue in both cases, and all forming an H-bond on K30 (Table 6 and Fig. 9). Finally, from this last series of tested ligands, these two large cyclopeptidic molecules were the only ones to dock in BD1 with well negative Eb values, in a manner very comparable to that of VLN, another large cyclopeptidic molecule.

For CLS, the lowest energy pose (CLS1) at -9.4 kcal/mol (Table 5 and Fig. 8A), was found at a very deep position in BD2 (Fig. 8B), similar to the other following clusters CLS3/4/5. CLS1 interacted with 13 residues (9 hotspot residues, 1 H-bond) (Table 5). Alternatively, the second cluster (CLS2, 6 poses), at -8.7 kcal/mol (Table 5 and Fig. 8A), was found lower in BD2 (Fig. 8C). It showed 12 interacting residues (6 hotspot residues, 1 H-bond) (Table 5), and shared no common residues with CLS1 (Table 6), from which it was totally apart (Fig. 9), consistently with a binding stoichiometry of 2.

IV. DISCUSSION

For the first time, a reliable crystal structure with a good resolution is available for an eukaryotic multidrug membrane full ABCB transporter. This is especially interesting because (i) it presents a fair sequence homology with the mammalian Pgp (ABCB1), in particular well-described to be responsible for the MDR phenotype in cancer cells, and (ii) it presents an inward-facing open conformation supposed to be competent for ligand recognition and binding as the first step of its transport cycle. This structure provides the valuable basis for an *in silico* approach aimed at delineating the binding modes of a variety of chemically unrelated molecules, which can be compared to the classical situation of specific ligand/receptor or enzyme/substrate recognition (the "key-in-the-lock" dogma). Actually, *in silico* computation is the only reliable mean for addressing the structural arrangement of a large series of ligands on their cognate membrane protein receptor since it cannot be expected to collect all the corresponding co-crystal structures for the considered complexes. This work contributes to raise and elaborate the vision of a multispecific large domain, situated in the inner chamber of the protein. This also participate in the determination of potent interacting residues offering a rich palette of continuously connected sub-sites, which can accommodate a large variety of chemically unrelated molecules of sizes ranging from 300 to 1250 Da.

1. Validation of *in silico* docking approach on Cel-Pgp-1

The general structure of the membrane ABC proteins has led to the generally-accepted dogma of energy coupling between ATP hydrolysis and substrate translocation (Sarkadi et al., 2006). Binding of a ligand to an ABC transporter, the first essential step for being eventually transported, is thus expected to induce a stimulation of its ATPase activity. Reciprocally, when a molecule is observed *in vitro* to stimulate Cel-Pgp-1 ATPase activity, it can be considered as a specific ligand that will likely be actively transported. We thus modeled *in silico* the binding on Cel-Pgp-1 structure of the 6 structurally unrelated compounds: VLN, VBL, ACD, DPM, PRG and PCT, reported to be ATPase activators by (Jin et al., 2012). The low AutoDock4 binding energies calculated for all of them ranged from -7 to -17 kcal/mol, in agreement with scoring values that are generally found for high affinity and specific binding of ligands to protein active sites using the AutoDock method (Morris et al., 1998). The binding energies of the best poses of these compounds proved to be consistent with their respective *in vitro* potency to stimulate ATPase activity of Cel-Pgp-1, giving confidence in the current docking protocol.

Furthermore, Jin et al., (2012) reported the half-activation concentration values, considered as their apparent affinity, of ACD and PCT on isolated membranes containing Cel-Pgp-1, to be similar, in the 0.03-0.05 µM range. The apparent affinity of an amphiphilic/hydrophobic ligand, interacting with a protein from the membrane lipid phase, can be considered as the combination of its membrane partition coefficient, which can be evaluated by logP, and the "true affinity" of its interaction with its specific binding site on the protein once in the membrane. Taking into account the very different hydrophobicity of ACD and PCT, whose logP values are estimated at -1.2 and 3.3, respectively (MarvinSketch 15.5.4) (Viswanadhan et al., 1989), it can be inferred that ACD should have a much higher true affinity for the protein than PCT to exhibit comparable experimental affinities. This is actually what is shown by the *in silico* docking calculations, which gave binding energy values of respectively -17 and -8.3 kcal/mol. This "semi-quantitative correlation", although based on only two drugs, contributes to the confidence in the *in silico* data obtained.

Since RHO is a non-activator of Cel-Pgp-1 ATPase, further information on the relationships between drug binding and drug transport can be gained by docking RHO as low affinity control. The ectopic binding sites (e.g. in the cytosolic antechamber) found by *in silico* calculations for the three first clusters of RHO are considered as irrelevant for active transport. By inference, this is consistent with the inner chamber forming the binding domain for initiating the active translocation cycle. Also, the low binding energy (-3.8 kcal/mol) of RHO for the docking site located within the inner chamber (4^{th} and 5^{th} clusters) corresponds to a very low affinity, that can be expected to require a too high concentration of RHO to be observable *in vitro*. In addition, RHO docking experiments provides a reference level in the energy scale as being biochemically non-relevant, compared to the set of transported ligands whose binding energies range between -7 and -17 kcal/mol.

Another validation of our *in silico* data comes from the comparison of the docking data with what is known for drug binding to mammalian Pgp (ABCB1), the most studied multidrug ABC transporter on which numerous data are available. In particular, the list of the 62 "hotspot residues", that we updated from the prior list established by (Shilling et al., 2006) (see Computational methods section), provides a global insight to the residues in the protein sequence that are involved in its multispecific recognition capacity. In the total of 11 relevant positions calculated for the docking of the 6 Cel-Pgp-1 ATPase activators, we have a total of 155 contact residues, amongst which 87 (nearly 60%) are aligned with these hotspot residues. Such high ratio indicates that our *in silico* docking calculations give a fair convergence,

consistent with the structure/function relationship of these two proteins which appear to exhibit high similarity in the topology of their inner chambers.

As a whole, for the validation set of ligands, our *in silico* data shows a good overall agreement with the few *in vitro* data available, regarding the enzymological properties of Cel-Pgp-1 as well as the global drug recognition determinants in the homolog mammalian Pgp (hotspot residues). This gives to these *in silico* data the confidence required for applying the same strategy to other drugs, in order to reliably predict their binding characteristics on Cel-Pgp-1.

2. Further analysis of the *in silico*/*in vitro* correlation

We then included in our *in silico* docking series 5 more molecules chosen among those tested in Jin et al, (2012), which showed either an only marginal stimulation of Cel-Pgp-1 ATPase activity : VCR, VRP, or did not show any ATPase stimulation : CCH, DXR, DNR.

VCR showed its two lowest energy clusters, relevant for describing its binding mode, with Eb values very comparable to those found for VBL (VCR1 and VCR2 at -9.9 and -9.8 kcal/mol and VBL1 and VBL2 at -10.5 and -9.9 kcal/mol, respectively). Taking into account the very close chemical structures of these two molecules, and hence their close logP values, it can be considered that they display closely related behavior. Indeed, Jin et al. (2012) reported a small (about a factor 2 at 200 µM) Cel-Pgp-1 ATPase stimulation for VCR, although slightly more moderate than for VBL (about a factor 3). It can thus be concluded that they both significantly stimulate Cel-Pgp-1 ATPase, and hence are to be considered as "good" ligands for Cel-Pgp-1, according to *in vitro* as well as to *in silico* data.

CCH and DXR each presented a well-negative Eb value (their two lowest energy clusters between -8.3 and -9.3 kcal/mol), along with a logP value (1.5) showing a more hydrophilic behavior than any of the other tested molecules in our series (except ACD). The fact than Jin et al. (2012) did not observe any ATPase stimulation for these two drugs, tested at not too high concentrations (200 and 500 µM, respectively), can be reconciled with the *in silico* data by the fact that they are supposed to partition into the membrane lipid phase before being recognized by the transporter. This pathway is actually consistent with the position of the aperture of its inner chamber at the level of the membrane plane, as typically described for efflux ABC transporters. These two molecules have thus to be considered as theoretical ligands, as shown by the *in silico* calculations that only take into account the various aspects of protein/ligand interactions without any integration of the surrounding membrane lipid phase. However, such

hydrophilic theoretical ligands are not biochemically relevant ligands because of their too low apparent affinity for this membrane transport system, due to a too low membrane partition, which is hardly compatible with practically reachable concentrations for *in vitro* experiments.

At variance, VRP gave a much less negative Eb value (-6.5/-6.1 kcal/mol), which puts it within the gap range between very low-affinity ligands (> -5 kcal/mol) and good ligands (< -7 kcal/mol) as deduced from the validation set. However, it has a logP value (5.0) showing its clearly hydrophobic character. As a result, its high membrane partition explains that Jin et al (2012) observed a small ATPase stimulation at the highest tested concentration of 5 mM. Thus, *in silico* as well as *in vitro* data both led to the conclusion that VRP is a weak ligand, probably of too low affinity to be pharmacologically relevant.

DNR gave an even higher Eb value than VRP (lowest energy cluster at -5.6 kcal/mol), along with a logP value (2.3) showing a rather hydrophilic character. Both parameters are converging to be consistent with the absence of any detectable ATPase stimulation. It should be thus considered as a weak (below detectable level) or non-ligand of Cel-Pgp-1.

In conclusion, thanks to these additional 5 tested molecules, we observed a very satisfying *in silico*/*in vitro* correlation when considering a total of 12 molecules tested by Jin et al, which encompassed a large range of theoretical and apparent affinities, at the only condition to take into account the quantitative parameter of membrane partition of the ligands of potential interest.

In the very particular case of ACD, although it is clearly hydrophilic, it nevertheless efficiently binds Cel-Pgp-1 thanks to a remarkably high “theoretical binding” (Eb at least 6 kcal/mol more negative than the best other ligands), and this will be specifically discussed below.

3. Molecular properties of the multispecific binding domain of Cel-Pgp-1

Our series of docking calculations, performed to determine the binding modes of a total of 18 tested ligands, comprised 11 molecules presenting miscellaneous chemical structures (the prediction set) that extended the initial validation set (7 molecules). This provided a valuable mean to investigate and finely delineate the binding site(s) of Cel-Pgp-1, which otherwise remains uncharacterized in the absence of ligand-bound structure of this protein receptor.

a) Cel-Pgp-1 is a multidrug transporter

The first point to be mentioned is that some tested molecules appeared to be very weak affinity or non-ligands for Cel-Pgp-1: among our series, they are RHO, CAM and DNR. Actually, they were characterized by low negative Eb values, i.e. -3.8, -5.3 and -5.6 kcal/mol, respectively. They can be clearly distinguished from the high affinity ligands (Eb < -7 kcal/mol) and the medium affinity ligands (Eb in the range -7/-6 kcal/mol). This demonstrates that the binding domain of this drug transporter does not behave like a "sticking paper", which would non-selectively bind any molecule hydrophobic enough to reach its transmembrane segments. In contrast, it actually presents some molecular characteristics that allow distinguishing between all its potential ligands. This justifies quoting it as endowed of a "multispecific" recognition capacity.

As a matter of fact, this multispecific transporter shows a striking capacity to handle various unrelated chemical structures. Indeed, among the 18 tested molecules, 14 were found as high affinity ligands and one medium affinity ligand. If we also consider an additional series of 13 anthelmintic drugs that we recently tested for docking according to the same *in silico* procedure, including 7 macrocyclic lactones (related to ivermectin) and 6 drugs of other classes, we found 11 high affinity ligands (including the 7 macrocyclic lactones) and 2 medium affinity ligand (manuscript under review, Part I, B). This means a total of 25 high affinity ligands and 3 medium affinity ligands out of 31 tested molecules. This strongly reinforces the vision initially given by Jin et al. (2012) (claiming 5 "positive drugs" for 30 compounds tested on ATPase) of Cel-Pgp-1 being a multidrug transporter.

b) Importance of the inner chamber for ligand recognition

In agreement, and actually confirmation, of the result coming from the validation set which established that the relevant drug binding domain is the inner chamber, the supplementary series of tested molecules showed that it indeed plays a central role in multidrug recognition because all the considered lowest energy clusters of poses are positioned within this inner chamber. Conversely, the two very weak or non-ligands, CAM and DNR, showed some of their lowest energy clusters (CAM1, and DNR2 and DNR3, at around -5.5 kcal/mol) located outside the chamber, in a similar fashion to that of the cluster RHO1. These ectopic docking positions were detected because we chose to consider an exploration box large enough to contain all the residues of the TM segments, thus leading to docking poses of molecules that

interact with some of these residues while protruding for most of their surface out of the inner pocket. As these external poses were found only for very weak or non-ligands, they give the illustration that the energy range ≥ -5.5/-5.0 kcal/mol corresponds to non-specific binding, while docking poses for Eb ≤ -6/-7 kcal/mol can be considered with confidence as specific.

The first and pivotal property of the inner chamber constituting the multispecific binding domain is its large size. It can actually be envisioned by performing the superimposition of all the docked ligands, realizing its "internal moulding" (see Fig. 10). When we additionally superimposed the 13 anthelmintic drugs *in silico* tested elsewhere (David et al., 2016), the global superimposition was similar (not shown), showing that the structural diversity of the 15 ligands here reported for their relevant docking is sufficient to describe the whole multispecific binding domain of Cel-Pgp-1. This image allows pointing out three molecular characteristics:

(i) Small (< 500-550 Da) as well as larger (as compared to the chamber size) ligands can accommodate a set of anchoring points among all the offered interacting residues lining the chamber. Their interaction was actually done by different manners for each of them, sometimes giving the possibility of binding stoichiometry of 2. This multiplicity of binding solutions explains well the multispecific recognition capacity of the transporter.

(ii) The wide chamber of Cel-Pgp-1 does not present any sign of symmetry with respect to TMD1/TMD2 that limit it. This is consistent with the few observed cases of double stoichiometry that do not result from a symmetrical binding onto each of the two halves of the pseudo-symmetrical tandem structure (typical of the full-size ABC transporters).

(iii) Regarding its shape, the chamber appears constituted by two close sub-domains, BD1 and BD2, separated by an intermediate constriction zone. These two sub-domains, situated at the level of the cytosolic leaflet and in the core of the membrane, accommodate large (≥ 1000 Da) and smaller ligands, respectively. This 3D, "bi-lobulated" arrangement of the inner chamber is strikingly reminiscent of the double pharmacophoric model previously published (Garrigues et al., 2002). The latter was established from enzymological data on a mammalian Pgp and 3D ligand superimpositions, and revealed consensus recognition elements and mutual relationships of a set of specific ligands (although it was built without any structural information on the receptor multispecific binding domain). The possible relationships between nematode and mammalian Pgps will be discussed below.

Taking into account the protein transconformation that leads from "open inward-facing" to "closed inward-facing" and then "outward-facing" during its enzymatic transport cycle, it can be hypothesized that, for steric reasons, the large ligands bound onto the BD1 sub-domain will

probably have to transfer to another binding site to allow the TMDs to get closer and eventually induce their release to the acceptor compartment (while the BD2 ligands will possibly be able to directly be released once their affinity will have dropped during this transconformation). This could be correlated with the observation by Jin et al of the highest ATPase stimulation ratios (factor 8-10) for the three largest ligands of the validation set, ACD, VLN and PCT (all three binding on BD1), meaning that they are the tested substrates that are translocated the most rapidly by the active transporter.

c) Ligand access and binding

A key point in the molecular mechanisms of multidrug recognition by Cel-Pgp-1 is the accessibility of its ligands to the putative multispecific drug binding domain, i.e. the inner chamber. In particular, the short N-terminal helices TMa and TMb form a helical hairpin in Cel-Pgp-1, that is not present in many Pgp homologs, including mammalian Pgp. They thus close one lateral opening between the membrane lipid phase and the inner chamber, between TM4 and TM6. Therefore, only the other lateral aperture, on the opposite side between TM10 and TM12, remains available as an access pathway for the various hydrophobic ligands to their specific binding domain. Incidentally, this lateral aperture presents a rather favorable configuration for ligand entrance due to the bending of TM10. Indeed, to model ligand pathway, two possible access channels were computed by Mole 2.0 and found between TM10 and TM12, with a channel radius of about 3 Å (not shown). This can mediate the access of cumbersome ligands that display sufficient flexibilities, such as VLN, CSP and ACD.

Our series of 18 tested ligands encompasses a large spectrum of molecular sizes, covering the range 314 to 1255 Da. The validation set first gave the indication of a fair correlation between binding energy Eb and molecular weight (Table 1), but this was not confirmed by the whole series. Indeed, some small ligands, such as CCH, CST and HST (all around 400 Da), displayed good Eb values (all around -9 kcal/mol), whereas CAM is a large molecule (≈ 1000 Da) that was found to be a very weak affinity ligand to non-ligand (Table 3 and 5). In the same line, the two structurally close molecules, DXR and DNR, gave very different Eb values. This absence of a "size effect" confirms that the binding strength of a given ligand mainly depends on the nature of the interactions with the protein, and relatively very few on the number of interacting residues. In particular, H-bonds specifically contribute to the energy score calculation, but they all do not have the same weight depending on the distance between the donor and the acceptor of hydrogen found by AutoDock.

In contrast, there seems to be a ligand size effect when considering the location of ligand binding, in particular the distribution between the two sub-domains BD1 and BD2. Indeed, all the large ligands tested, i.e. ACD, VLN, CSP and VSP (all > 1000 Da), as well as the cluster PCT1 (850 Da) bound to BD1. Interestingly, all these five docking positions (and also KTC2) interacted with the residue K30 born by TMb, pointing out the structural importance of this small hairpin helix, supplementarily found in Cel-Pgp-1. In addition, among the 6 non-ML anthelmintic drugs in-silico tested elsewhere for docking, only the largest one, emodepside (1119 Da), bound on BD1 (manuscript under review, Part I, B). However, such agreement was not further confirmed when considering the docking of the series of 7 MLs tested, which all converged to bind in BD2 with good Eb values. It thus seems to be necessary to take into account ligand flexibility to discuss these data.

d) Comparison with mammalian Pgp

As far as primary sequence conservation is concerned, Cel-Pgp-1 shares high sequence identity with four human ABC proteins, namely ABCB1/Pgp (46%), ABCB4 (44%), ABCB5 (40%) and ABCB11 (39%) which all are full transporters, whereas a lower identity was found with other members of the B sub-family that are not involved in drug transport (and are "half-transporters"), such as ABCB2 (31% and 34% for the N- and C-term half of Cel-Pgp-1, respectively). Sequence identity is even lower with multidrug transporters of the C sub-family, such as ABCC2/Multidrug Resistance Protein 2 (MRP2) and ABCC5/MRP5 (26%). This is consistent with, at least partial, structural conservation of the multispecific binding domain restricted to the B sub-family (and most likely only considering the full transporters), and obviously not extended to other ABC sub-families.

In a more detailed comparison of Cel-Pgp-1 with mammalian Pgp, we considered the 62 hotspot residues that have been collected from various experimental approaches in order to describe the mammalian Pgp residues involved in multiple drug recognition (Shilling et al., 2006; Bessadok et al., 2011). Among these residues aligned in Cel-Pgp-1, the conservation ratio with respect to human Pgp is 16/62 (=26%). However, it appears that the number of hotspot residues actually contributing in interactions with the protein that are conserved in their nature is even lower, and especially low (for the validation set: 18/87 ≈ 20% hotspot conservation) when compared to the global sequence identity of 46% (and even the TMD-restricted identity of 37%). This hotspot divergent evolution is not surprising for two transporters that had independently evolved to handle different drug diversity profiles. This could be illustrated

firstly by the different ratios found between the hydrophobic/polar residues among the 62 human hotspot residues and their counterparts in Cel-Pgp-1: this ratio is clearly lower in Cel-Pgp-1 (37/20 = 1.85) than in Has-Pgp (44/14 = 3.1), indicating that Cel-Pgp-1 is probably more suited to recognize and transport amphiphilic and globally less hydrophobic compounds than mammalian Pgp. Taking into account the relatively high fraction of interacting residues aligned on hotspot residues (87/155 for the validation set, 236/403=59% for the whole series of 15 ligands), this indicates that the shape of the binding domain is preserved through evolution, but not the functional residues lining the internal cavity, resulting in a divergence of profiles for recognized drugs.

More specifically, in the case of the 5 molecules, i.e. VBL, ACD, VRP, CCH and DNR, some residues have been determined on mammalian Pgp to participate in their specific recognition ("specific hotspot residues", respectively quoted Mv, Ma, CV, Mc and Md in Table S1) (Shilling et al., 2006). We then determined which of the corresponding aligned residues in Cel-Pgp-1 were contributing to their docking. For VBL, among the 9 specific hotspot residues determined on mammalian Pgps, only 3 are found as interacting residues in Cel-Pgp-1, amongst which 2 are conserved in terms of the nature of the residue involved. For ACD, among the 6 specific hotspot residues, none is found in Cel-Pgp-1. For VRP, the respective values are 9/3/0 for determined hotspot residues in mammalian Pgp/found interacting in Cel-Pgp-1/conserved in terms of nature; for CCH, the respective values are 13/5/1; for DNR, the respective values are 19/1/0. These very low values of conservation of interacting residues between the two transporters clearly reveal that a given ligand can be differently recognized by both transporters in their related, but qualitatively different multispecific binding chambers.

e) Versatility of ligand binding in the inner chamber, and application to rational design of an optimized ligand

Another consequence of the large size of the multispecific inner chamber is that most of the tested ligands gave two pertinent energy clusters (among the lowest energy ones, analyzed as described in Computational methods section) corresponding to two different docking positions. Two notable exceptions were the two large ligands, VLN and CSP, which each showed only one binding site (in BD1). Indeed, in some cases of small ligands, such as PRG, CCH, KTC and CST, these pairs of docking positions corresponded to two distinct, separate binding sites (all in BD2), leading to a stoichiometry of 2. This was also the case for the larger molecule PCT, which in fact bound on two sites distributed between BD1 and BD2. Otherwise, the

somehow larger molecules, such as VBL, VCR, DPM, DXR, HST and VSP, are more prone to dock to two binding sites displaying partial overlaps. This situation leads to a dual binding site, with a stoichiometry of 1 but presenting the possibility of ligand oscillation between two approximately equiprobable (since the corresponding Eb values are close) positions. Such an ambivalent binding pattern is clearly at variance with a classical ligand/receptor relationship showing a restricted specific binding site, according to the "lock-and-key" mechanism. Here, our model provides an illustration of the wide palette of interactions offered by all the possible contact residues lining the inner chamber, together with the large size of the multidrug binding domain. This explains well the multispecific capacities offered by the inner pocket of this multidrug transporter, which can apply as well to various chemical structures of tested ligands.

The number of tested ligands in our series allowed us to consider some pairs of molecules with close chemical structures, and to draw valuable insights. The couple of small ligands PRG/CST (~ 350 Da) showed well different Eb values (≈ 2 kcal/mol more negative for CST), in both cases with a stoichiometry of 2, but different binding sites. The couple DXR/DNR (≈ 540 Da) showed an Eb value much less negative (> 3kcal/mol) by alteration of the H-bond pattern for DNR docking, accompanied by a clear change of their second docking positions (the clusters DXR1 and DNR1 having quite the same position). At variance, the couple of larger ligands VBL/VCR (≈ 820 Da) showed quite the same Eb value, with docking sites only slightly changed. Finally, the couple of very large ligands CSP/VSP (≈ 1210 Da) showed comparable Eb values with quite the same binding sites. Such a ligand size effect on these structure-binding relationships indicates the high difficulty to apply a classical QSAR approach, especially for the relatively small ligands (< 600 Da), in order to study chemical derivatives of a given ligand. Indeed, due to the richness of possible interactions with the various anchoring points within the inner chamber, a slight chemical modification may well lead to a marked alteration of the binding mode of the considered compound. This hence impedes the rational optimization of a the chemical structure of a ligand to improve its binding characteristics.

Conversely, however, these data prompted us to pay more attention to the large (> 900-1000 Da) ligands. First, their binding modes seem much less sensitive to small chemical variations. Furthermore, an important parameter appears to be their structural flexibility. Indeed, the large cyclic (pseudo)peptides, such as VLN, CSP and VSP, as well as emodepside (manuscript under review, Part I, B), known to be highly conformationally flexible, all display good Eb values (around -10 kcal/mol). In contrast, the large polycyclic, rather rigid, molecules tested display either remarkably bad, i.e. CAM, at -5.3 kcal/mol, or excellent binding properties,

such as ivermectin and derivatives, at around -12/-13 kcal/mol (manuscript under review, Part I, B) and ACD at about -16 kcal/mol. This striking observation can be interpreted by the fact that a large flexible molecule will always find, by fully exploring its conformational space, a suitable binding site, especially in BD1, whereas a large rigid molecule will encounter much more constraints to find a possible combination of a large number of suited interacting residues, at variance with smaller ligands that will most often find a binding solution. Thus, this will lead to a clearcut situation of either "frustrated" non-ligands or "excellent" ligands, particularly well-fitting a large portion of the inner chamber walls (both considering its shape and its residue composition). As a consequence, we can propose that the prototype of an optimized ligand of highest affinity would be a large, rather rigid and polycyclic molecule (sufficiently hydrophobic to ensure a satisfying membrane partition), on which it could be expected to possibly perform slight chemical structure improvements.

f) Biochemical/enzymological applications

(i) Fluorescent substrates : In a perspective of *in vitro* biochemical/enzymological experiments, our data strongly indicate that for performing transport measurements of a fluorescent substrate mediated by Cel-Pgp-1, present either in cultured (heterogeneously transfected) cells or in membrane vesicles, one should prefer to use HST rather than CAM or RHO. For RHO, it is worth mentioning that, since it is a very broadly used fluorescent substrate for mammalian Pgp, it has also been proposed for testing transport function of some nematode Pgps. However, this should not apply for Cel-Pgp-1, because RHO was found both non-activating ATPase *in vitro* and weak- to non-ligand *in silico*, and this could illustrate an aspect of functional difference with mammalian Pgp. By the way, RHO has been reported to stimulate ATPase activity of Cgr-Pgp with an apparent Km of about 40 µM (Eytan et al., 1997), with actually reflects a not so good affinity, especially considering its fair hydrophobicity. Thus, from a quantitative view, it could be speculated that such a low affinity site on mammalian Pgp could have easily evolved to a "degenerated", non-specific site in Cel-Pgp-1.

(ii) Interaction with cholesterol: Our data clearly identify a membrane lipid component, cholesterol, as a good ligand for Cel-Pgp-1. This lipid is not an obligatory endogenous constituent of the cell membranes in nematodes, which depend on its environmental conditions, as there are auxotrophic for sterols. Our result however brings the indication that, when it is present in the protein environment, it can interfere with drug handling by the transporter by a competitive mechanism. Indeed, for ligands binding on BD2 and presenting a less negative Eb

value (> -9/-8 kcal/mol), they will likely be sensitive to cholesterol presence in the membranes (the observed competition of course actually depending on their respective local concentrations). Alternatively, drugs binding either on BD1 or presenting a highly negative Eb value (< -10 kcal/mol) should be unsensitive to cholesterol presence.

In conclusion, our model of multispecific drug recognition and binding provides valuable insights on the molecular mechanisms involved, which appear clearly different from the classical specific ligand-receptor dogma, and gives a new vision of the inner chamber of this multidrug ABC transporter of B sub-family. Actually, this model extends some previous reports addressing the structural bases for multispecific drug recognition by soluble proteins (Mariuzza, 2006), which were mainly regulators of expression of multidrug transporters such as QacR (Schumacher et al., 2004), BmrR (Bachas et al., 2011) and PXR (Watkins et al., 2001; Chrencik et al., 2005). In all cases, this work is expected to contribute in the elaboration of a better approach for rational drug design strategies aimed at overcoming these transporters of great pharmacological importance.

Acknowledgements: We thank Fabien Jourdan and Clément Frainay for extensive technical support, and the GenoToul Bioinformatics hardware infrastructure that was used for computing. This work was supported by the Natural Sciences and Engineering Research Council of Canada, the FRQNT Centre for Host-Parasite Interactions, Quebec, and EMIDA ERA-NET project CARES n° 11-EMID-003-02.

References

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.

Ambudkar, S.V., Dey, S., Hrycyna, C.A., Ramachandra, M., Pastan, I., Gottesman, M.M., 1999. Biochemical, cellular, and pharmacological aspects of the multidrug transporter. Annu Rev Pharmacol Toxicol 39, 361-398.

Bachas, S., Eginton, C., Gunio, D., Wade, H., 2011. Structural contributions to multidrug recognition in the multidrug resistance (MDR) gene regulator, BmrR. Proc Natl Acad Sci U S A 108, 11046-11051.

Bessadok, A., Garcia, E., Jacquet, H., Martin, S., Garrigues, A., Loiseau, N., Andre, F., Orlowski, S., Vivaudou, M., 2011. Recognition of sulfonylurea receptor (ABCC8/9) ligands by the multidrug resistance transporter P-glycoprotein (ABCB1): functional similarities based on common structural features between two multispecific ABC proteins. J Biol Chem 286, 3552-3569.

Boer, R., Dichtl, M., Borchers, C., Ulrich, W.R., Marecek, J.F., Prestwich, G.D., Glossmann, H., Striessnig, J., 1996. Reversible labeling of a chemosensitizer binding domain of p-glycoprotein with a novel 1,4-dihydropyridine drug transport inhibitor. Biochemistry 35, 1387-1396.

Chrencik, J.E., Orans, J., Moore, L.B., Xue, Y., Peng, L., Collins, J.L., Wisely, G.B., Lambert, M.H., Kliewer, S.A., Redinbo, M.R., 2005. Structural disorder in the complex of human pregnane X receptor and the macrolide antibiotic rifampicin. Mol Endocrinol 19, 1125-1134.

Eckford, P.D., Sharom, F.J., 2009. ABC efflux pump-based resistance to chemotherapy drugs. Chem Rev 109, 2989-3011.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32, 1792-1797.

Eytan, G.D., Regev, R., Oren, G., Hurwitz, C.D., Assaraf, Y.G., 1997. Efficiency of P-glycoprotein-mediated exclusion of rhodamine dyes from multidrug-resistant cells is determined by their passive transmembrane movement rate. Eur J Biochem 248, 104-112.

Garrigues, A., Loiseau, N., Delaforge, M., Ferte, J., Garrigos, M., Andre, F., Orlowski, S., 2002. Characterization of two pharmacophores on the multidrug transporter P-glycoprotein. Mol Pharmacol 62, 1288-1298.

Haubertin, D.Y., Madaoui, H., Sanson, A., Guerois, R., Orlowski, S., 2006. Molecular dynamics simulations of E. coli MsbA transmembrane domain: formation of a semipore structure. Biophys J 91, 2517-2531.

Higgins, C.F., 1992. ABC transporters: from microorganisms to man. Annu Rev Cell Biol 8, 67-113.

Janssen, I.J., Krucken, J., Demeler, J., von Samson-Himmelstjerna, G., 2013. Caenorhabditis elegans: modest increase of susceptibility to ivermectin in individual P-glycoprotein loss-of-function strains. Exp Parasitol 134, 171-177.

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

Li, J., Jaimes, K.F., Aller, S.G., 2013. Refined structures of mouse P-glycoprotein. Protein Sci 23, 34-46.

Loo, T.W., Clarke, D.M., 2001. Defining the drug-binding site in the human multidrug resistance P-glycoprotein using a methanethiosulfonate analog of verapamil, MTS-verapamil. J Biol Chem 276, 14972-14979.

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

Loo, T.W., Bartlett, M.C., Clarke, D.M., 2006a. Transmembrane segment 7 of human P-glycoprotein forms part of the drug-binding pocket. Biochem J 399, 351-359.

Loo, T.W., Bartlett, M.C., Clarke, D.M., 2006b. Transmembrane segment 1 of human P-glycoprotein contributes to the drug-binding pocket. Biochem J 396, 537-545.

Mariuzza, R.A., 2006. Multiple paths to multispecificity. Immunity 24, 359-361.

Martin, C., Berridge, G., Higgins, C.F., Mistry, P., Charlton, P., Callaghan, R., 2000. Communication between multiple drug binding sites on P-glycoprotein. Mol Pharmacol 58, 624-632.

Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., Olson, A.J., 1998. Automated docking using lamarckian genetic algorithm and an empirical binding free energy function. J Comput Chem 19, 1639-1662.

Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., Olson, A.J., 2009. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J Comput Chem 30, 2785-2791.

Orlowski, S., Garrigos, M., 1999. Multiple recognition of various amphiphilic molecules by the multidrug resistance P-glycoprotein: molecular mechanisms and pharmacological consequences coming from functional interactions between various drugs. Anticancer Res 19, 3109-3123.

Sarkadi, B., Homolya, L., Szakacs, G., Varadi, A., 2006. Human multidrug resistance ABCB and ABCG transporters: participation in a chemoimmunity defense system. Physiol Rev 86, 1179-1236.

Schumacher, M.A., Miller, M.C., Brennan, R.G., 2004. Structural mechanism of the simultaneous binding of two drugs to a multidrug-binding protein. EMBO J 23, 2923-2930.

Shapiro, A.B., Ling, V., 1997. Positively cooperative sites for drug transport by P-glycoprotein with distinct drug specificities. Eur J Biochem 250, 130-137.

Sharom, F.J., 2006. Shedding light on drug transport: structure and function of the P-glycoprotein multidrug transporter (ABCB1). Biochem Cell Biol 84, 979-992.

Shilling, R.A., Venter, H., Velamakanni, S., Bapna, A., Woebking, B., Shahi, S., van Veen, H.W., 2006. New light on multidrug binding by an ATP-binding-cassette transporter. Trends Pharmacol Sci 27, 195-203.

Sousa, S.F., Fernandes, P.A., Ramos, M.J., 2006. Protein-ligand docking: current status and future challenges. Proteins 65, 15-26.

Sousa, S.F., Ribeiro, A.J., Coimbra, J.T., Neves, R.P., Martins, S.A., Moorthy, N.S., Fernandes, P.A., Ramos, M.J., 2013. Protein-ligand docking in the new millennium--a retrospective of 10 years in the field. Curr Med Chem 20, 2296-2314.

Szakacs, G., Paterson, J.K., Ludwig, J.A., Booth-Genthe, C., Gottesman, M.M., 2006. Targeting multidrug resistance in cancer. Nat Rev Drug Discov 5, 219-234.

Viswanadhan, V.N., Ghose, A.K., Revankar G.R., Robins, R.K., 1989. Atomic physicochemical parameters for three dimensional structure directed quantitative structure-activity relationships. 4. Additional parameters for hydrophobic and dispersive interactions and their application for an automated superposition of certain naturally occurring nucleoside antibiotics. J Chem Inf Comput Sci 29, 163-172.

Wang, E.J., Casciano, C.N., Clement, R.P., Johnson, W.W., 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.

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, 13386-13391.

Watkins, R.E., Wisely, G.B., Moore, L.B., Collins, J.L., Lambert, M.H., Williams, S.P., Willson, T.M., Kliewer, S.A., Redinbo, M.R., 2001. The human nuclear xenobiotic receptor PXR: structural determinants of directed promiscuity. Science 292, 2329-2333.

Figures

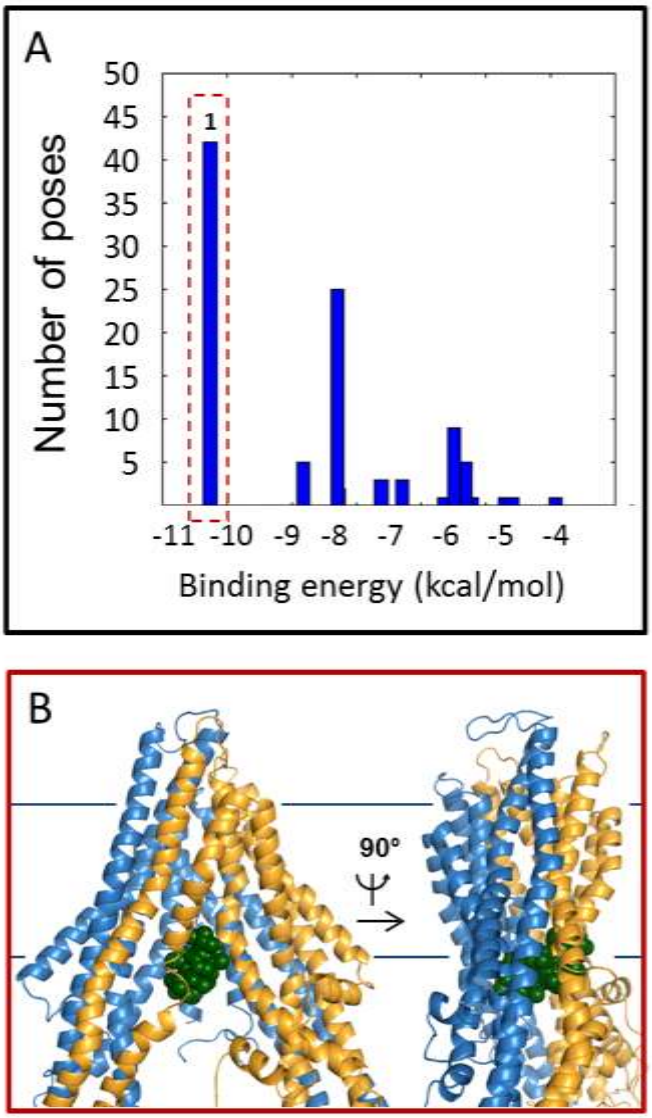


Figure 1. Valinomycin binding to Cel-Pgp-1. A. Energy clustering. **B.** Binding site of the lowest energy cluster generated using PyMol. Valinomycin is represented in dark green spheres, Cel-Pgp-1 in light blue (N-term) and yellow (C-term) ribbon.

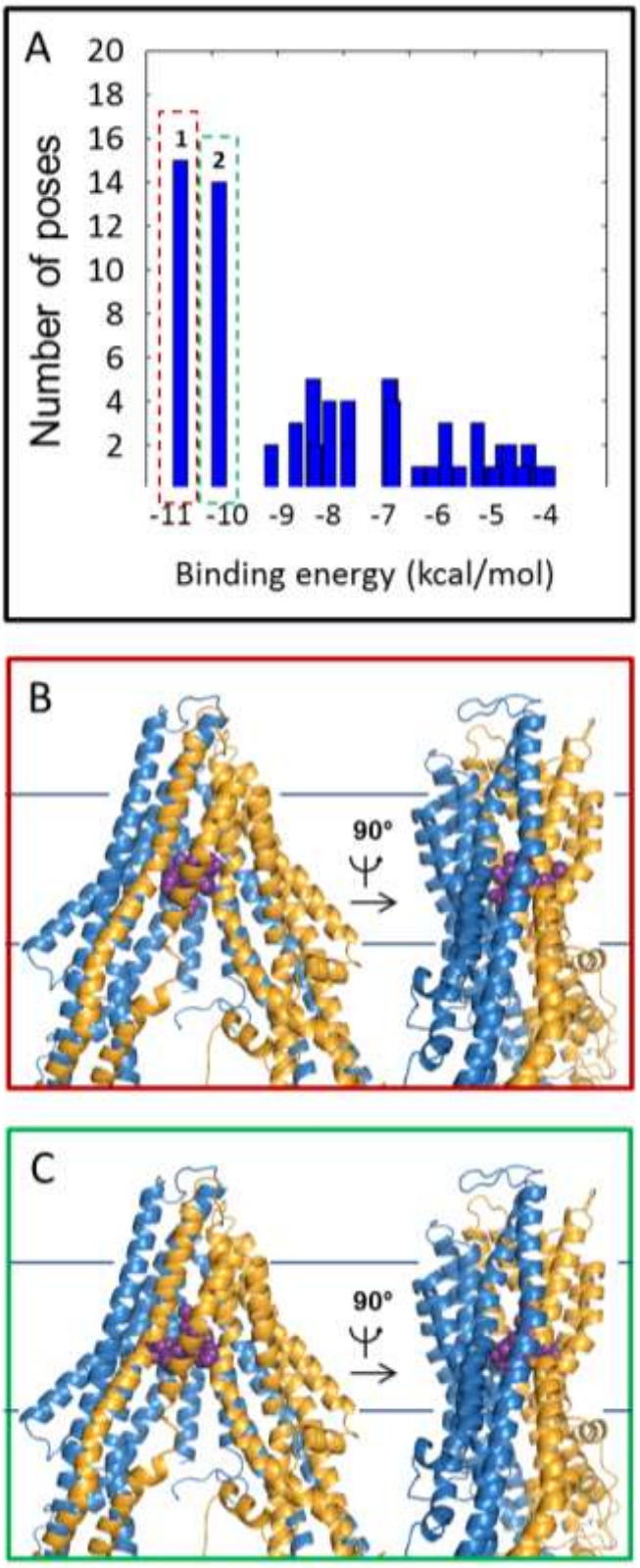


Figure 2. Vinblastine binding to Cel-Pgp-1. A. Energy clustering. **B. C.** Binding sites of the 1st (B) and 2nd (C) lowest energy clusters generated using PyMol. Vinblastine is represented in purple spheres, Cel-Pgp-1 in light blue (N-term) and yellow (C-term) ribbon.

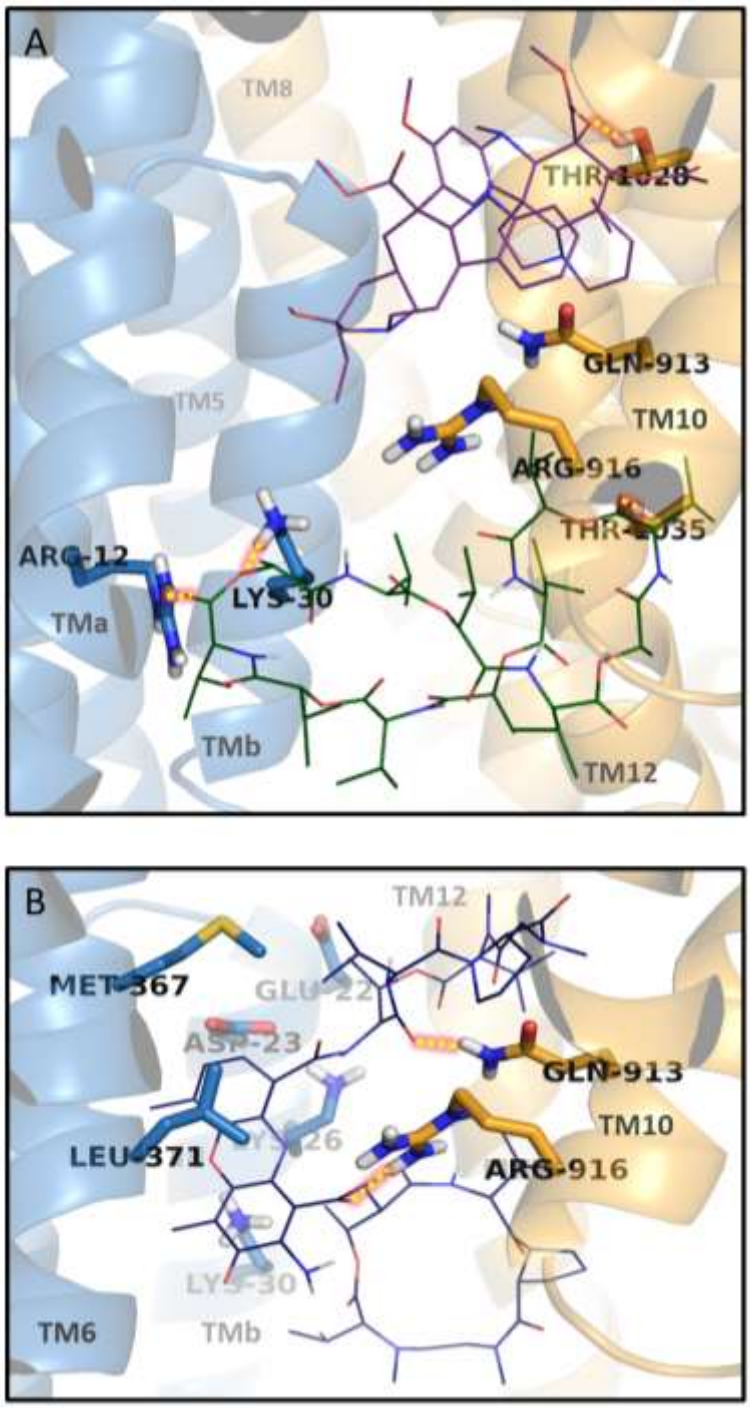


Figure 3. Key residues of Cel-Pgp-1 for interaction with substrates. A. Residues forming H-bonds with VLN1 (green lines) and VBL1 (purple lines), in addition to Q913 and R916 interacting with both molecules (G1032 cannot be visible as it does not have any secondary chain). **B.** Cel-Pgp-1 residues forming the narrow site around the polycyclic core of ACD (dark blue lines): E22, D23, K26, K30, M367; M371; Q913 and R916, the latter two forming H-bonds. Cel-Pgp-1 is represented in light blue (N-term) and yellow (C-term) transparent ribbons for TM helices and sticks for specifically-interacting residues. Atoms of Cel-Pgp-1 residues and substrates are colored in blue for N, red for O and grey for H. H-bonds are represented in yellow dotted lines with red shadow. TM helices numbers are indicated as TM# on each helix in black for closest helices and in clearer grays as the distance increases. Images were generated with PyMol.

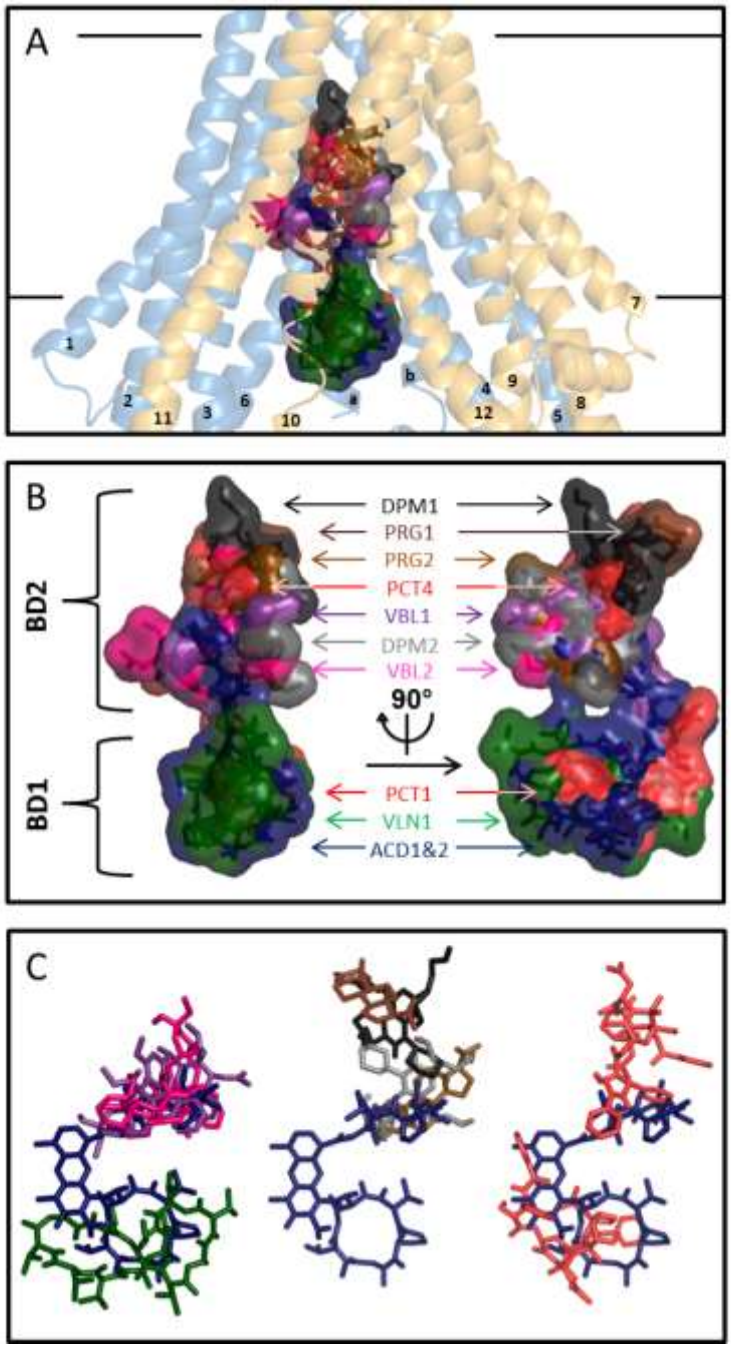


Figure 4. Binding domains characteristics of Cel-Pgp-1 substrates. A. Front view of Cel-Pgp-1 represented in transparent light blue (N-term) and yellow (C-term) ribbon, with the binding sites of the first and second lowest energy clusters of actinomycin D (ACD1 and ACD2, dark blue), valinomycin (VLN1, green), vinblastin (VBL1, purple and VBL2, pink), progesterone (PRG1, dark brown and PRG2, light brown), dipyridamole (DPM1, black and DPM2, grey), paclitaxel (PCT1 and PCT4, light red), all represented in sticks and transparent surfaces. TM helices numbers are indicated as their number only, on each helix, in gray characters. **B.** Zoom, without Cel-Pgp-1, on the front and lateral views of all substrates docking positions in the two binding sub-domains (BD1 and BD2). **C.** Overlaps between substrates: ACD1 (dark blue) represented in all panels to interact with VBL1 (purple), VBL2 (pink) and VLN1 (green) in the left panel; with PRG1 (dark brown), PRG2 (light brown), DPM1 (black) and DPM2 (grey) in the middle panel; and with PCT1 and PCT4 (light red) in the right panel. All molecules are represented in sticks. Images were generated with PyMol.

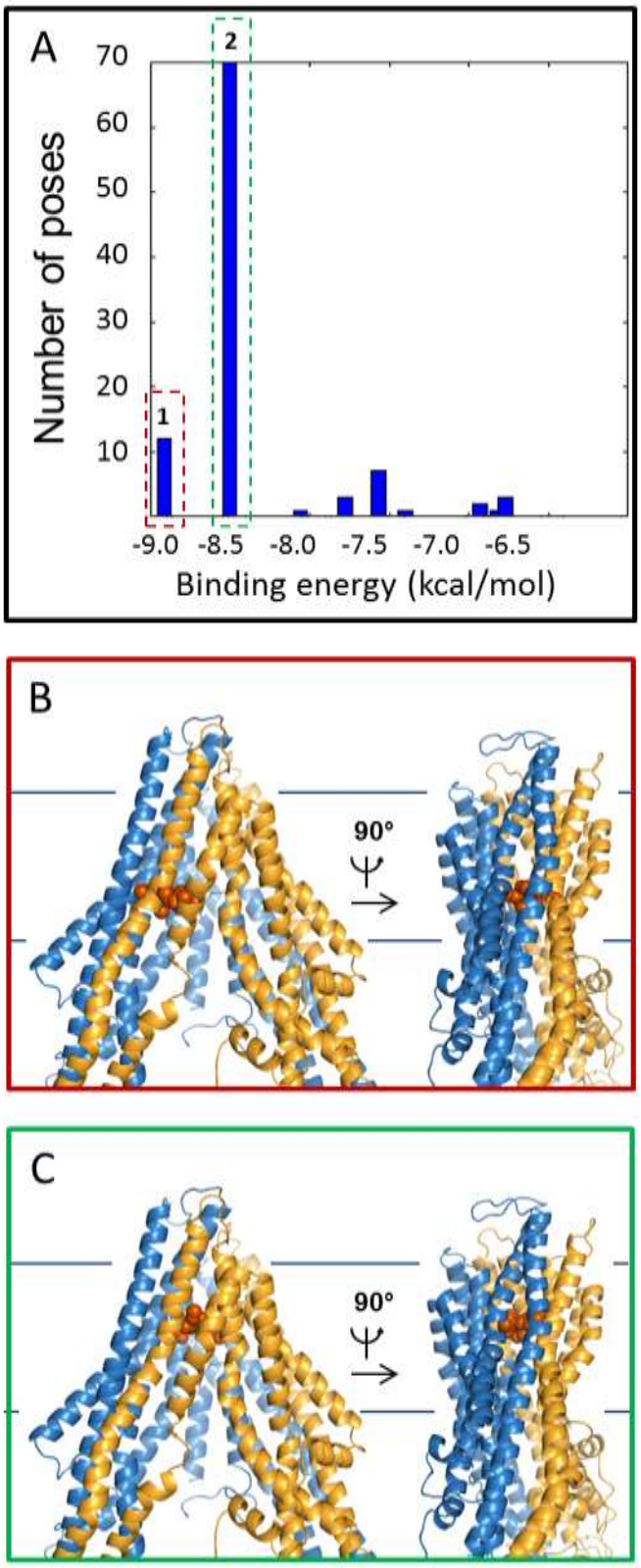


Figure 5. Colchicine binding to Cel-Pgp-1. A. Energy clustering. **B. C.** Binding sites of the 1st (B) and 2nd (C) lowest energy clusters generated using PyMol. Colchicine is represented in orange spheres, Cel-Pgp-1 in light blue (N-term) and yellow (C-term) ribbon.

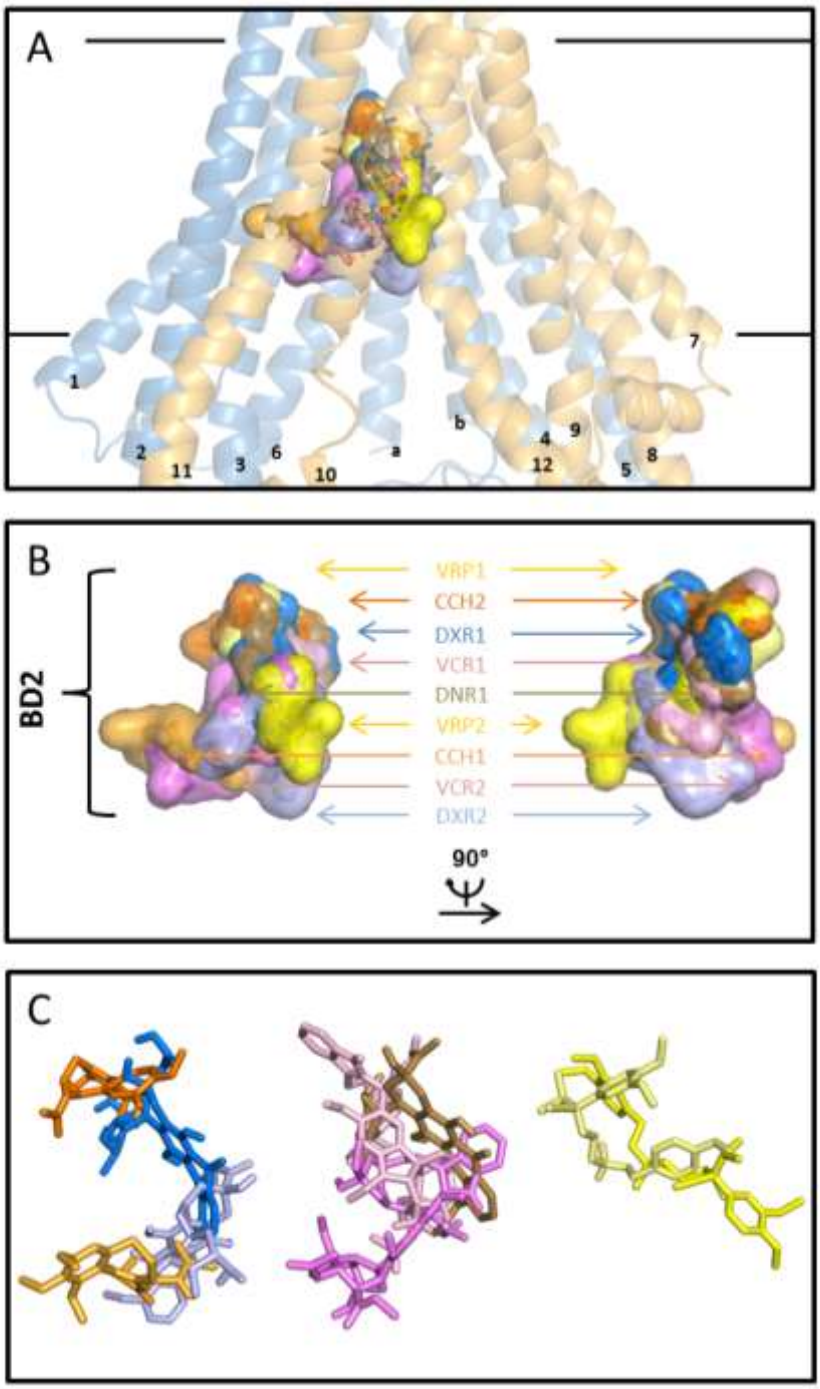


Figure 6. Binding domains characteristics of supplementary tested compounds for docking on Cel-Pgp-1. A. Front view of Cel-Pgp-1 represented in transparent light blue (N-term) and yellow (C-term) ribbon, with the binding sites of the first and second lowest energy clusters of colchicine (CCH1, light orange and CCH2, bright orange), vincristine (VCR1, light pink and VCR2, bright pink), verapamil (VRP1, light yellow and VRP2, bright yellow), doxorubicin (DXR1, bright blue and DXR2, light blue) and daunorubicin (DNR1, light brown), all represented in sticks and transparent surfaces. TM helices numbers are indicated as their number only, on each helix, in gray characters. **B.** Zoom, without Cel-Pgp-1,on the front and lateral views of all substrates docking positions in binding sub-domain 1 (BD1). **C.** Overlaps between substrates using the same colors as in panel A and B. All molecules are represented in sticks. Images were generated with PyMol.

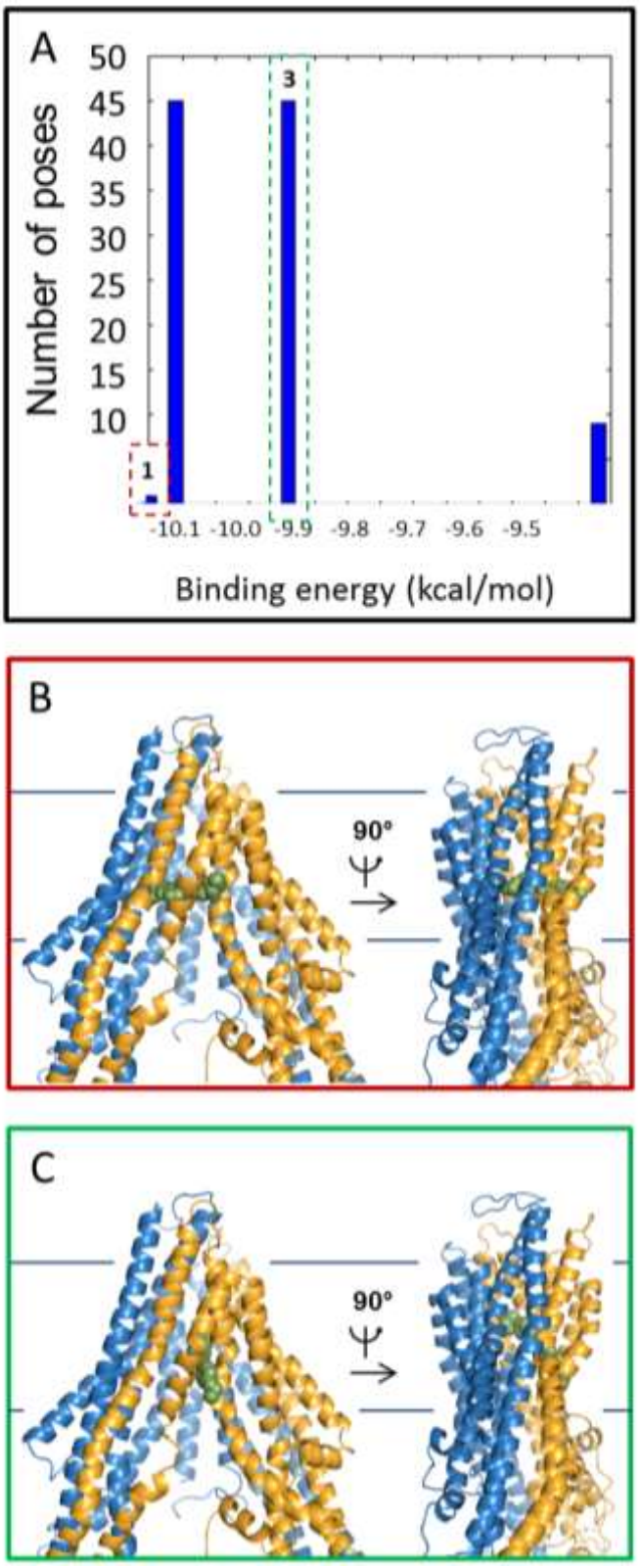


Figure 7. Hoechst 33342 binding to Cel-Pgp-1. A. Energy clustering. **B. C.** Binding sites of the 1st (B) and 2nd (C) lowest energy clusters generated using PyMol. Hoechsst 33342 is represented in light green spheres, Cel-Pgp-1 in light blue (N-term) and yellow (C-term) ribbon.

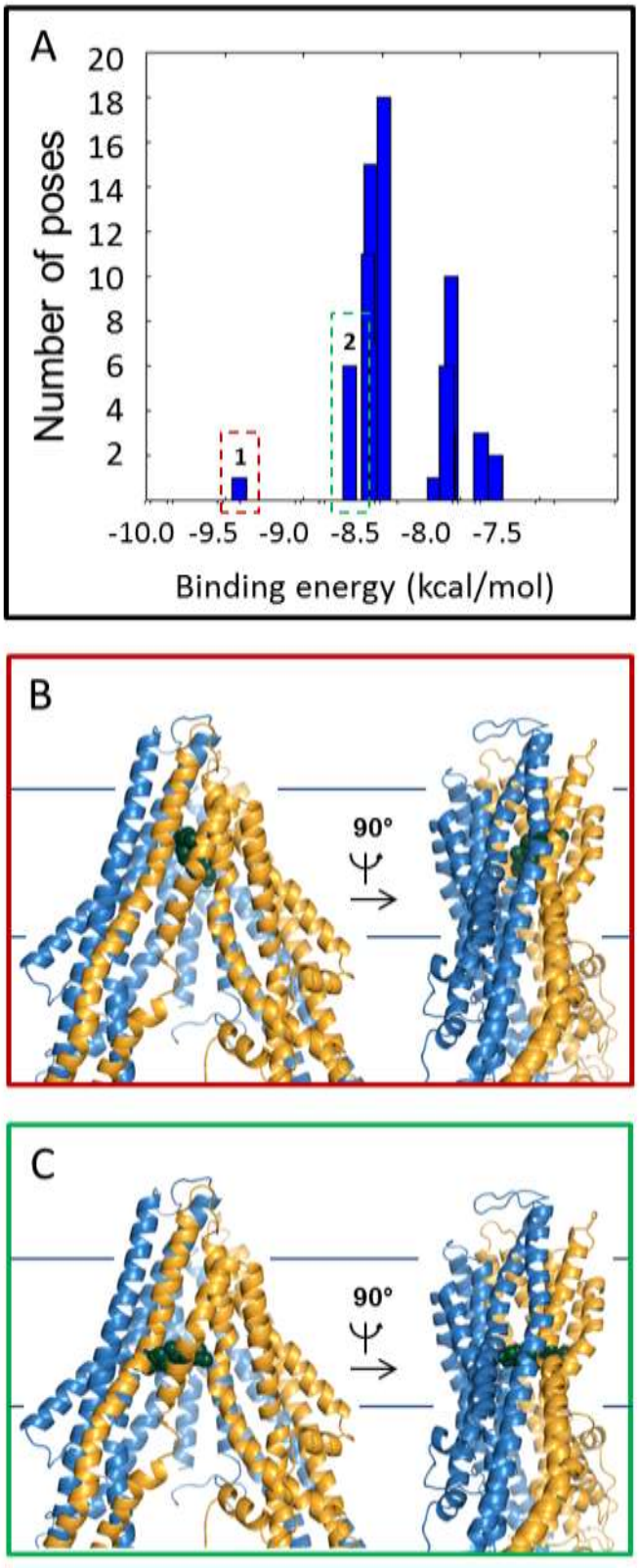


Figure 8. Cholesterol binding to Cel-Pgp-1. A. Energy clustering. **B.C.** Binding sites of the 1st (B) and 2nd (C) lowest energy clusters generated using PyMol. Cholesterol is represented in dark green, Cel-Pgp-1 in light blue (N-term) and yellow (C-term) ribbon.

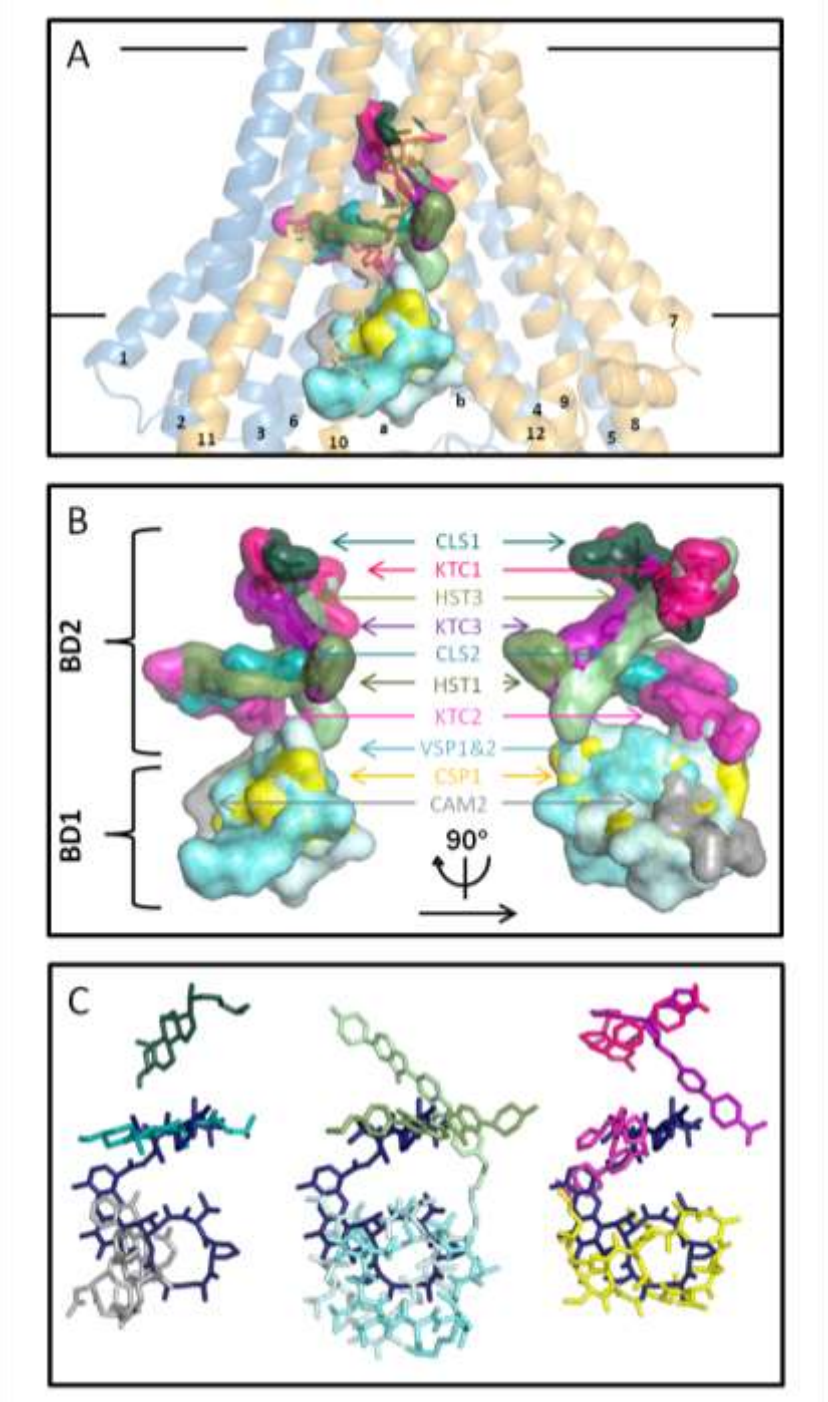


Figure 9. Binding domains characteristics of last tested compounds for docking on Cel-Pgp-1. A. Front view of Cel-Pgp-1 represented in transparent light blue (N-term) and yellow (C-term) ribbon, with the binding sites of the first and second lowest energy clusters of hoechst 33342 (HST1, green and HST2, light green), cholesterol (CLS1, dark turquoise and CLS2, light turquoise), calcein-AM (CAM2, grey), ketoconazole (KTC1, bright pink, KTC2, light pink and KTC3, purple), ciclosporin A (CSP1, yellow) and valspodar (VSP1&2, light blue), all represented in sticks and transparent surfaces. TM helices numbers are indicated as their number only, on each helix, in gray characters. **B.** Zoom, without Cel-Pgp-1, on the front and lateral views of all substrates docking positions in the two binding sub-domains (BD1 and BD2). **C.** Overlaps between substrates with the same color as in panel A and B. All molecules are represented in sticks. Images were generated with PyMol.

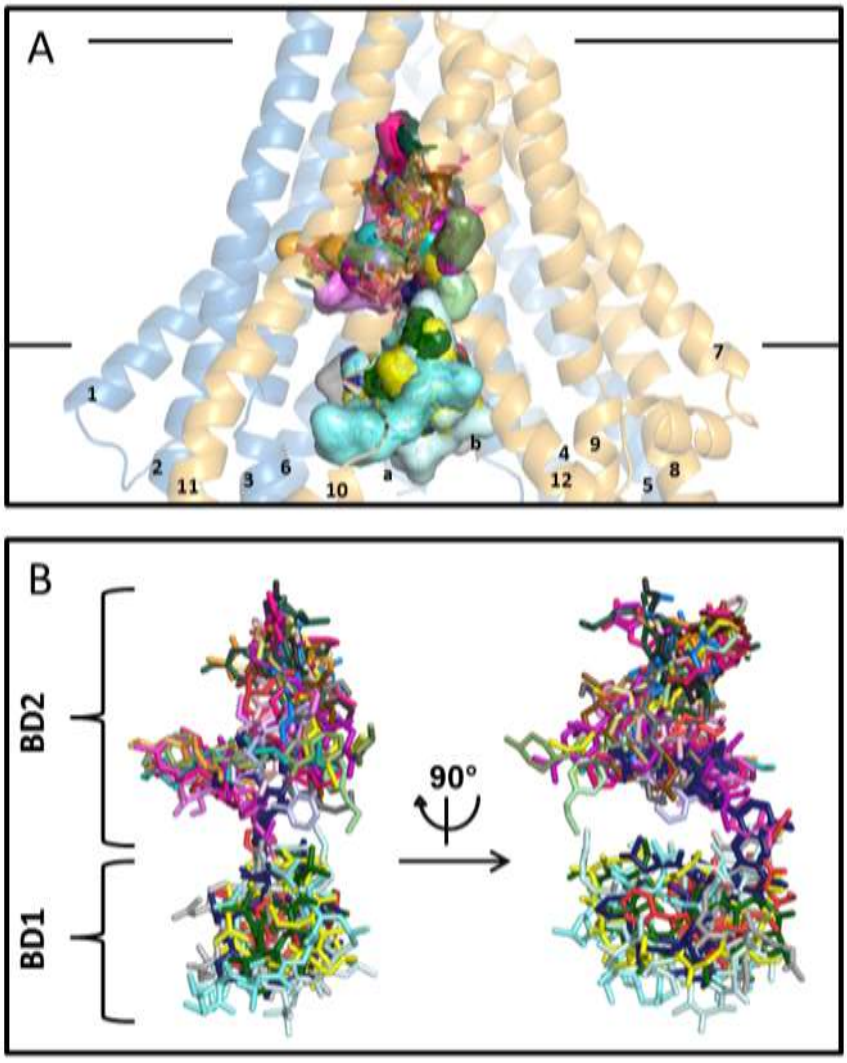


Figure 10. All binding poses of tested substrates to Cel-Pgp-1. A. Front view of Cel-Pgp-1 represented in transparent light blue (N-term) and yellow (C-term) ribbon, with the binding sites of the first and second lowest energy clusters of all compounds tested for docking, represented in sticks and transparent surfaces, with colors similar as in Figure 4, Figure 6 and Figure 9. **B.** Zoom, without Cel-Pgp-1, on the front and lateral views of all substrates docking positions in the two binding sub-domains (BD1 and BD2). All molecules are represented in sticks. Images were generated with PyMol.

Tables

Table 1. Physico-chemical, enzymological properties and docking characterization of the set of molecules tested for the validation step of the in silico procedure on Cel-Pgp-1: valinomycin (VLN), vinblastine (VBL), actinomycin D (ACD), dipyridamole (DPM), progesterone (PRG) and paclitaxel (PCT).

Molecule	VLN	VBL		ACD		DPM		PRG		PCT		RHO	
MW (Da)	1111	811		1255		504		314		854		381	
logP [a]	5.9	3.7		-1.2		1.9		4.6		3.3		3.3	
Fold activation of Cel-Pgp-1 ATPase [b]	8	3		10		4		4		8		1	
Cluster rank	1	1	2	1	2	1	2	1	2	1	4	1*	4
Binding Energy (kcal/mol)	-10.3	-10.5	-9.9	-17.0	-14.9	-7.1	-7.0	-7.2	-7.2	-8.3	-7.5	-5.0	-3.8
Binding Domain	1	2	2	1	1	2	2	2	2	1	2	-	2
Nb of poses	42	15	14	1	1	3	4	13	85	1	1	43	10
Nb of inter-acting residues	14	16	12	19	19	13	12	10	9	14	17	10	12
Nb of hotspot residues	3	10	7	8	7	11	7	10	6	3	15	1	11
Nb of H-bonds	4	2	1	2	0	2	4	2	0	2	0	3	2

(a) as calculated using Marvin Sketch with the consensus method (b) as reported in Jin et al, 2012. (*) cluster not positioned whithin the inner chamber.

Table 2. List of interacting residues of each transmembrane helix of Cel-Pgp-1 with the indicated lowest energy clusters of valinomycin (VLN), vinblastine (VBL), actinomycin D (ACD), dipyridamole (DPM), progesterone (PRG) and paclitaxel (PCT).

Molecule	VLN	VBL		ACD		DPM		PRG		PCT	
Cluster rank	1st	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	4th
TMa-b	L11 R12 K30	 D23 K26 	 	 D15 D23 K26 K30 E33	 D15 E22 D23 K26 K30 	 P21 E22 	 E22 K26 	 	 E22 K26 	 R12 D15 S18 D23 K26 K30 	 E22
TM1		**L91**	**L91**								**L91** **M94** **Q98**
TM3	K213		Q219 Q223								
TM5								**F334**			
TM6		**M364** **M367**		**M367** G370 L371 P374	**M367** G370 L371 P374	**M363**		**F359** **M363**		**M367** G370 L371 P374	**S360** **M363** **M364** **M367**
TM7						**Y771**		**Y771** **F775**			
TM10	Q913	L906 **V909** Q913	**V909** Q913	**V909** Q913	Q913		L906 A910		L906		L906

	Y914										
	R916	R916	R916	R916	R916					R916	
	G917				G917						
	R918				R918						
		A986		A986							
		L990	L990	L990	L990	L990		L990			L990
			Y991								
TM11									L993		L993
						N994		N994			N994
						Y998					
						M1021					
						Y1022		Y1022			Y1022
							I1024		I1024		
		T1025	T1025			T1025	T1025		T1025		T1025
						I1026					I1026
		T1028	T1028	T1028	T1028		T1028		T1028		T1028
TM12		S1029		S1029	S1029	S1029	S1029		S1029		S1029
							T1030				
							L1031				
	G1032	G1032	G1032	G1032	G1032		G1032		G1032		
	F1033			F1033	F1033		F1033			F1033	
	T1035			T1035						T1035	
	S1036				S1036					S1036	

Bold: hotspot residues. Underscored: residues establishing a H-bond. Green: common interacting residues to VLN. Red: common interacting residues to VBL. Blue: common interacting residues to ACD. Orange: common interacting residues to DPM. Purple: common interacting residues to PRG. Black: interacting residues never shared with any other substrate cluster.

Table 3. Physico-chemical, enzymological properties and docking characterization of the set of molecules tested for the validation step of the in silico procedure on Cel-Pgp-1: Vincristine (VCR), verapamil-S (VRP), colchicine (CCH), doxorubicin (DXR) and daunorubicin (DNR).

Molecule	VCR		VRP		CCH		DXR		DNR
MW (Da)	825		455		399		544		528
logP [a]	3.1		5.0		1.5		1.5		2.3
Fold activation of Cel-Pgp-1 ATPase [b]	2		2		1		1		1
Cluster rank	1	2	1	2	1	2	1	2	1
Binding Energy (kcal/mol)	-9.9	-9.8	-6.5	-6.1	-8.9	-8.5	-9.3	-8.3	-5.6
Binding Domain	2	2	2	2	2	2	2	2	2
Nb of poses	19	4	6	5	12	70	3	7	2
Nb of inter-acting residues	13	16	9	13	14	14	14	11	12
Nb of hotspot residues	10	10	8	10	6	14	12	7	10
Nb of H-bonds	3	3	2	2	1	1	4	4	4

(a) as calculated using Marvin Sketch with the consensus method. (b) as reported in Jin et al, 2012.

Table 4. List of interacting residues of each transmembrane helix of Cel-Pgp-1 with the indicated lowest energy clusters of: vincristine (VCR), verapamil-S (VRP), colchicine (CCH), doxorubicin (DXR) and daunorubicin (DNR).

Molecule	VCR		VRP		CCH		DXR		DNR
Cluster rank	1	2	1	2	1	2	1	2	1
Tma-b	E22		E22	E22			P21 E22	K26	P21 E22
TM1	M94 Q98	L91			T87 G88 L91	L91 M94 S95 Q98	Q98		Q98
TM2					M149				
TM3		Q219 Y220			Q219 Y220 Q223				
TM5						F334			
TM6	L356 S360 M363	M367 L371	S360 M363	F359		F359 S360 M364	M363		M363
TM7				Y771		Y771 F775	Y771		
TM10	Q913	L906 V909 Q913 R916		L906 V909 A910	Q913		V909	V909 G912 Q913	V909

TM11	L990	Y983 A986 S987 L990 L993	L993 N994	L990 L993	A986 S987 V989 L990 Y991	N994	L990 L993 Y998	V989 L990	L990 L993 N994
TM12	Y1022 T1025 I1026 T1028 S1029	T1025 T1028	Y1022 T1025 T1028 S1029	Y1022 T1025 I1026 T1028 S1029		Y1022 T1025 I1026	Y1022 T1025 I1026 T1028 S1029	T1025 T1028 S1029 G1032 F1033	T1025 I1026 T1028 S1029

Bold: hotspot residues. Underscored: residues establishing a H-bond. Green: common interacting residues to VCR1. Red: common interacting residues to VCR2. Blue: common interacting residues to VRP1. Orange: common interacting residues to VRP2. Purple: common interacting residues to CCH1. Brown: common interacting residues to DXR1. Pink: common interacting residues to DXR2. Black: interacting residues never shared with any other substrate cluster.

Table 5. Physico-chemical, enzymological properties and docking characterization of the set of molecules tested for the validation step of the in silico procedure on Cel-Pgp-1: Calcein-AM (CAM), hoechst 33342 (HST), cholesterol (CLS), ketoconazole (KTC), ciclosporin A (CSP) and valspodar (VSP).

Molecule	CAM		HST		CLS		KTC			CSP		VSP
MW (Da)	995		453		387		531			1203		1215
logP [a]	1.5		4.8		7.1		4.2			3.6		4.7
Cluster rank	1*	2	1	3	1	2	1	2	3	1	1	2
Binding Energy (kcal/mol)	-5.4	-5.3	-10.2	-10.0	-9.4	-8.7	-9.98	-9.96	-9.85	-10.3	-9.5	-9.2
Binding Domain	-	1	2	2	2	2	2	2	2	1	1	1
Nb of poses	1	2	1	45	1	6	1	1	15	2	6	25
Nb of inter-acting residues	13	13	15	9	13	12	17	16	13	17	18	14
Nb of hotspot residues	2	2	8	8	9	6	15	6	10	2	5	3
Nb of H-bonds	2	6	0	1	1	1	1	3	1	2	2	1

(a) as calculated using Marvin Sketch with the consensus method. (*) cluster not positioned whithin the inner chamber.

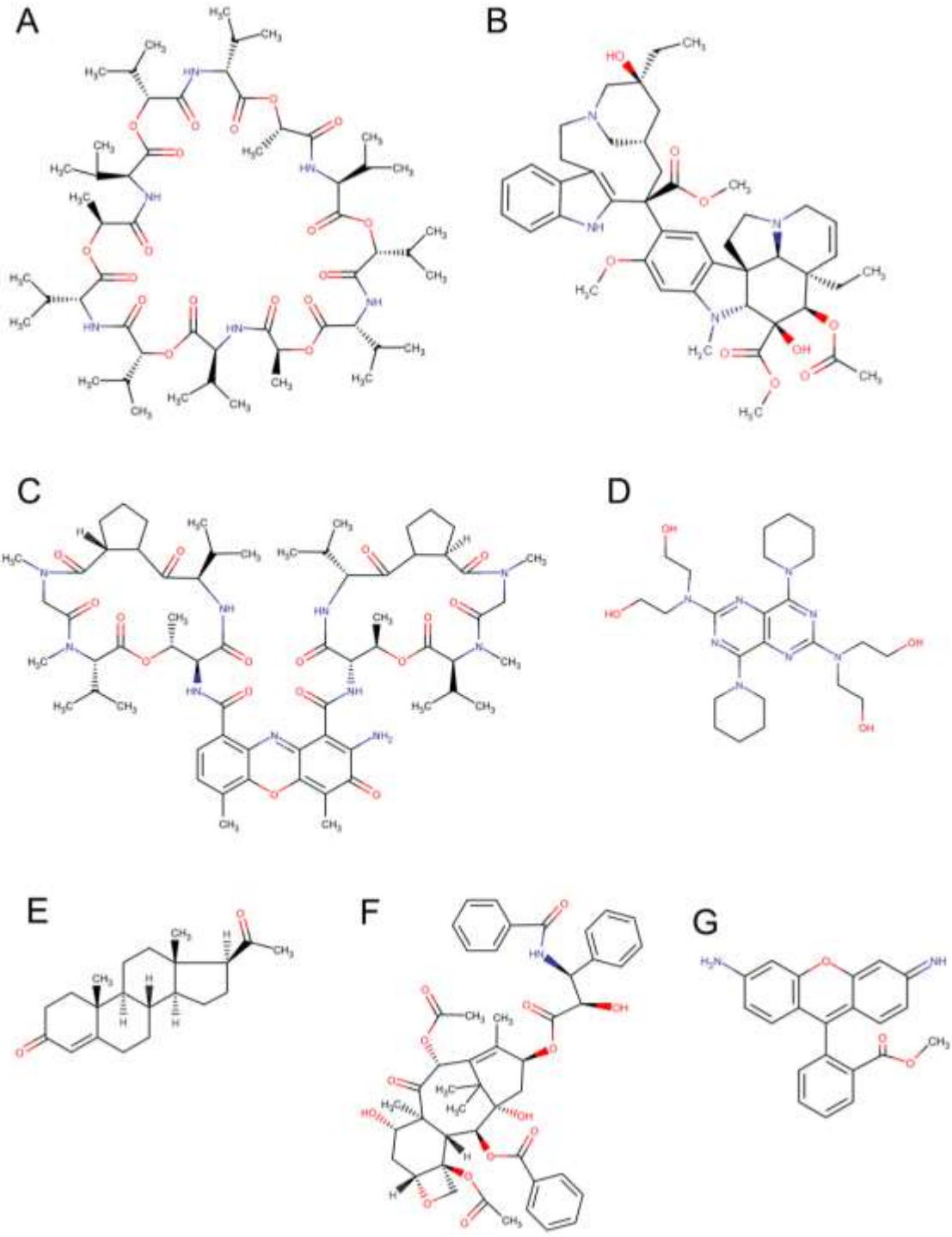
Table 6. List of interacting residues of each transmembrane helix of Cel-Pgp-1 with the indicated lowest energy clusters of: Calcein-AM (CAM), hoechst 33342 (HST), cholesterol (CLS), ketoconazole (KTC), ciclosporin A (CSP) and valspodar (VSP). Bold: hotspot residues.

Molecule	CAL	HCT		CTL			KTZ		CSP	VSP	
Cluster rank	2	1	3	1	2	1	2	3	1	1	2
N-term	R8									R8	
	L11										
	R12								R12	R12	
							D15		D15	D15	
							S18				
				P21		P21					
Tma-b				E22							
							D23				
							K26		K26	K26	K26
	K30						K30		K30	K30	K30
	E33									E33	
	D34										
		T87			**T87**		**T87**				
					G88						
TM1		**L91**			**L91**		**L91**				
								M94			
				Q98		**Q98**		**Q98**			
	K213								K213		K213
	M216										
TM3		Q219			Q219						
					Y220		Y220				
					Q223		Q223				
TM5										**N330**	
										F331	
			F359			**F359**					
						S360		**S360**			
						M363					
						M364					
TM6							**M367**				
	G370								G370		
	L371						L371		L371		
									P374	P374	
	Q375								**Q375**		**Q375**
TM7			**Y771**			**Y771**					
			F775			**F775**					
TM9				I902							
TM10		L906						L906			

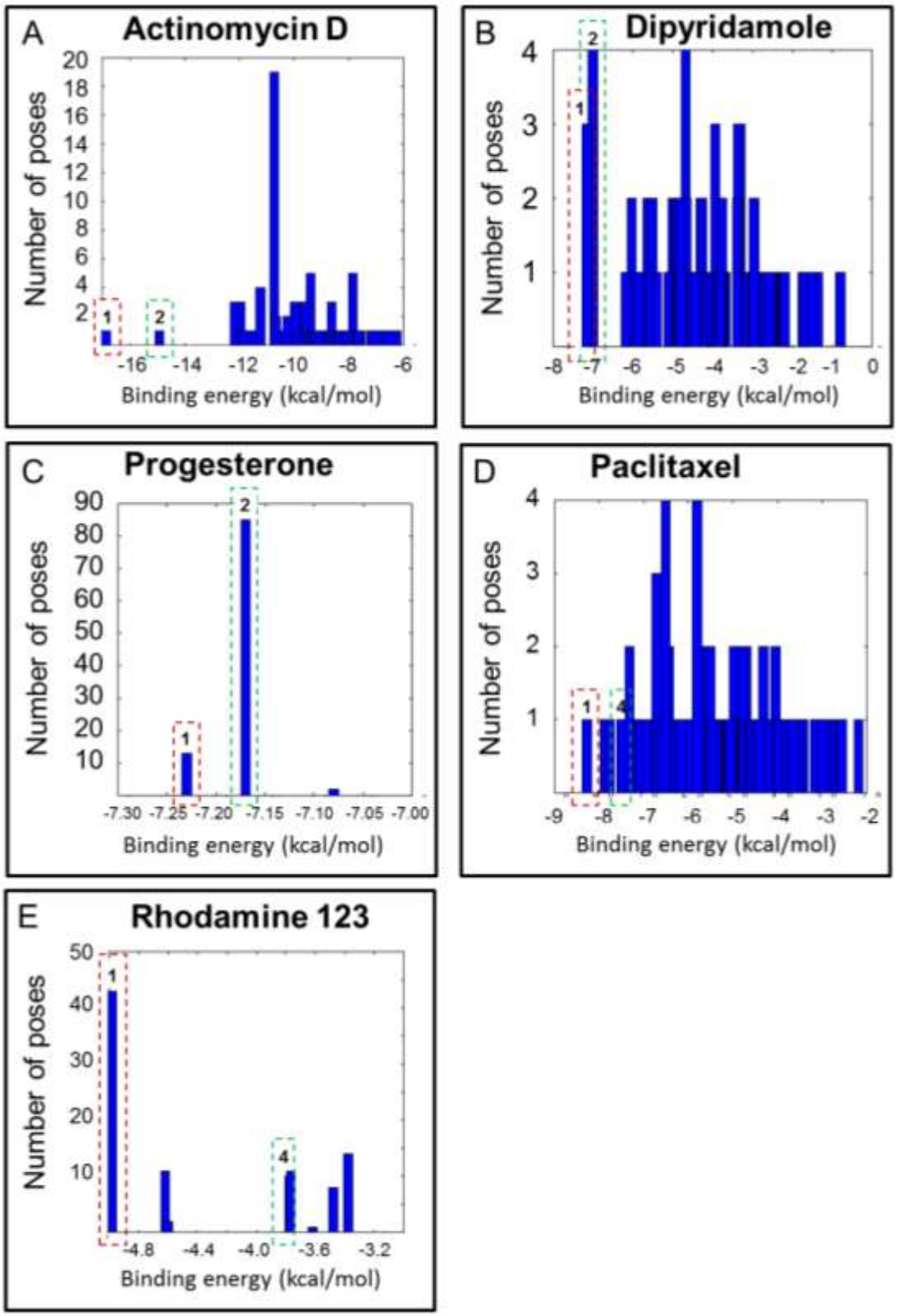
		V909		V909					
		A910		A910		A910			
		G912							
		Q913		Q913	Q913		Q913	Q913	Q913
							Y914		
	R916				R916		R916	R916	R916
							G917	G917	G917
									R918
							R919		R919
IL10-11									F920
									N924
		A986							
		S987		S987	S987				
		V989		V989					
		L990		L990	L990	L990			
TM11		Y991			Y991				
			L993	L993					
			N994	N994		N994			
			A997	A997					
			Y998	Y998		Y998			
			M1021	M1021		M1021			
			Y1022	Y1022		Y1022			
			T1025	T1025		T1025			
			I1026	I1026					
		T1028	T1028			T1028			
TM12			S1029	S1029	S1029				
		L1031	L1031			L1031			
			G1032					G1032	G1032
	F1033							F1033	F1033
							T1035	T1035	T1035
							S1036	S1036	
							P1039	P1039	
C-term								E1040	

Underscored: residues establishing a H-bond. Green: common interacting residues to CAM2. Red: common interacting residues to HST1. Blue: common interacting residues to HST3. Orange: common interacting residues to CLS1. Purple: common interacting residues to CLS2. Brown: common interacting residues to KTC1. Pink: common interacting residues to KTC2. Khaki green: common interacting residues to CSP1. Black: interacting residues never shared with any other substrate cluster.

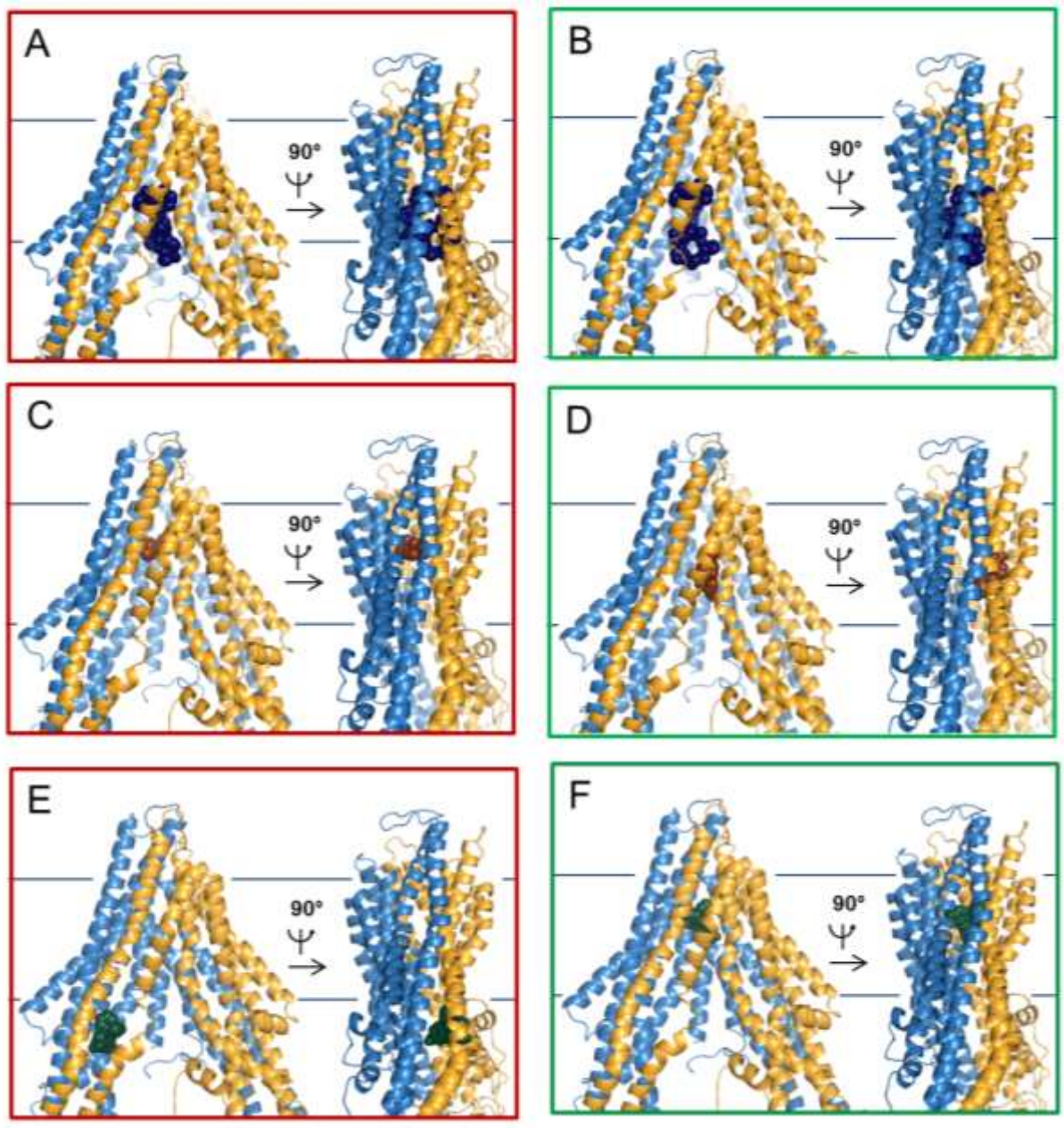
Supplementary Figures



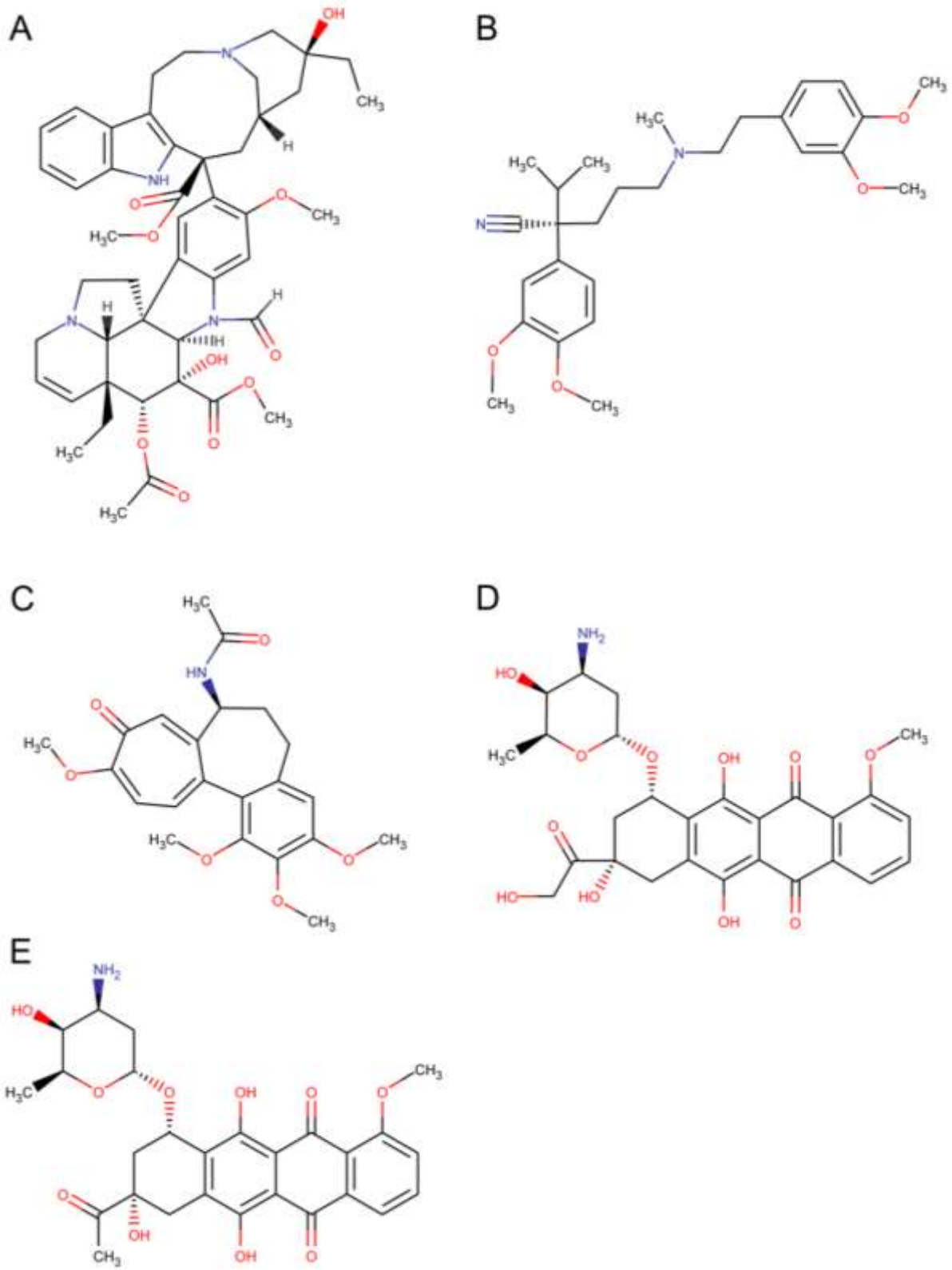
Supplementary Figure S1. Chemical structure of various substrates of Hsa-Pgp docked on Cel-Pgp-1. Valinomycin (CID 21493802) (A), vinblastine (DB00570) (B), actinomycin D (DB00970) (C), dipyridamole (DB00975) (D), progesterone (DB00396) (E), paclitaxel (DB01229) (F) and rhodamine 123 (CID65217) (G).



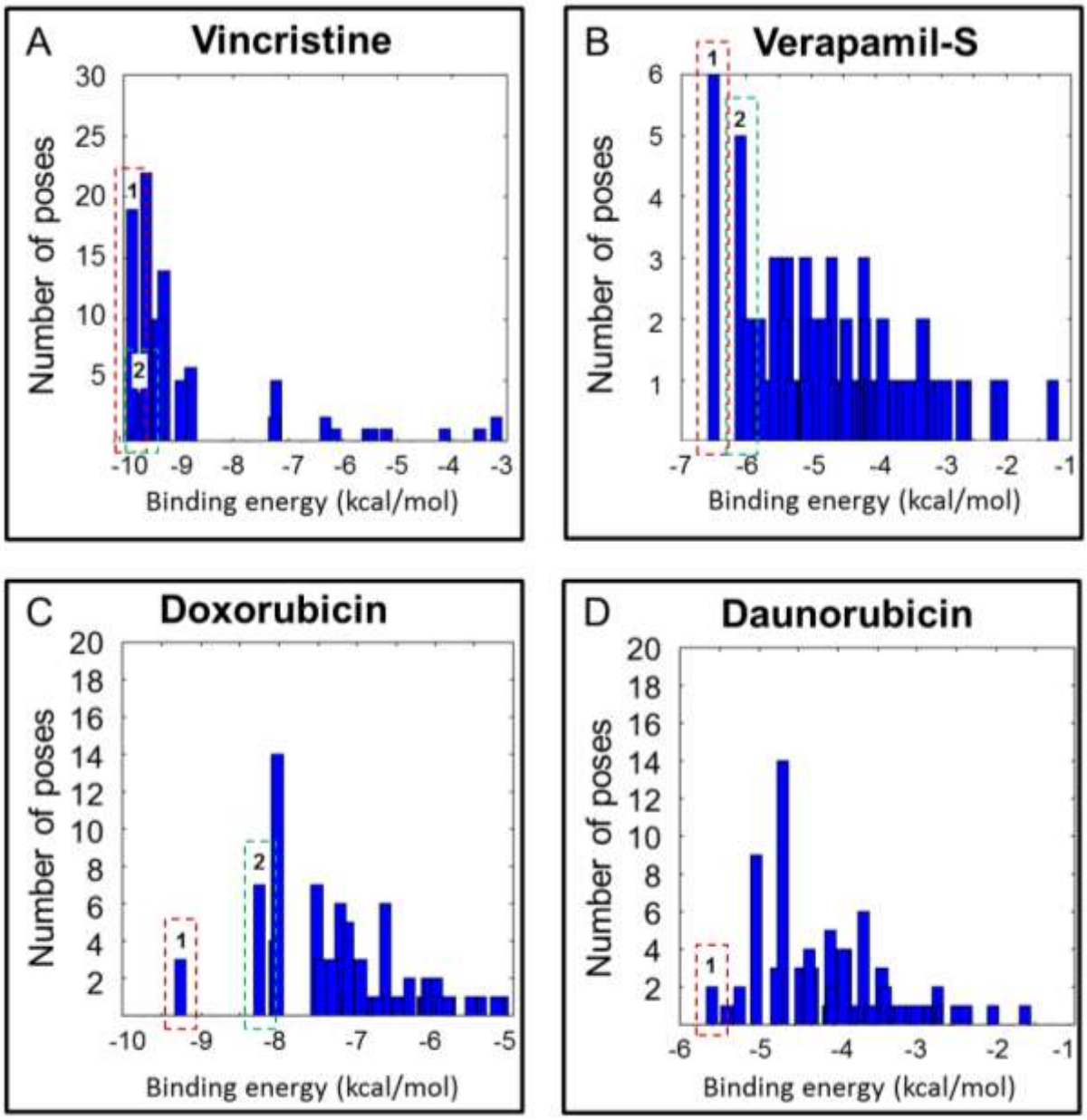
Supplementary Figure S2. Energy clustering histograms obtained for the docking of actinomycin D (A), dipyridamole (B), progesterone (C), paclitaxel (D) and rhodamine 123 (E) on Cel-Pgp-1.



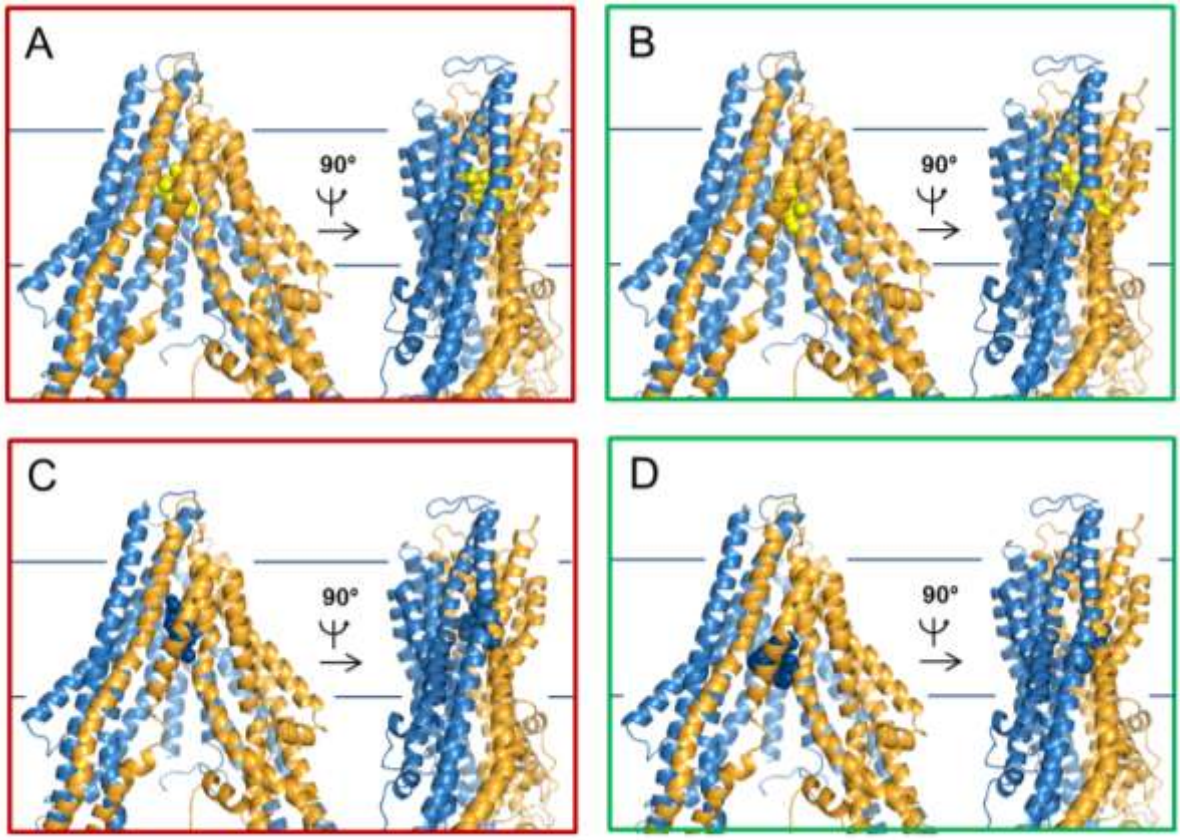
Supplementary Figure S3. Main binding sites of actinomycin D (A, B), progesterone (C, D) and rhodamine 123 (E, F) to Cel-Pgp-1. Binding sites of the 1st (A, C, E) and 2nd (B, D, F) lowest energy clusters generated using PyMol. Actinomycin D is represented in dark blue spheres, prgesterone in brown spheres and rhodamine in dark green spheres. Cel-Pgp-1 is represented in light blue (N-term) and yellow (C-term) ribbon.



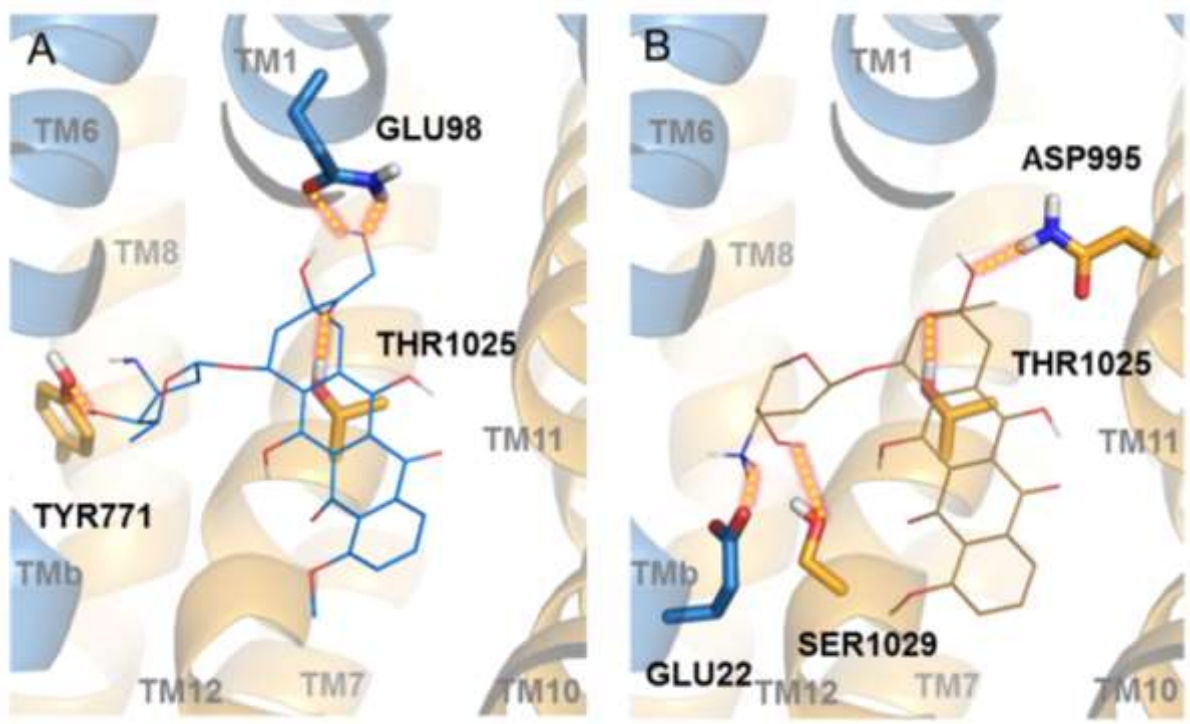
Supplementary Figure S4. Chemical structure of various substrates of Hsa-Pgp docked on Cel-Pgp-1. Vincristine (CID5978) (A), verapamil-S (CID2520) (B), colchicine (CID6167) (C), doxorubicine (CID31703) (D) and daunorubicine (CID30323) (E).



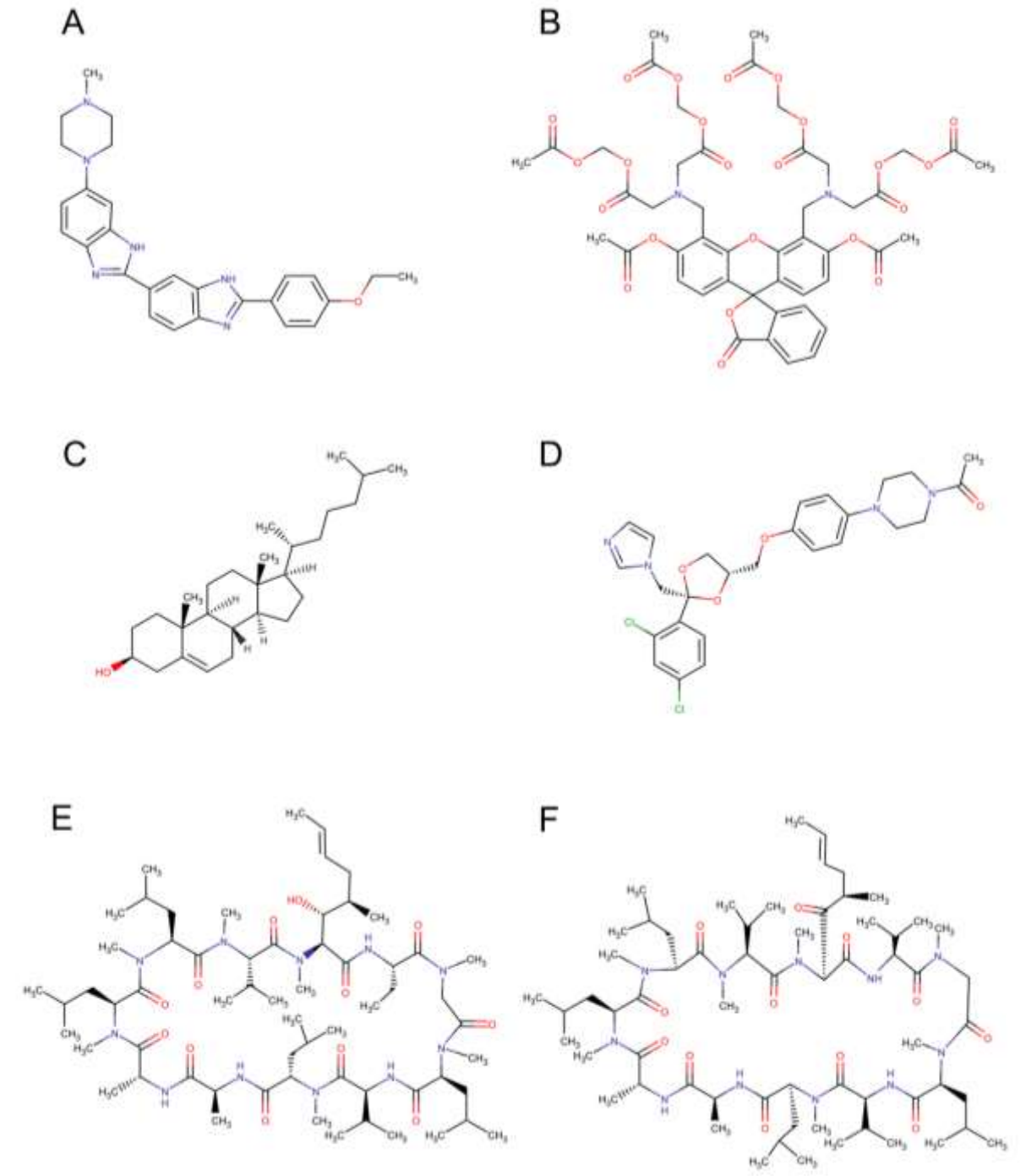
Supplementary Figure S5. Energy clustering histograms obtained for the docking of vincristine (A), verapamil-S (B), doxorubicin (C), and daunorubicin (D) on Cel-Pgp-1.



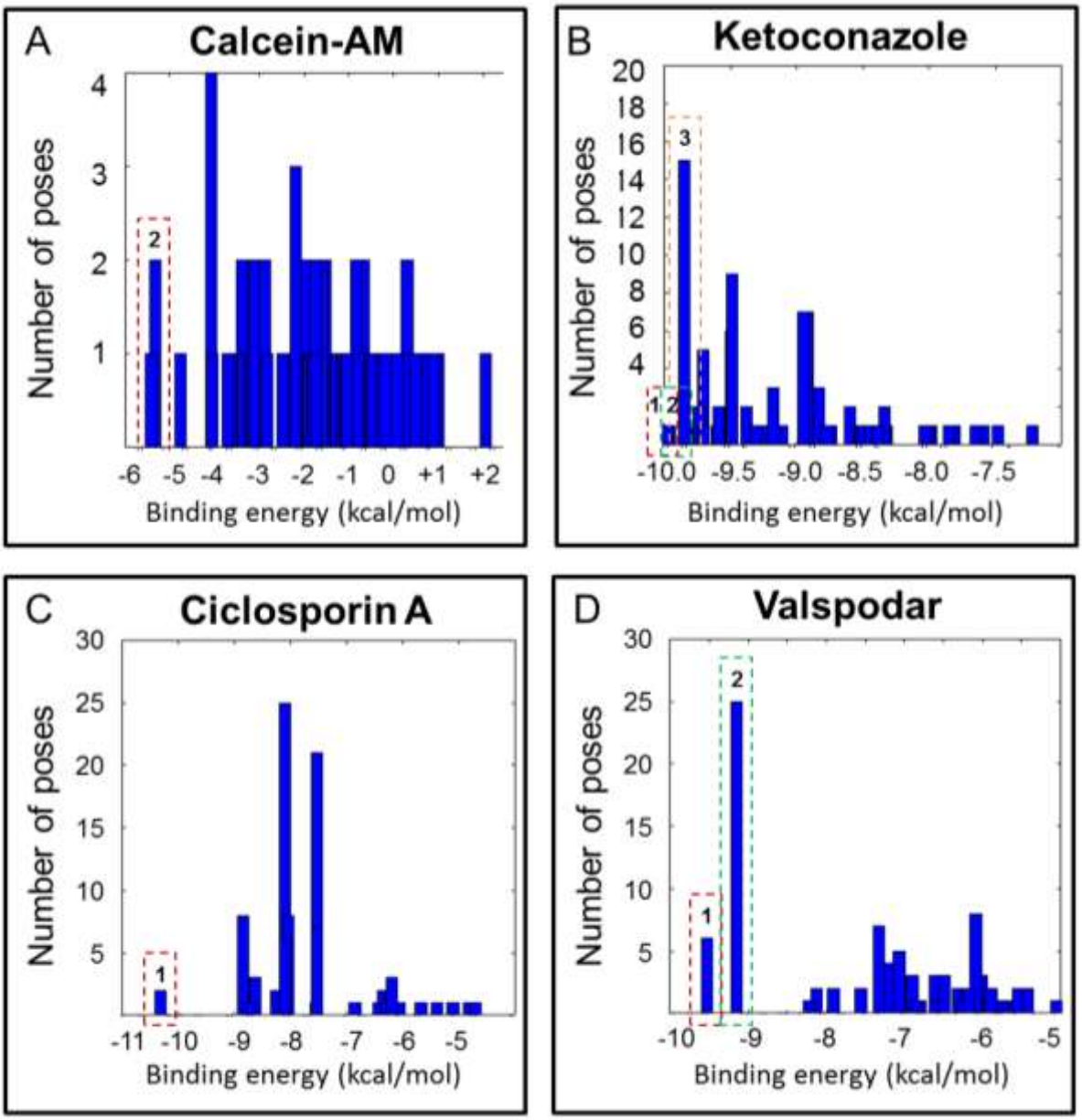
Supplementary Figure S6. Main binding sites of verapamil-S (A, B) and doxorubicin (B, C) binding to Cel-Pgp-1. Binding sites of the 1st (A, C) and 2nd (B, D) lowest energy clusters generated using PyMol. Verapamil-S is represented in yellow spheres and doxorubicin in blue spheres. Cel-Pgp-1 is represented in light blue (N-term) and yellow (C-term) ribbon.



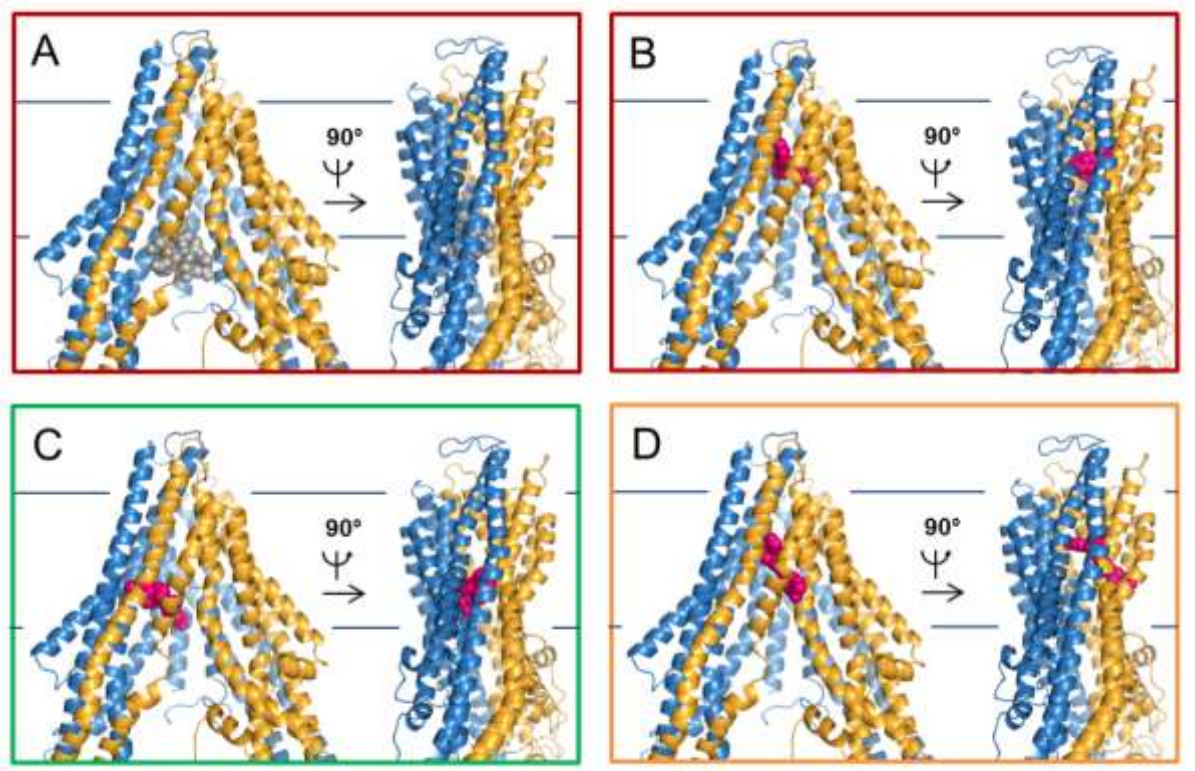
Supplementary Figure S7. Residues forming H-bonds with DXR1 (blue lines) (A) and DNR1 (brown lines) (B). Cel-Pgp-1 is represented in light blue (N-term) and yellow (C-term) transparent ribbons for TM helices and sticks for specifically-interacting residues. Atoms of Cel-Pgp-1 residues and substrates are colored in blue for N, red for O and grey for H. H-bonds are represented in yellow dotted lines with red shadow. TM helices numbers are indicated as TM# on each helix in dark gray for closest helices and in clearer grays as the distance increases. Images were generated with PyMol.



Supplementary Figure S8. Chemical structure of various substrates of Hsa-Pgp docked on Cel-Pgp-1. Hoechst 33342 (CID1464) (A), calcein-AM (CID4126474) (B), cholesterol (CID5997) (C), ketoconazole (DB01026) (D), cyclosporin A (chemspider 4447449) (E) and valspodar (chemspider 4445174) (F).



Supplementary Figure S9. Energy clustering histograms obtained for the docking of calcein-AM (A), ketoconazole (B), cyclosporin A (C) and valspodar (D) on Cel-Pgp-1.



Supplementary figure S10. Main binding sites of ketoconazole (A, B, C) and calcein-AM (D) binding to Cel-Pgp-1. Binding sites of the lowest energy cluster generated using PyMol (A, D), 2nd lowest energy cluster (B) and 3rd lowest energy cluster (C). Ketoconazole is represented in pink spheres and calcein-AM is represented in gray spheres, Cel-Pgp-1 is represented in light blue (N-term) and yellow (C-term) ribbon.

Supplementary Table

Supplementary table S1. List of "hotspot residues" that have been identified as being involved in multispecific substrate recognition in mammalian Pgp, and their aligned residues in Cel-Pgp-1.

P-gp	Hsa-Pgp	Mmu-ABCB1a	Cgr-ABCB1	Mmu-ABCB1b	Origin	Cel-Pgp-1
TM1	**H61**	H60	H60	H60	Mv	T87
	G64	A63	A63	L63	Mv	G90
	L65	L64	L64	L64	**Mv/CV/SL**	L91
	M68	**M67**	M67	L67	SL	M94
	M69	**M68**	M68	M68	S	S95
	F72	**F71**	F71	F71	SL	Q98
TM2	Y118	**Y114**	Y115	Y117	SL	Y142
TM4	**S222**	S218	S219	S221	CV	T246
TM5	I299	**M295**	M296	I298	SL	F323
	F303	**F299**	F300	Y302	SL	Q327
	L304	**L300**	L301	L303	SA	A328
	I306	I302	I303	V305	**M/SL**	N330
	Y307	**Y303**	Y304	Y306	S	F331
	Y310	**Y306**	Y307	Y309	SL	F334
TM6	**F335**	F331	F332	F334	Mva	F359
	F336	**F332**	F333	F335	S	S360
	V338	V334	V335	I337	Mv	V362
	L339	**L335**	L336	L338	**CV/S**	M363
	I340	**I336**	I337	L339	**CR/S**	M364
	G341	G337	G338	G340	M	G365
	A342	A338	A339	T341	CV	S366
	F343	**F339**	F340	F342	S	M367
	S351	D347	D348	**D350**	M	Q375
TM7	N721	**N717**	N718	N719	SL	G764
	Q725	**Q721**	Q722	Q723	S	Y768
	F728	**F724**	F725	F726	**CV/S**	Y771
	F732	**F728**	F729	F730	S	F775
TM8	L762	**L758**	L759	M760	SA	L803
	F770	**F766**	L767	Y768	SL	S811
TM9	T837	**F833**	T834	T835	SA	I878
	I840	I836	**I837**	V838	M	L881
	A841	A837	A838	A839	CR	V882
	N842	N838	**N839**	N840	M	S883
TM10	I864	I860	I861	**I862**	M	I905
	I867	I863	I864	**L865**	M	I908

	I868	I864	I865	I866	CV	V909
	I870	I866	I867	**L868**	M	F911
	A935	A931	A932	**A933**	M	A976
	F938	**F934**	F935	F936	M	Q979
	F942	**F938**	F939	F940	**M/CV**	Y983
	S943	**S939**	S940	**S941**	Mva	G984
	T945	**T941**	T942	T943	**Mv/CV**	A986
TM11	Q946	**Q942**	Q943	Q944	M	S987
	M949	**M945**	M946	M947	SL	L990
	Y950	**Y946**	Y947	Y948	Mva	Y991
	S952	**S948**	S949	S950	Ma	L993
	Y953	**Y949**	Y950	Y951	**Ma/S**	N994
	F957	**F953**	F954	F955	**Mva/SL**	Y998
	L975	L971	L972	M973	**M/CR/SL**	L1018
	F978	**F974**	F975	F976	**M/S**	M1021
	S979	**S975**	S976	S977	S	Y1022
	V981	I977	I978	V979	**M/CR**	I1024
	V982	**V978**	V979	V980	**CR/S**	T1025
	F983	F979	F980	F981	**M/SL**	I1026
	G984	G980	G981	G982	CV	S1027
TM12	A985	**A981**	A982	A983	SA	T1028
	M986	**M982**	M983	M984	S	S1029
	A987	**A983**	A984	A985	SL	T1030
	G989	**G985**	G986	G987	SA	G1032
	Q990	**Q986**	Q987	N988	S	F1033
	V991	**V987**	V988	T989	SL	A1034
	S993	**S989**	S990	S991	SA	S1036

Hotspots are distributed among the various TM helices, as quoted in the first column. The orthologs: human (Hsa), mouse (Mmu) a and b isoforms and Chinese hamster (Cgr), in which the residue has been indicated, are shown in columns 2-5 by a bold residue symbol. Hotspots were determined, in the literature, by different experimental approaches (point mutagenesis, chemical labelling, co-crystallization, respectively denoted as M, C, S), as quoted in the sixth column; when the residue has been found by at least two different approaches, the symbols are highlighted in bold. M: cytotoxicity assay after Pgp site-directed mutagenesis using, among other drugs, VBL (v) or ACD (a); C : MTS-mediated chemical labelling by a MTS-derivative of verapamil (V) or rhodamine (R) of Cys-free Pgp specifically bearing Cys residues obtained by scanning mutagenesis; S: contact residues of the co-crystallized ligands QZ59RRR and QZ59SSS in the Pgp crystal structures published in Aller et al. (2009) and Li et al (2014), but presenting few differences (A: residues found in the 2009 structure alone, L: residues found in the 2014 structure alone). See Computational Methods section for the relevant references. Last column: corresponding residues found in Cel-Pgp-1 after sequence alignment with mammalian Pgp.

B. Manuscript n°2 published in the International Journal for Parasitology: Drugs and Drug Resistance:

In silico analysis of anthelmintics binding to *Caenorhabditis elegans* P-glycoprotein 1

Marion David[a,b], Stéphane Orlowski[c], Roger K. Prichard[b], Shaima Hashem[c], François André[c,#] and Anne Lespine[a,#]

[a] Toxalim (Research Centre in Food Toxicology), Université de Toulouse, INRA, ENVT, INP-Purpan, UPS, Toulouse, France

[b] Institute of Parasitology, McGill University, Sainte-Anne-De-Bellevue, Canada

[c] Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS UMR 9198, Univ Paris-Sud, Université Paris-Saclay, 91198, Gif-sur-Yvette cedex, France

[#] Corresponding authors:

Anne Lespine: anne.lespine@toulouse.inra.fr
Tel: +33 (0)5 82 06 63 52
Fax: +33 (0)5 61 28 53 10
Address: INRA, UMR 1331 TOXALIM - Equipe E6-TMR: Transporteurs Membranaires & Résistance ; 180, chemin de Tournefeuille - BP 93173 ; F-31027 TOULOUSE Cedex 3, France

François André: francois.andre@cea.fr
Tel: +33 (0)1 69 08 44 32
Fax: +33 (0)1 69 08 87 17
Address: CEA, IBiTec-Saclay, SB2SM and UMR9198 CNRS, I2BC
91191 Gif-sur-Yvette cedex, France.

International Journal for Parasitology: Drugs and Drug Resistance xxx (2016) 1–15

Contents lists available at ScienceDirect

International Journal for Parasitology: Drugs and Drug Resistance

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

Invited article

In silico analysis of the binding of anthelmintics to *Caenorhabditis elegans* P-glycoprotein 1

Marion A. David [a, b], Stéphane Orlowski [c], Roger K. Prichard [b], Shaima Hashem [c], François André [c, **], Anne Lespine [a, *]

[a] Toxalim (Research Centre in Food Toxicology), Université de Toulouse, INRA, ENVT, INP-Purpan, UPS, Toulouse, France
[b] Institute of Parasitology, McGill University, Sainte-Anne-De-Bellevue, Canada
[c] Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS UMR 9198, Univ Paris-Sud, Université Paris-Saclay, 91198, Gif-sur-Yvette cedex, France

ARTICLE INFO

Article history:
Received 4 May 2016
Received in revised form
7 September 2016
Accepted 9 September 2016
Available online xxx

Keywords:
Anthelmintics
Macrocyclic lactones
Nematodes
Caenorhabditis elegans
ABC transporters
Molecular docking

ABSTRACT

Macrocyclic lactones (ML) are important anthelmintics used in animals and humans against parasite nematodes, but their therapeutic success is compromised by the spread of ML resistance. Some ABC transporters, such as P-glycoproteins (Pgps), are selected and overexpressed in ML-resistant nematodes, supporting a role for some drug efflux proteins in ML resistance. However, the role of such proteins in ML transport remains to be clarified at the molecular level. Recently, *Caenorhabditis elegans* Pgp-1 (Cel-Pgp-1) has been crystallized, and its drug-modulated ATPase function characterized *in vitro* revealed Cel-Pgp-1 as a multidrug transporter. Using this crystal structure, we have developed an *in silico* drug docking model in order to study the binding of ML and other anthelmintic drugs to Cel-Pgp-1. All tested ML bound with high affinity in a unique site, within the inner chamber of the protein, supporting that ML may be transported by Cel-Pgp-1. Interestingly, interacting residues delineate a ML specific fingerprint involving H-bonds, including T1028. In particular, benzofurane and spiroketal moieties bound to specific sub-sites. When compared with the aglycone ML, such as moxidectin and ivermectin aglycone, avermectin anthelmintics have significant higher affinity for Cel-Pgp-1, likely due to the sugar substituent(s) that bind to a specific area involving H-bonds at Y771. Triclabendazole, closantel and emodepside bound with good affinities to different sub-sites in the inner chamber, partially overlapping with the ML binding site, suggesting that they could compete for Cel-Pgp-1-mediated ML transport. In conclusion, this work provides novel information on the role of nematode Pgps in transporting anthelmintics, and a valuable tool to predict drug-drug interactions and to rationally design new competitive inhibitors of clinically-relevant nematode Pgps, to improve anthelmintic therapeutics.

1. Introduction

Macrocyclic lactones (ML) are the most important anthelmintic (AH) drugs used today for the control of parasitic nematodes and ectoparasites in animals and humans. Since the discovery of ivermectin (IVM) in 1980, several structurally-related ML drugs, belonging to the avermectin or milbemycin families, have been marketed. They all exert their anthelmintic effect by binding to glutamate-gated chloride channels expressed in the nervous system, causing paralysis of somatic, pharyngeal or excretory cell muscles that leads to paralysis, starvation, loss of tolerance to host immunity and death of the nematode (Cully et al., 1994; Forrester et al., 2003; Moreno et al., 2010). Overall, besides their exceptional safety margin and broad-spectrum activity, ML were previously effective against parasite strains resistant to older deworming

Abbreviations: ABC, ATP-binding cassette; ABA, abamectin; AH, anthelmintic; Cel, *Caenorhabditis elegans*; Ceg, *Cylicocylus elongatus*; Cgr, *Cricetulus griseus*; CID, Chemspider; CLO, closantel; DB, Drugbank; DOR, doramectin; EMD, emodepside; EPR, eprinomectin; Hco, *Haemonchus contortus*; Hsa, *Homo sapiens*; IVM, ivermectin; IVA, ivermectin-aglycone; LEV, levamisole; M, Merck Index; MDR, multidrug resistance; ML, macrocyclic lactone(s); Mmu, *Mus musculus*; MNP, monepantel; MOX, moxidectin; NBD, nucleotide binding domain; Oar, *Ovis aries*; PDB, Protein Data Bank; Pgp, P-glycoprotein; RMSD, root mean square deviation; SEL, selamectin; TBZ, thiabendazole; TCZ, triclabendazole; TMD, transmembrane domain.

* Corresponding author. INRA, UMR 1331 TOXALIM - Equipe E6-TMR: Transporteurs Membranaires & Résistance ; 180, chemin de Tournefeuille - BP 93173, F-31027, Toulouse Cedex 3, France.
** Corresponding author. CEA, IBiTec-Saclay, SB2SM and UMR9198 CNRS, I2BC, 91191, Gif-sur-Yvette cedex, France.
E-mail addresses: francois.andre@cea.fr (F. André), anne.lespine@toulouse.inra.fr (A. Lespine).

http://dx.doi.org/10.1016/j.ijpddr.2016.09.001

Please cite this article in press as: David, M.A., et al., *In silico* analysis of the binding of anthelmintics to *Caenorhabditis elegans* P-glycoprotein 1, International Journal for Parasitology: Drugs and Drug Resistance (2016), http://dx.doi.org/10.1016/j.ijpddr.2016.09.001

medications. However, the long-term use of ML has led to the development of drug resistance in animal parasites (Jabbar et al., 2006; Kaplan, 2004), and this phenomenon has also emerged in human parasites (Osei-Atweneboana et al., 2007, 2011). Because there are few therapeutic alternatives and new drug discovery is a very long and expensive process, a challenge for sustainable control of nematode parasites is to retain the efficiency of the existing AH drugs, including ML, by overcoming the process of AH resistance. This indicates the urgency to decipher the molecular mechanisms of anthelmintic resistance in nematodes.

Some active membrane transporters from the ATP-binding cassette (ABC) family are associated with ML resistance, and possibly resistance against other AH drugs, in parasite nematodes (Lespine et al., 2012). Many ABC transporters are present in various kingdoms, including different infectious agents (Koenderink et al., 2010; Lage, 2003), where they lead to various multidrug resistance phenotypes (Jones and George, 2005). Interestingly, all these transporters have two nucleotide binding domains (NBDs), which are highly conserved sequences giving the hallmark for belonging to the ABC family (Higgins, 1992). These NBDs bind ATP for hydrolysis, providing energy for substrate translocation during the transport cycle. In addition, ABC efflux transporters are all similarly organized in two transmembrane domains (TMDs) of 6 α-helices each. These TMDs delineate an "inner chamber" within the membrane where substrates can bind. Although TMDs harbor less conserved sequences when compared with NBDs, they generally present a fairly conserved conformational folding between subfamilies and species. These typical structural characteristics support the notion that these proteins share similar transport mechanisms among species (Seeger and van Veen, 2009). In mammals, multidrug ABC transporters such as P-glycoprotein (MDR1/ABCB1/Pgp) can efflux various, structurally unrelated drugs, and Pgp is responsible for the multidrug resistance (MDR) in mammalian tumor cells (Eckford and Sharom, 2009; Leonard et al., 2003). Pgp is also physiologically expressed in the tissue barriers and is involved in intestinal absorption, cerebral distribution and biliary and urinary excretions. Thereby, it contributes in a significant manner to the pharmacokinetics of a broad range of molecules of pharmacological interest (Schinkel, 1997). Noticeably, amongst many drugs, IVM and other ML are transported by mammalian Pgp (Lespine et al., 2007; Roulet et al., 2003; Schinkel et al., 1994), which contributes to modulating their AH efficacy in the host.

The model free-living nematode, *Caenorhabditis elegans (Cel)* genome contains 14 homologs of ABCB/Pgps gene products and 10 are found in the parasitic nematode *Haemonchus contortus (Hco)* (Laing et al., 2011, 2013). There is little information on their functions, except that the loss of each of the 14 Pgps increases susceptibility of *C. elegans* to IVM to various degrees (Ardelli and Prichard, 2013; Janssen et al., 2013). Also, induction of expression of some Pgp genes after selection under IVM pressure is associated with IVM resistance in *C. elegans* and in several parasitic nematodes, which can be partly reversed by using mammalian Pgp inhibitors (James and Davey, 2009; Lespine et al., 2012; Menez et al., 2016). In addition, ML were shown to inhibit parasitic nematode Pgp-mediated drug transport in heterologous recombinant systems overexpressing Hco-Pgp-2, Hco-Pgp-9.1, Hco-Pgp-16, *Cylicocylus elongatus* (*Ceg*)-Pgp-9 or *Dirofilaria immitis* (*Dim*)-Pgp-11 (Godoy et al., 2015a, 2015b, 2016; Kaschny et al., 2015; Mani et al., 2016). Nevertheless, all these data give only indirect evidence that ML can be substrates of nematode Pgps. Considerable progress was made when Cel-Pgp-1 was purified, allowing functional characterization. Indeed, its stimulation of ATPase activity by mammalian Pgp-substrate drugs provides the most compelling evidence of a multidrug transport function for a nematode Pgp, but AH drugs have not been tested on this relevant experimental system. Therefore, direct information about the molecular relationships between AH drugs and ABC transporters in nematodes is still crucially lacking.

In this context, the determination of the crystal structure of Cel-Pgp-1 at a good atomic resolution (3.4 Å, Protein Data Bank PDB structure 4F4C) realizes a clear breakthrough as it is a unique and accurate molecular description of a nematode Pgp (Jin et al., 2012). Plus, it presents an open inward-facing conformation, defining a large cavity (the "inner chamber"), which is expected to be competent for substrate uptake as the first step of its transport enzymatic cycle. Thus, it provides a reliable structural basis to undertake *in silico* investigations of drug binding by molecular docking approaches. In this work, we studied the capacity of ML and other AHs of therapeutic interest to interact with Cel-Pgp-1 using *in silico* docking techniques. In order to validate and strengthen the computational modelling approach, we first performed docking calculations with the drugs identified as substrates of Cel-Pgp-1 on the basis of their ability to stimulate its ATPase activity, as previously demonstrated *in vitro* (Jin et al., 2012). Briefly, using AutoDock 4 scoring function, we found for all 6 ATPase activators, valinomycin, vinblastine, actinomycin D, dipyridamole, progesterone and paclitaxel, a good to very good binding energy (in the range from −7.0 to −17.0 kcal mol^{-1}) for the best docking poses, which were all located within the inner chamber of the protein (David et al., work in process). Furthermore, docking calculations for positively charged rhodamine 123, taken as a negative control since it did not stimulate Cel-Pgp-1 ATPase activity, gave two poses that were either clearly outside the inner chamber or presented a very weak binding energy (−3.8 kcal mol^{-1}). This good qualitative and quantitative agreement for the correlation between *in vitro* enzymological data and *in silico* calculations provided confidence for using our modelling strategy to investigate putative interactions with Cel-Pgp-1 of compounds belonging to several AH classes. For the first time, we determined the parameters of nematode Pgp-AH interaction, including binding energy and location of the binding sites in the protein, with a model nematode Pgp. In particular, we delineated their predictive binding sites by identifying amino acid residues that interact with different drug substituents of importance in the specific binding of each ML. We thus proposed a molecular view of the binding of several ML, which showed a unified handling by the transporter since they all share a common binding domain in the inner chamber. Using this model, we compared ML binding with the binding modes of other AH classes on Cel-Pgp-1, which all presented different binding sites within the inner chamber. Our findings thus provide a significant breakthrough in understanding how AHs bind to nematode Pgps, and provide strong evidence to indicate that ML can be transported by parasite nematode Pgp-1 homologs.

2. Computational methods

2.1. Structure of Cel-Pgp-1

The Cel-Pgp-1 X-ray structure, determined at a resolution of 3.4 Å (PDB code 4F4C) (Jin et al., 2012), was used in all docking calculations. The whole chain A was taken into account, with the exception of two detergent molecules (undecyl 4-O-alpha-D-glucopyranosyl-1-thio-beta-D-glucopyranoside, PDB entry name 0SA), bound in the inner chamber, and that have been removed for grid maps calculations. The 4F4C structure includes the full glycosylated Cel-Pgp-1 sequence (1321 amino acids), but N-terminal (M1-R3) and C-terminal (G1307-K1321) segments are missing in the structure, as well as a short segment (A52-E54) located in an extended loop of the first TMD domain, and a 49-residues segment (K666-E715) belonging to the linker region. Interestingly, an additional

Please cite this article in press as: David, M.A., et al., *In silico* analysis of the binding of anthelmintics to *Caenorhabditis elegans* P-glycoprotein 1, International Journal for Parasitology: Drugs and Drug Resistance (2016), http://dx.doi.org/10.1016/j.ijpddr.2016.09.001

helix-turn-helix motif (Q9-V32) is present in the N-terminal domain, a structural feature that has not been observed in other Pgp structures released in the PDB.

Orientations of proteins in membranes database (http://opm.phar.umich.edu/) allowed the determination of the position of the lipid bilayer relative to Cel-Pgp-1, which defined the grid box position. The atomic coordinates PDB file was then converted into a PDBQT file by AutoDock Tools 4 (Morris et al., 2009) for docking calculations. PyMOL (The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC) was used as visualization tools for various tasks (3D alignments, ligands and hotspots location, grid box positioning for AutoGrid 4) and for structure rendering in figures.

2.2. Preparation and conformational analysis of ligands

The molecular structures of ligands were extracted from Chemspider (CID), Drugbank (DB) or Merck Index (M) (Suppl. Fig. S1 and S3), depending on the availability of structures. Each molecular structure was carefully scrutinized for chirality, and sometimes corrected when inconsistencies were found in the literature. For abamectin (ABA), the dry compound is a mixture of B1a (substituent isobutyl on C25) and B1b (substituent isopropyl on C25) forms. As B1a is largely predominant in natural mixtures, the docking calculations were performed only with B1a compound. Several AHs, i.e. levamisole (LEV), closantel (CLO) and monepantel (MNP) are used *in vivo* as racemic mixtures for treatment, which led us to investigate the docking of each enantiomer.

In the semi-flexible mode, the ligand is handled as flexible around all the rotatable bonds. However, the conformational space of the ligand can be poorly explored when it contains ring structures as in avermectins, since AutoDock does not consider single bonds in non-aromatic cycles as rotatable bonds. To overcome this limitation, in order to better sample the initial conformational space accessible to the ligand, we generated for each ring-containing compound 10 different low energy conformations. For this, we used Marvin Sketch and the minimization under the MMFF94 force field provided in Marvin Suite (https://www.chemaxon.com/products/marvin/marvinsketch/). The diversity of the 10 lowest energy conformers was evaluated by their pairwise root mean square deviations (RMSD) for atomic positions after superimposition under PyMOL. For each obtained cluster of close conformers, a representative one, defined as the center of the group according to the calculated RMSDs, was selected. One to five different conformers were thus selected as starting points for further docking procedures. In all cases, the selected conformers presented rather close energies, corresponding to rapidly interconverting forms of the molecule, and the docking results were generally comparable between each conformer. Thus, the most representative conformer for docking results was chosen for data presentation.

2.3. Docking calculations

Molecular docking experiments were performed using AutoDock 4 (release 4.2.6) in the semi-flexible mode with the Cel-Pgp-1 4F4C PDB structure kept rigid, and prepared with AutoDock Tools (Morris et al., 2009). AutoDock, which is the most cited docking software (Sousa et al., 2006, 2013), has a free-energy scoring function based on AMBER force field and a large set of diverse protein-ligand complexes with known inhibition constants. Few residues of the protein could have been declared as flexible in the PDBQT file, but the program restrains the total number of torsional degrees of freedom to 32, shared between the ligand and the receptor. This is a drawback in the case of the Cel-Pgp-1 structure, since the inner chamber is large, and in the absence of consistent indications about the exact location of the binding sites of the various drugs, different cavities have to be taken into account in the calculation. Thus, we privileged an approach based on better coverage of ligand flexibility and a grid box extended to the whole membrane part of the receptor protein. Indeed, for all ligands tested, the docking box, in which grid maps were computed using program AutoGrid 4, encompassed all the TM helices and the whole internal cavity, including lateral access channels and protein surface, to allow a large sampling of potential poses. The grid built by AutoGrid included 100, 124, and 126 points in x, y, and z directions, with a grid spacing of 0.375 Å to allow a good compromise between resolution of the explored volume and the size of the binding area (box dimensions $37.5 \times 46.5 \times 47.3$ Å, centered in the inner cavity of Cel-Pgp-1, at the point $x = 22.2$ Å; $y = 77.6$ Å; $z = -1.4$ Å). For each ligand conformer, 100 independent calculations were performed using the Lamarckian genetic algorithm. All the other parameters were set at the default value.

The 100 generated poses were assigned a score calculated by AutoDock that can be considered as an estimated free energy of ligand binding (indicative of binding affinity). They were then clustered as a function of the closeness of their positions and conformations with RMSD set at 2.0 Å, and finally ranked by their binding score (for the best pose in the cluster). The results are displayed in an energy scores histogram, which reproducibility could be assessed by comparing docking calculations performed on close ligand conformers, or on a truly duplicated calculation. As a result, binding energies (positions of the best pose in each cluster of the histogram) were found to fall within a range of 0.25 kcal/mol, and the number of poses in a cluster within 10%. This gives an indication of the accuracy of the histogram parameters in our series of runs, i.e., the binding energies and the overall distribution of clusters.

2.4. Data analysis

Different parameters and observables can be used in the interpretation of docking data issued by AutoDock: binding energies (i.e. docking scores), histogram bars energy range, profile of the histogram, and location of calculated positions in the protein structure. The position of clusters in figures and tables corresponds to the binding energy value of the lowest energy pose in the cluster. The spreading of clusters in the energy scores histogram was considered as a general indication for ligand docking calculation reliability: the less scattered the energies, the higher specificity of binding can be expected. Conversely, a pseudo-Gaussian profile for a group of histogram bars may suggest a non-specific docking. However, in some cases, very negative binding energies (i.e. very good docking scores) were found in poorly populated clusters whereas the general aspect of the histogram was scattered. We considered these poses as relevant, despite the lack of sampling, in view of the gap energy with the next clusters, revealing high binding site specificity.

The question arose as to whether the lowest energy cluster or the most populated cluster (highest histogram bar) had to be considered. The AutoDock docking score is based on an empirical free-energy force field which has been parameterized using a large number of protein-inhibitor complexes for which both structure and inhibition constants were known, and thus should reliably reflect the affinity of the ligand for the receptor, and the stability of the ligand-protein complex. In contrast, the number of poses found in a cluster reflects the number of times that conformers are found in very close (within RMSD) binding sites, without any anticipation of their stability. Thus considering the highest cluster in the energy histogram may be not relevant *per se* for identifying the most

Please cite this article in press as: David, M.A., et al., *In silico* analysis of the binding of anthelmintics to *Caenorhabditis elegans* P-glycoprotein 1, International Journal for Parasitology: Drugs and Drug Resistance (2016), http://dx.doi.org/10.1016/j.ijpddr.2016.09.001

probable docking site of the considered ligand. Practically, for each ligand, when the difference in the energies between the two best-scored clusters was more than 2 kcal/mol, the lowest energy one was considered as the most representative, since the other ones correspond to ligand-protein complexes associated to negligible lifetime. Alternatively, when the energy gap was narrower, we manually analyzed all the clusters within 2 kcal/mol below the lowest one for the localization of the included poses in the protein 3D structure. Finally, one or two main clusters were selected as representative, and they corresponded most often to the first and second minimum energy clusters, except when clusters corresponding to non-relevant positions in the protein were found interleaved in the ranking. In some cases, this protocol of validation led us to consider a double binding site on the protein, leading to a stoichiometry of one or two depending on whether these two docking positions were partially overlapping or not. Such a possibility simply reflects the large size of the multispecific binding domain, which likely encompasses the whole "inner chamber".

Lastly, criteria for discriminating relevant from non-relevant docking poses had to be defined. This delineation was made necessary by the fact that we enlarged the zone of docking search to the whole membrane part of the protein. The large size of the grid box ensured an exhaustive conformational exploration not biased by preliminary beliefs, although it contributed to the scattering of the docking results. Poses that were outside of the expected ligand binding pocket were not considered for further analysis. These were essentially poses located either "outside" of the Cel-Pgp-1 structure, i.e., at the protein/lipid interface, or lowermost in the transmembrane domain, i.e., at the level of interface with the cytosolic medium ("cytosolic antechamber"), where the phospholipid polar headgroups are likely to be invaginated into the cavity between the transmembrane helices in the inward-facing Cel-Pgp-1 conformation (Haubertin et al., 2006). These poses can have a functional relevance, for example as allosteric modulator sites, or not, but likely not as binding sites for transport. Finally, a search of the possible access channels allowing ligands to reach the inner pocket was performed using Mole 2.0. (http://mole.upol.cz/). In addition to the wide opening of the protein towards the cytosolic interface, the result showed two lateral tunnels located between TM10 and TM12, communicating with the cytosolic leaflet of the membrane (data not shown).

For each lowest energy pose of selected clusters, the number and nature of interacting residues were analyzed within the protein. Among these, particular interest was given to residues belonging to the "hotspots for drug binding", described hereafter.

2.5. Determination of the residues constituting the "hotspots for drug binding"

A number of experimental works have been conducted in order to determine the key residues responsible for multidrug recognition by mammalian Pgp (human and rodent isoforms). They initially included various directed-mutagenesis analyses that allowed identification of residues whose mutation led to alterations of the MDR profile, in contrast to numerous mutations that led to unspecific global decrease of Pgp function. The significance of these mutations was shown by cytotoxicity assays, which were a good indication of their involvement in the multi-specific drug recognition and binding. Furthermore, some additional data were collected with a chemical cross-linking approach, using a few drug derivatives bearing a moiety that could be activated, and ensuring specificity by testing protection by the native drug. All these data have been compiled in a review by Shilling et al. (2006), and were completed by including some references by Loo & Clarke (Bessadok et al., 2011; Loo et al., 2006a, b; Loo and Clarke, 2001, 2002). Finally, the only crystal structures released so far for mammalian Pgps are of murine Pgp. They were co-crystallized with hydrophobic cyclic peptide inhibitors than can be considered as functionally relevant ligands, and which pointed to a set of contact residues. However, two versions of the interpretation of the experimental X-ray diffraction data have been published (Aller et al., 2009; Li et al., 2014) that differed in the orientation of some transmembrane helices and side chains. Therefore, we consider the lists of contact residues from the two versions equally. All these identified residues form a collection of 62 residues listed in Suppl. Table S1 with their corresponding numbering in the 4 mammalian proteins sequences (Hsa-ABCB1, Mmu-ABCB1a and B1b, Cgr-ABCB1). Moreover, 14 of these 62 residues are common between at least two different techniques. They are all situated in the transmembrane part of the protein. Altogether, they provide a frame in the inner chamber that offers a set of anchoring points for multi-specific recognition and binding, and eventual translocation, of various transport ligands.

Multiple protein sequence alignments have been performed on Cel-Pgp-1, human ABCB1, murine ABCB1a and B1b, and Chinese hamster ABCB1, using Muscle software (Edgar, 2004). We checked that the transmembrane segments were satisfactorily aligned. Among the 62 human ABCB1 hotspot residues, 16 (26%) were found identical and 12 homologous with the corresponding residues in Cel-Pgp-1, representing a global conservation ratio of 45%. As a comparison, human ABCB1 and Cel-Pgp-1 display a global similarity of 63% (BLASTP positive matches), and still 58% when considering the TMDs only. This indicates that hotspots are subjected to more genetic variation than the overall sequence. More precisely, this set of residues is mainly hydrophobic, but less markedly for Cel-Pgp-1: the hydrophobic (F-Y-A-L-I-V-M)/hydrophilic (S-T-N-Q-H) ratio is 44/14 for Hsa-Pgp and 37/20 for Cel-Pgp-1, respectively.

3. Results

3.1. Binding mode of macrocyclic lactones on Cel-Pgp-1

We have first docked seven ML of pharmacological interest, whose structures are presented in Suppl. Fig. S1. A model for the binding site of ivermectin (IVM), and four other avermectins, abamectin (ABA), eprinomectin (EPR), doramectin (DOR) and selamectin (SEL), the ivermectin-aglycone derivative (IVA), and the milbemycin, moxidectin (MOX) on Cel-Pgp-1 was hence proposed.

The energy clustering histograms obtained from the docking calculations allowed us to identify two major possible positions for each ML (except one for DOR), based on the lowest (i.e. most negative) binding energies calculated (Fig. 1A and Suppl. Fig. S2) that were considered relevant for predicting ML binding site. The binding energies for all the considered clusters (13 for the 7 ML) ranged from −13.0 to −9.4 kcal/mol, which are indicative for high affinity binding to Cel-Pgp-1. All positions were found in the Cel-Pgp-1 deepest part of the inner transmembrane cavity, at the top of the inner chamber, and the identification of drug interacting residues, some forming H-bonds, allowed us to delineate a specific ML binding site, which involved several amino acids identified as hotspot residues on mammalian Pgp (detailed in Table 2 and Suppl. Table S2).

3.1.1. Docking of avermectins and the milbemycin, moxidectin

For IVM, the positions of the two lowest energy clusters on Cel-Pgp-1 were very close, with IVM1 (−12.2 kcal/mol, 20 poses) being slightly deeper in the inner chamber than IVM2 (−11.1 kcal/mol, 6 poses) (Table 1 and Fig. 1). Overall, they interacted with 19 (including 14 hotspot residues, 2 H-bonds) and 13 residues (6 hotspot residues, 3 H-bonds), respectively, and they shared 6

Please cite this article in press as: David, M.A., et al., *In silico* analysis of the binding of anthelmintics to *Caenorhabditis elegans* P-glycoprotein 1, International Journal for Parasitology: Drugs and Drug Resistance (2016), http://dx.doi.org/10.1016/j.ijpddr.2016.09.001

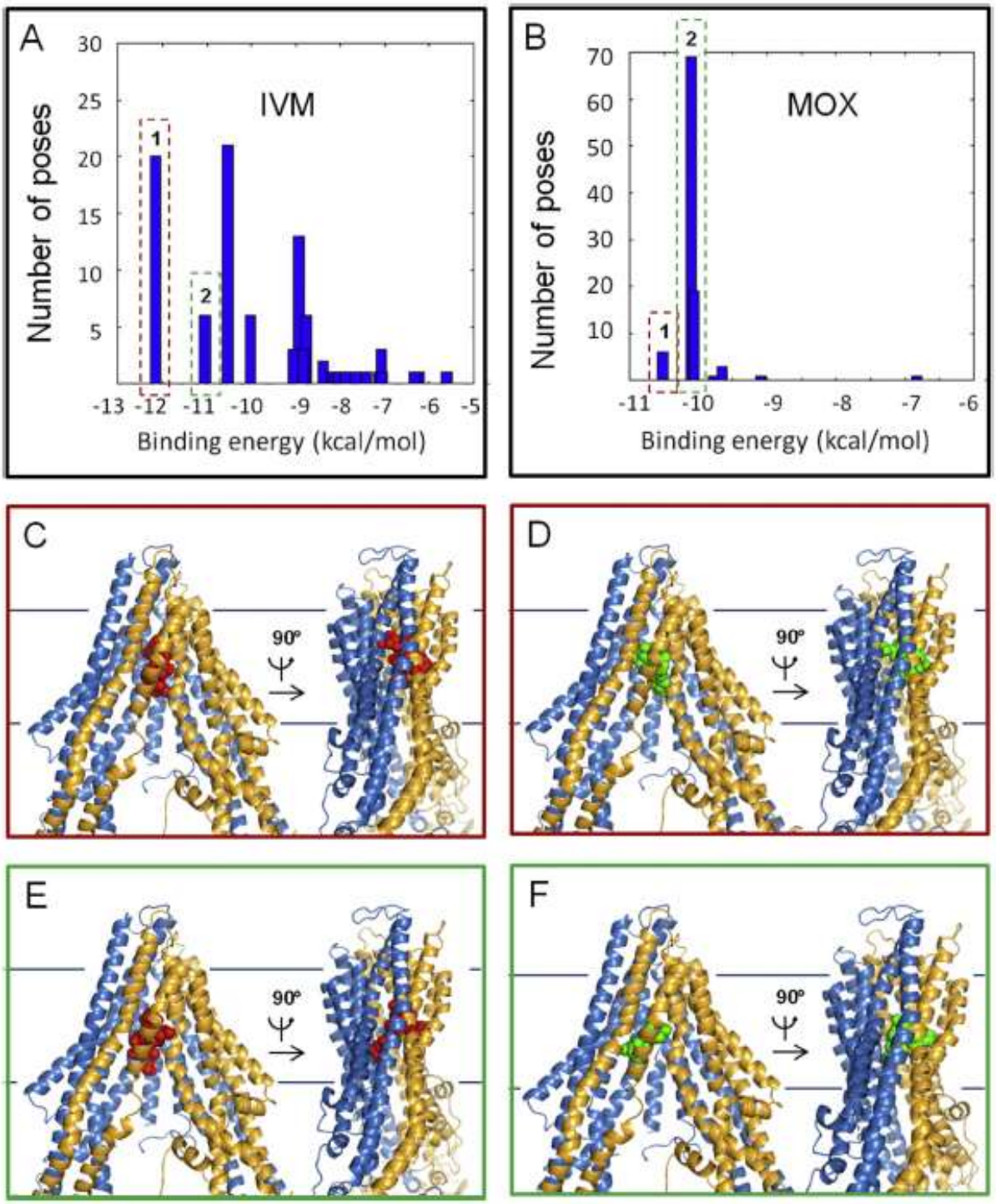


Fig. 1. Ivermectin and moxidectin binding to Cel-Pgp-1. **A** and **B**. Histograms of energy scores, clusterized at RMSD = 2 Å. The first (1) and second (2) lowest energy clusters are framed with red and green dotted lines, respectively. **C** and **D**. Binding sites of the lowest energy clusters of IVM (C) and MOX (D). **E** and **F**. Binding sites of the 2nd lowest energy clusters of IVM (E) and MOX (F). IVM is represented in red spheres, MOX is represented in green spheres, Cel-Pgp-1 in light blue (N-term) and yellow (C-term) ribbon. Images were generated using PyMol. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

common interacting residues, including 4 hotspot residues and 1 H-bond (Table 2 and Suppl. Table S2).

For ABA, the two lowest-energy positions identified (−12.3 and −11.1 kcal/mol, 8 and 37 poses, respectively) interacted with 21 and 15 residues, including 17 and 10 hotspot residues, respectively, and forming 2 H-bonds (Table 1). They shared 8 common interacting residues (including 6 hotspot residues and 0 H-bond) that closely overlapped with IVM positions (Table 2 and Suppl. Table S2).

For EPR, the clustering was more scattered than that of ABA or IVM (Suppl. Fig. S2), but nevertheless two lowest energy clusters were found at −12.5 kcal/mol (EPR1, 7 poses) and −12.0 kcal/mol (EPR2, 6 poses) (Table 1). They both bound in the deep pocket in the inner chamber, with EPR2 being deeper than EPR1. These two clusters interacted with 14 and 12 hotspot residues (6 were shared by both positions), respectively, among 20 and 17 interacting residues (8 shared), and formed 1 and 2 H-bonds, respectively (Table 2 and Suppl. Table S2).

For DOR, only one energy cluster was considered (−13.0 kcal/

Please cite this article in press as: David, M.A., et al., *In silico* analysis of the binding of anthelmintics to *Caenorhabditis elegans* P-glycoprotein 1, International Journal for Parasitology: Drugs and Drug Resistance (2016), http://dx.doi.org/10.1016/j.ijpddr.2016.09.001

Table 1
Physico-chemical properties and docking characterization of macrocyclic lactones to Cel-Pgp-1: ivermectin (IVM), abamectin (ABA), eprinomectin (EPR), doramectin (DOR), selamectin (SEL), ivermectin-aglycone (IVA) and moxidectin (MOX).

Molecule	IVM		ABA		EPR		DOR	SEL		IVA		MOX	
MW (Da)	875		873		914		899	770		587		640	
logP	4.3		4.4		4.2		4.7	5.2		3.4		5.4	
Cluster rank	1	2	1	2	1	2	1	1	2	1	2	1	2
Binding Energy (kcal/mol)	−12.2	−11.1	−12.3	−11.1	−12.5	−12.0	−13.0	−12.9	−12.4	−9.6	−9.4	−10.5	−10.1
Nb of poses	20	6	8	37	7	6	19	52	23	33	49	6	68
Nb of inter-acting residues	19	13	21	15	20	17	17	14	18	10	10	19	16
Nb of hotspot residues	14	6	17	10	14	12	14	13	12	7	8	14	8
Nb of H-bonds	2	3	2	2	1	2	1	2	2	3	2	3	2

Table 2
List of interacting residues of each transmembrane helix of Cel-Pgp-1 with the first or second lowest energy clusters of ivermectin (IVM1), abamectin (ABA1), eprinomectin (EPR1), doramectin (DOR1), selamectin (SEL2), ivermectin-aglycne (IVA2) and moxidectin (MOX2). Bold: hotspot residues. Underscored: residues establishing a H-bond. Black: residues interacting with macrocycle. Red: residues interacting with benzofurane. Blue: residues interacting with spiroketal unit. Green: residues interacting with the mono/di-saccharide moiety.

Molecule	IVM	ABA	EPR	DOR	SEL	IVA	MOX
Cluster Rank	1st	1st	1st	1st	2nd	2nd	2nd
TMa-b	E22				E22	E22	E22
					D23		
	K26				K26		
TM1				L91			L91
			M94	M94	M94		
		Q98	Q98	Q98	Q98		
TM5		F334					
TM6			L356	L356			
	F359	F359	F359	F359			
	S360	S360	S360	S360			
	M363	M363	M363				
		M364	M364	M364			
		M367			M367	M367	M367
							L371
TM7	Y771	Y771	Y771				
	F775	F775	F775	F775			
TM10	L906				L906	L906	L906
	V909	V909	V909		V909	V909	V909
	A910	A910	A910				A910
			G912				G912
	Q913	Q913	Q913	Q913	Q913		Q913
		R916	R916				R916
TM11		A986	A986	A986			A986
		S987	S987				S987
		V989	V989	V989			V989
	L990	L990	L990	L990	L990	L990	L990
					L993	L993	
					N994	N994	
					Y998		
TM12					M1021		
	Y1022	Y1022	Y1022	Y1022		Y1022	
	T1025	T1025		T1025	T1025	T1025	T1025
	I1026	I1026		I1026			
	T1028	T1028	T1028	T1028	T1028	T1028	T1028
	S1029			S1029	S1029		
					L1031		
	G1032						
	F1033						

mol, 19 poses) (Table 1), as it was significantly separated from all the other clusters (Suppl. Fig. S2). DOR1 interacted with 17 residues in Cel-Pgp-1 transmembrane inner chamber, including 14 hotspot residues, with 1 H-bond formed (Table 2 and Suppl. Table S2).

For SEL, the two lowest energy clusters were found at −12.9 kcal/mol (SEL1, 52 poses), and −12.4 kcal/mol (SEL2, 23

Please cite this article in press as: David, M.A., et al., *In silico* analysis of the binding of anthelmintics to *Caenorhabditis elegans* P-glycoprotein 1, International Journal for Parasitology: Drugs and Drug Resistance (2016), http://dx.doi.org/10.1016/j.ijpddr.2016.09.001

poses) (Table 1). These two clusters were positioned close to each other, with 14 and 18 interacting residues, 13 and 12 being hotspot residues, respectively. They each formed 2 H-bonds with different hotspot residues, and they shared 8 common interacting residues, including 7 hotspot residues (Table 2 and Suppl. Table S2).

For IVA, the two lowest energy clusters together comprising most of the poses (82%), indicating a good specificity of binding to this site, were close in energy, at −9.6 and −9.4 kcal/mol, and both bound very similarly to a unique binding site (Table 1). IVA1 and IVA2 shared 6 common interacting residues out of 10 for each, with 4 hotspot residues out of 7 and 8, respectively, and formed 3 and 2 H-bonds, respectively (Table 2 and Suppl. Table S2).

For MOX, the two lowest energy clusters were found at −10.5 kcal/mol (MOX1, 6 poses) and −10.1 kcal/mol (MOX2, 68 poses) (Table 1 and Fig. 1). They interacted with 19 and 16 residues, 8 being common to the 2 clusters, including 14 and 8 hotspot residues, respectively; 5 of which being common and forming 3 and 2 H-bonds, respectively (Table 2 and Suppl. Table S2).

In summary, the very low binding energies calculated in the docking modes (from −13.0 to −11.1 kcal/mol), strongly suggest that avermectins can bind with high affinity to Cel-Pgp-1 (Table 1). IVA and MOX positions showed significantly less negative binding energies (from −10.5 to - 9.4 kcal/mol) than the avermectins, reflecting a lower affinity, but they were still consistent with specific binding to Cel-Pgp-1. Interestingly, all ML comparably bound to a unique site located deep in the transmembrane domain of the nematode transporter, and subtle differences were observed in relation with the specific substituents of each ML (Table 2).

3.1.2. *Relation between ML structure and their binding characteristics*

We showed above that all ML bound to a similar and specific binding site (Fig. 2A) on the transmembrane domain of Cel-Pgp-1, with very negative energies reflecting high affinity binding (Table 1). In addition, ML binding involved some hotspot residues, which have been previously identified to be important for drug binding to mammalian Pgp (Table 2 and Suppl. Table S2). Nevertheless, it was possible to point out noticeable specificities between each ML, due to their different substituents that appeared important in the respective orientation of each ML in their binding site. Thus, we could propose several common features and some specific aspects for each ML in terms of binding characteristics, based on the analysis of their two lowest energy clusters.

3.1.2.1. Influence of the macrocyclic ring. The docking of each ML on Cel-Pgp-1 revealed that a total of 42 residues interacted with ML (Table 2). Among them, 14 residues, including 10 hotspot residues, were found to interact at least 5 times when considering the 10 most clearly overlapping cluster positions, which delineated a specific binding pocket, and gave a specific "fingerprint" for ML binding (Table 2 and Suppl. Table S2). Remarkably, the following ML cluster positions: IVM1, ABA1, EPR1, DOR, SEL2, IVA2 and MOX2 tightly superimposed (Fig. 2B and C), and they aligned quite perfectly at the macrocyclic ring level. Therefore, the binding characteristics of these relevant clusters were further analyzed. In particular, the residues E22, M367 and L990 were respectively interacting with the C12–13, C14–15 and C17-18 of the macrocyclic ring: interactions of E22 with IVA and MOX; M367 with ABA, IVA and MOX; and L990 with DOR, SEL and IVA (black colored in Table 2).

3.1.2.2. Influence of the benzofurane moiety. The benzofurane moiety is also a common structural feature that superimposed for all ML on their binding site in Cel-Pgp-1, although it was found in an inverted orientation for SEL and IVA when compared to the other

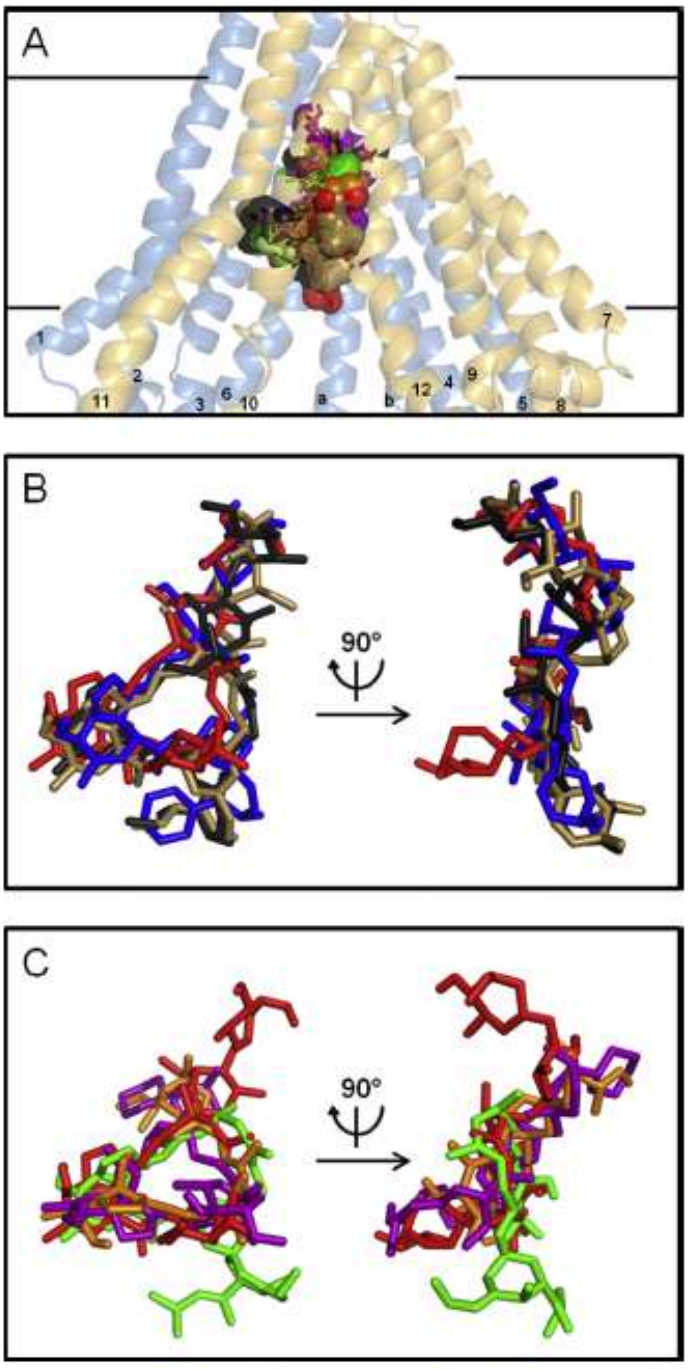


Fig. 2. Binding domain characteristics of macrocyclic lactones. A. Front view of Cel-Pgp-1 represented in transparent light blue (N-term) and yellow (C-term) ribbon, with the binding sites of the first and second lowest energy clusters of abamectin (ABA1 and ABA2, black), ivermectin (IVM1 and IVM2, red), eprinomectin (EPR1 and EPR2, sand), doramectin (DOR1, blue), selamectin (SEL1 and SEL2, purple), ivermectin-aglycone (IVA1 and IVA2, orange), and moxidectin (MOX1 and MOX2, green), all represented in sticks and transparent surfaces. TM helices numbers are indicated as their number only, on each helix. **B.** Zoom, without Cel-Pgp-1, on the front and lateral views of the overlap between the binding sites of the first lowest energy clusters of abamectin (ABA1, black), ivermectin (IVM1, red), eprinomectin (EPR1, sand), doramectin (DOR1, blue). **C.** Zoom, without Cel-Pgp-1, on the front and lateral views of the overlap between the binding sites of second lowest energy clusters of selamectin (SEL2, purple), ivermectin-aglycone (IVA2, orange), moxidectin (MOX2, green), with the lowest energy cluster of ivermectin (IVM1, red) as a marker of the binding position of other ML. All molecules are represented in sticks. Images were generated with PyMol. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

ML. Indeed, in most of the tested ML, the benzofurane moiety interacted with 6 residues, i.e. L906, V909, A910, L990, T1028 and

Please cite this article in press as: David, M.A., et al., *In silico* analysis of the binding of anthelmintics to *Caenorhabditis elegans* P-glycoprotein 1, International Journal for Parasitology: Drugs and Drug Resistance (2016), http://dx.doi.org/10.1016/j.ijpddr.2016.09.001

S1029 (red colored in Table 2). Interestingly, all ML interacted with T1028 on TM12, and in all the considered cluster positions, the C5-hydroxyl group formed an H-bond with T1028 (Table 2 and Fig. 3). T1028 has also been identified as a hotspot residue, and therefore we considered it as an important factor for the binding of all ML.

3.12.3. Influence of the spiroketal moiety. The spiroketal moiety, commencing at C17, differs between ML in various ways, and it introduces stereochemical diversity that impacts the binding to Cel-Pgp-1. ABA (an avermectin B1) is the natural compound from which most of avermectins are derived, and it has a C22–C23 double bond instead a saturated bond for IVM (C22–C23 dihydro-avermectin B1). Consequently, the spiroketal moiety of ABA has higher structural rigidity, and it bound to Cel-Pgp-1 in a slightly different orientation than that of IVM (Fig. 2B); both sharing only one common interacting residue (blue colored in Table 2). This reveals that the saturation of the C22-23 bond played a role in the conformation of the ML, and impacted the binding characteristics of the IVM spiroketal moiety. EPR (4′-epiacetylamino-4′-deoxy-avermectin B1) differs from ABA at the spiroketal moiety level by having at C25 a terminal oleandrose group. Nevertheless, the spiroketal moieties of EPR and ABA matched strictly on their binding site (Fig. 2B), interacting with residues mainly present on TM10 (Q913, R916) and TM11 (A986, S987, V989, L990) (Table 2). MOX has a substituted olefinic side chain at the 25-position and a methoxime group at the 23-position, which are two characteristics specific to this drug, not found in other commercial milbemycins or avermectins. Remarkably, although the spiroketal moieties of MOX, DOR, ABA and EPR differ considerably (Suppl. Fig. 1), the substituents bound to the same binding site with similar orientation for the four drugs (Fig. 2B and C), interacting with TM10 and TM11, and sharing Q913, A986 and V989 as common interacting residues (Table 2). Conversely, although DOR and SEL have similar spiroketal moieties characterized by a C25-cyclohexyl substituent, they did not overlap on the Cel-Pgp-1 binding site (Fig. 2 B and C). SEL and IVA have also similar spiroketal moieties that are both partly aligned with the disaccharide group of the four di-glycosylated avermectins (Fig. 2 B and C), interacting partly with TM11 and partly with TM12 (Table 2). This clearly revealed that common structural features, other than the spiroketal moiety, determine the orientation of ML on the binding site. This could be the presence of one or two sugars, or the saturated bond on C22-23 in SEL and IVM.

3.12.4. Influence of the disaccharide moiety. The four avermectins IVM, ABA, EPR and DOR have in common a disaccharide moiety linked to the C13 of the macrolide. This substituent mostly superimposed for these four ML (Fig. 3B), and interacted with TM1, TM6, TM7 and TM12 (green colored in Table 2). They all shared a number of interacting hotspot residues, i.e. F359, S360, F775 and Y1022, plus Q98, M363, M364, Y771, T1025 and I1026 when considering

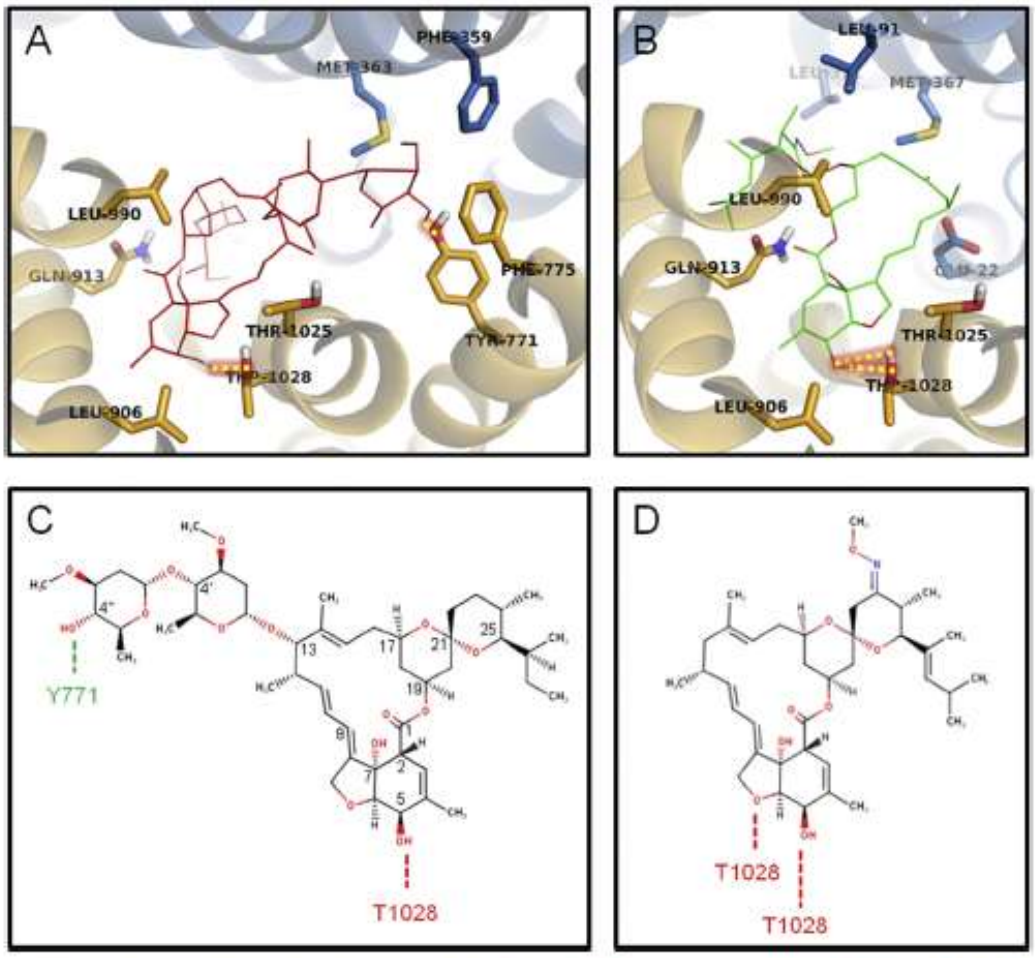


Fig. 3. Key residues of Cel-Pgp-1 for interaction with macrocyclic lactones. A and **B**. Cel-Pgp-1 key residues surrounding the binding site or forming H-bonds with ivermectin (IVM1, red lines) (A) and moxidectin (MOX2, green sticks) (B). Cel-Pgp-1 is represented in light blue (N-term) and yellow (C-term) transparent ribbons for TM helices and sticks for specifically interacting residues. Atoms of Cel-Pgp-1 residues and substrates are colored in blue for N, red for O and grey for H. H-bonds are represented in yellow dotted lines with red shadow. **C** and **D**. Ivermectin (C) and moxidectin (D) plane structures with H-bonds to residues of Cel-Pgp-1 represented as dotted lines. Images were generated with PyMol. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Please cite this article in press as: David, M.A., et al., *In silico* analysis of the binding of anthelmintics to *Caenorhabditis elegans* P-glycoprotein 1, International Journal for Parasitology: Drugs and Drug Resistance (2016), http://dx.doi.org/10.1016/j.ijpddr.2016.09.001

3 ML out of the 4, that clearly characterized the binding fingerprint specific of di-glycosylated avermectins. Interestingly, ABA and IVM disaccharide moieties both formed an H-bond with the hotspot residue Y771, showing the importance of this residue in the orientation of ML that contain a disaccharide (Table 2 and Fig. 3A and C). Apart from Y1022 and T1025, which interacted with spiroketal or benzofurane, none of the residues listed above interacted with IVA and MOX. Given the absence of glucide substituent on these two ML, and the monosaccharide of SEL, these residues are interacting specifically with the distal sugar unit of the disaccharidic avermectins. Indeed, the orientation of the monosaccharide of SEL and the hydroxyl on the C13 of IVA superimposed in a different orientation than that of the corresponding structure of other ML, which overlapped with the spiroketal moiety of IVM, and shared with it the residues E22 and K26 in TMa-b, and Q913 in TM10 as common interacting residues (Fig. 2 and Table 2). Because their only structural difference relies on the presence or absence of disaccharide, the comparison of IVM and IVA binding provided interesting information. Remarkably, 7 of the 10 residues interacting with IVA1 were all found to interact with IVM1, with a common H-bond on T1028, showing that the binding site of these two closely related molecules are very similar (Suppl. Table S2). It should be mentioned that the three ML lacking the disaccharide (SEL, IVA and MOX) exhibited an energetic clustering clearly less scattered than the four di-glycosylated avermectins (see Fig. 1A and Suppl. Fig. S2), suggesting that the absence of the two sugar units significantly increased the selectivity of the calculated interactions with the recognition domain. In addition, the two unglycosylated ML (IVA and MOX) have significant higher (i.e. less negative) binding energies when compared with other ML, which supports the important role of sugar in anchoring IVM onto the binding site.

Globally, these data reveal that all ML bound with good energy in the same pocket buried within the Cel-Pgp-1 inner chamber, and were clearly superimposed. Consideration of the residues involved in the binding of each chemical moiety of the ML, presented in different colors in Table 2, clearly supports a typical and unique binding pocket for all ML (Figs. 2 and 3). Furthermore, as the four ML that contain a disaccharide moiety were aligned for their lowest energy clusters, while this happens only for the second clusters of SEL, IVA and MOX, we propose that the main determinant of the orientation of ML, in their binding pocket, is the disaccharide moiety.

3.2. Binding mode of other anthelmintic drugs with Cel-Pgp-1

3.2.1. Docking results

The structures of all AHs tested are shown in Suppl. Fig. S3. For predicting AH binding site, a similar approach was used to choose the optimal cluster energy from the histogram obtained from the docking calculations (Suppl. Fig. S4).

For triclabendazole (TCZ), the lowest energy cluster (TCZ1, 97 poses) appeared at −8.0 kcal/mol, whereas the only other cluster (TCZ2, 3 poses) was found at −7.4 kcal/mol (Table 3). On the one hand, TCZ1 was found deep in the inner chamber of Cel-Pgp-1 (Fig. 4), interacting with 8 residues, (all hotspot residues, 1 H-bond) (Table 4). On the other hand, the binding site of the second cluster was in the middle of the inner chamber, at the level of the middle of the transmembrane bilayer (Fig. 4). For this position, 9 interacting residues were found (6 hotspot residues, 1 H-bond) (Table 4). These two positions were totally separated (Table 4 and Fig. 4), hence consistent with a binding stoichiometry of 2.

For thiabendazole (TBZ), the lowest energy cluster (TBZ1, 94 poses) was at −7.2 kcal/mol (Table 3), but it was found at a very lateral position in the TM helices of Cel-Pgp-1 (Fig. 4). TBZ1 interacted with 12 residues (3 hotspot residues, no H-bond) (Table 4). Alternatively, the second cluster (TBZ2, 3 poses), found at −5.7 kcal/mol, was found deep in the inner chamber, totally apart from TBZ1 (Fig. 4). It showed 7 interacting residues, (6 hotspot residues, 2 H-bonds) (Table 4) and shared 5 common residues, including 4 hotspots, with TCZ1 (Table 4).

For levamisole (LEV), the lowest energy cluster, (LEV1, 31 poses) of the two enantiomers R and S, which gave similar results for all clusters, was found at −6.8 kcal/mol (Table 3). It bound to Cel-Pgp-1 on a very lateral position in the TM helices bundle, close to TBZ1 position, and sharing 3 common interacting residues with TBZ1 (Table 4 and Fig. 4). LEV1 interacted with 12 binding residues (5 hotspot residues, 1 H-bond) (Table 4). The second lowest energy cluster, LEV2 (65 poses), formed a cluster at −6.1 kcal/mol (Table 3). It interacted deep in the inner chamber, very differently from LEV1, and overlapping with TCZ1 and TBZ2 (Fig. 4). It bound to 8 residues, (7 hotspot residues, 2 H-bonds), and showed 7 common interacting residues, including 6 hotspot residues, with TBZ2, and 5 common interacting residues with TCZ1, all hotspot residues (Table 4).

For closantel (CLO), similar lowest energy clusters CLO1 were found whatever the enantiomer considered, S or R. For the latter, it contained the maximum number of poses (80) at −10.3 kcal/mol (Table 3 and Fig. 5A). The binding site of this cluster was in the middle of the inner chamber (Fig. 5B), similar to the three other clusters found between 9.5 and 9.9 kcal/mol, and showed 10 interacting residues (6 hotspot residues, 2 H-bonds) (Table 4). It shared 3 common interacting hotspot residues with TCZ1, 2 with TBZ2 and LEV2, as well as 2 common interacting residues with TCZ2 (Table 4 and Fig. 4).

For monepantel (MNP), two main docking positions were found for both enantiomers S and R. For MNP-S, on one hand, the lowest energy cluster (MNP1, 3 poses) was found at −8.1 kcal/mol (Table 3), and bound "horizontally" at the level of the middle of the bilayer (Fig. 4). The 2 following clusters docked at very similar positions. On the other hand, the 4th cluster (MNP4, 1 pose) showed an energy of −7.8 kcal/mol, and was very close to most of the clusters LEV2, TBZ2 and TCZ2, at the deepest part of the inner chamber (Fig. 4). They both interacted with 14 residues, including 6 hotspot residues and 1H-bond for MNP1, and 12 hotspot residues without forming any H-bond for MNP4 (Table 3). These two binding

Table 3
Physico-chemical properties and docking characterization of anthelmintics to Cel-Pgp-1: triclabendazole (TCZ), thiabendazole (TBZ), levamisole-R (LEV-R), closantel-R (CLO-R), monepantel (MNP-S) and emodepside (EMD). IVM data from Table 1 are given in italic for the sake of comparison. (*) cluster positioned outside the inner chamber.

Molecule	TCZ		TBZ		LEV-R		CLO-R	MNP-S		EMD	*IVM*	
MW (Da)	360		201		204		663	473		1119	*875*	
logP	5.8		2.2		2.9		7.6	5.6		7.3	*4.3*	
Cluster rank	1	2	1*	2	1*	2	1	1	4	1	*1*	2
Binding Energy (kcal/mol)	−8.0	−7.4	−7.2	−5.7	−6.8	−6.1	−10.3	−8.1	−7.8	−9.5	*−12.2*	*−11.1*
Nb of poses	97	3	94	3	31	65	80	3	1	5	*20*	6
Nb of inter-acting residues	8	9	12	7	12	8	10	14	14	17	*19*	*13*
Nb of hotspot residues	8	6	3	6	5	7	6	6	12	6	*14*	6
Nb of H-bonds	1	1	0	2	1	2	2	1	0	3	*2*	*3*

Please cite this article in press as: David, M.A., et al., *In silico* analysis of the binding of anthelmintics to *Caenorhabditis elegans* P-glycoprotein 1, International Journal for Parasitology: Drugs and Drug Resistance (2016), http://dx.doi.org/10.1016/j.ijpddr.2016.09.001

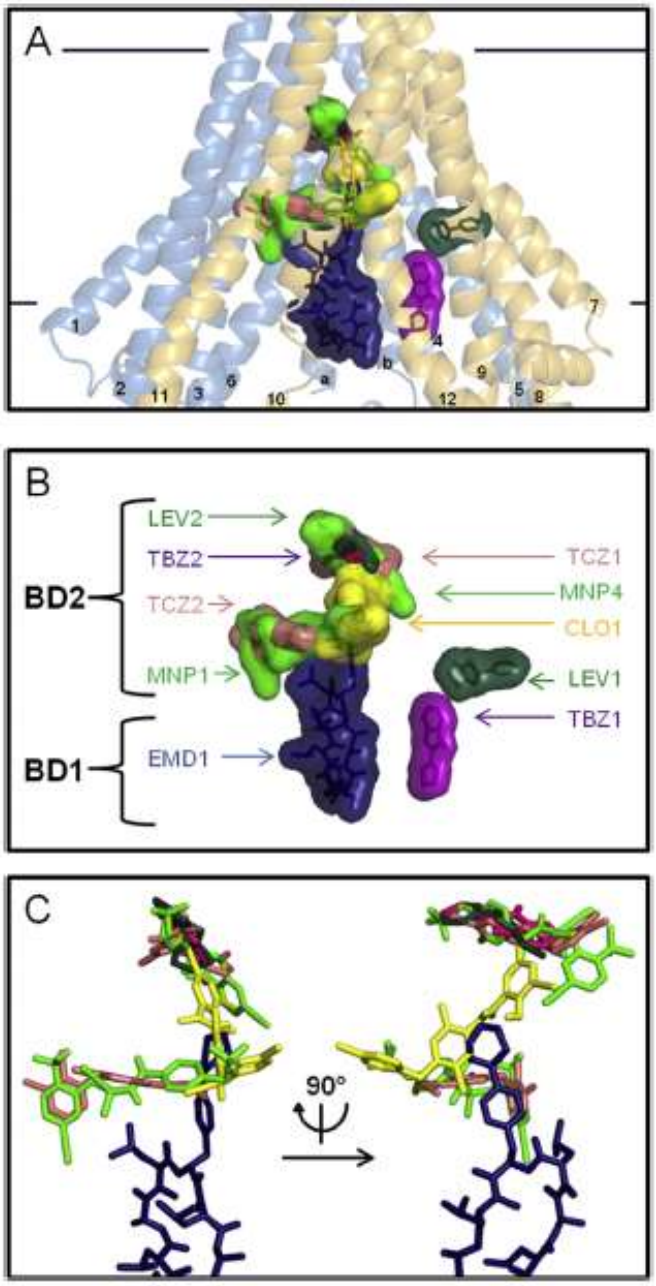


Fig. 4. Binding domains of anthelmintic drugs. A. Front view of Cel-Pgp-1 represented in transparent light blue (N-term) and yellow (C-term) ribbon, with the binding sites of the first and second lowest energy clusters of triclabendazole (TCZ1 and TCZ2, light pink), thiabendazole (TBZ1 and TBZ2, purple), levamisole-R (LEV1 and LEV2, dark green), closantel-R (CLO1, yellow), monepantel-S (MNP1 and MNP4, light green), emodepside (EMD1, dark blue), all represented in sticks and transparent surfaces. TM helices numbers are indicated as their number only, on each helix. **B.** Zoom, without Cel-Pgp-1, on the front view of all anthelmintics docking positions in the binding sub-domains, represented in transparent surfaces. **C.** Zoom, without Cel-Pgp-1, on the front and lateral views of the overlaps between the binding sites of triclabendazole (TCZ1 and TCZ2, light pink), thiabendazole (TBZ2, purple), levamisole (LEV2, dark green), closantel (CLO1, yellow), monepantel (MNP1 and MNP4, light green) and emodepside (EMD1, dark blue). All molecules are represented in sticks. Images were generated with PyMol. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

locations were totally separated (Table 4), consistent with a stoichiometry of 2. In addition, MNP1 shared 9 common interacting residues with TCZ2 and 2 with CLO1. MNP4 shared 7 common interacting residues with TCZ1 and TBZ2, 8 with LEV2, and 4 with CLO1 (Table 4 and Fig. 4).

For emodepside (EMD), the lowest energy cluster (EMD1, 5 poses) was found at −9.5 kcal/mol (Table 3). Its docking position was found near the cytosolic aperture of the inner chamber (Fig. 4). EMD1 bound to 17 residues (6 hotspot residues, 3 H-bonds) (Table 4), and shared 4 interacting residues with CLO1, 1 with TCZ2, and 2 with MNP1. All the other clusters bound very closely to the same site.

3.2.2. Analysis of the binding sites of AHs

The binding domain in Cel-Pgp-1 for AHs belonging to different classes could be analyzed by comparing their respective interacting sites and residues (Table 4 and Fig. 4). First, the two benzimidazoles, TCZ and TBZ, as well as the imidazothiazole LEV, displayed docking positions (TCZ1, TBZ2 and LEV2) that were remarkably highly overlapping (Table 4 and Fig. 4): they shared 5 common interacting residues (Q98, N994, M1021, Y1022, T1025), all hotspot residues. Their related chemical structures and small sizes made possible their interaction with the deepest part of the inner chamber of Cel-Pgp-1.

Second, this inner sub-domain is not specific to the structurally related AHs, TCZ, TBZ and LEV, since CLO1 and MNP also bound on it. Nevertheless, with its two binding positions, MNP, along with CLO1 and TCZ2, can describe a bigger surface of the inner chamber by filling various connected sub-sites (Fig. 4). Therefore, in the core of the chamber, the three AHs, TCZ, CLO and MNP share a high number of common interacting residues with IVM (taken as the representative molecule for all ML, gathering IVM1 and IVM2): 8 (including 7 hotspot residues) for TCZ1 & TCZ2, 8 (including 5 hotspot residues) for CLO1, and 9 (including 6 hotspot residues) for MNP1 and MNP4 (Table 4).

Finally, the large molecule EMD was well accommodated by the binding sub-domain close to the cytoplasmic entrance of the inner chamber of Cel-Pgp-1. Indeed, it also shared some common interacting residues with MNP1, TCZ2 and CLO1 (Fig. 4). Moreover, among its 17 interacting residues found in the docking, 9 (including 5 hotspot residues) were common to IVM1 and IVM2 (Table 4).

4. Discussion

Resistance against anthelmintic drugs in parasite nematodes, including against ML, compromises the therapeutic benefit of AH-based treatments (Lespine et al., 2012). Taking account of the serious concern for animal and human health caused by nematode infections, it is urgent to gain an understanding of the molecular factors involved in limiting drug efficacy in nematodes. It is now well established that some ABC transporters are central in limiting ML efficacy. Mammalian Pgp transports ML out of the host organism thereby limiting the active concentration reaching the nematode target (Kiki-Mvouaka et al., 2010). In addition, some nematode Pgps also interact with ML, and they are involved in the process of development of AH resistance (Lespine et al., 2012). Nevertheless, tools are needed to understand in detail the role of these proteins in ML transport, and to consider them as relevant targets for reversion of AH resistance. The release of the 3D crystal structure of Cel-Pgp-1, with good resolution (Jin et al., 2012) and QMEAN Z-score (Domicevica and Biggin, 2015) provides an opportunity to launch a sound *in silico* structural study of drug binding to this protein. This is particularly relevant in the view of multispecific recognition, and hence expected transport capacity, of this model protein. Indeed, the open inward-facing conformational state, in which Cel-Pgp-1 has been crystalized as revealed by the study of Jin et al. (2012), is by far the most frequent found conformation among the ABC proteins structurally determined today (Ward et al., 2007), and particularly the unique one found for mammalian Pgps (Aller et al., 2009; Li et al., 2014; Szewczyk et al., 2015). This conformation

Please cite this article in press as: David, M.A., et al., *In silico* analysis of the binding of anthelmintics to *Caenorhabditis elegans* P-glycoprotein 1, International Journal for Parasitology: Drugs and Drug Resistance (2016), http://dx.doi.org/10.1016/j.ijpddr.2016.09.001

Table 4
List of interacting residues of each transmembrane helix of Cel-Pgp-1 with the selected lowest energy clusters of triclabendazole (TCZ), thiabendazole (TBZ), levamisole-R (LEV-R), closantel-R (CLO-R), monepantel-S (MNP-S) and emodepside (EMD). The interacting residues with the reference ML, ivermectin (Table 1), are also presented for comparison. Bold: hotspot residues. Underscored: residues establishing a H-bond. Green: common interacting residues to TCZ1. Blue: common interacting residues to TCZ2. Orange: common interacting residues to TBZ1. Dark green: common interacting residues to TBZ2. Purple: common interacting residues to LEV2. Brown: common interacting residues to CLO1. Pink: common interacting residues to MNP1. Black: interacting residues never shared with any other AH cluster.

Molecule	TCZ		TBZ		LEV-R		CLO-R	MNP-S		EMD	IVM	
Cluster Rank	1st	2nd	1st	2nd	1st	2nd	1st	1st	4th	1st	1st	2nd
TMa-b			L25 I29 E33		L25				P21	L11 R12 K26	E22 K26	D23 K26 K30
TM1	Q98	L91		Q98		Q98		L91	Q98			
TM3		Q219						Q219 Y220		Y220		Y220
TM5					Q327							
TM6									F359 M363	L371 P374	F359 S360 M363	M367 G370 L371
TM7	Y771 F775				G764 Y768				Y771		Y771 F775	
TM8					Q807 C810 S811 M814							
TM9			R873 T876 V877 T880		T876 T879 T880							
TM10		V909 Q913					L906 V909 A910 Q913	L906 V909 A910 Q913 R916		Q913 R916 G917	L906 V909 A910 Q913	V909 Q913
TM11	N994	A986 S987 V989 L990 Y991		N994 A997 Y998		L993 N994 A997 Y998		Y983 A986 S987 V989 L990 Y991	L993 N994 A997 Y998		L990	L990 L993

seems to be stable and ubiquitous, and is now commonly admitted to be the competent one for substrate recognition and uptake. This is corroborated by the fact that the only Pgp structures co-crystalized with a drug involve Mmu-Pgp with QZ59 derivatives, all bound in the same location within the inner chamber. From all the structures today available, either from bacterial or eukaryotic organisms, the general scheme admitted for drug translocation by multidrug transporters implies that the initial drug chelation step, obviously necessary for drug transport, occurs before and independently from the following trans-conformational steps constituting the catalytic cycle of the active ATP-dependent transmembrane drug translocation.

However, a critical point is that our *in silico* approach needs to be validated by an *in silico*/*in vitro* correlation. Such a validating correlation could actually be achieved by considering the enzymatic data reported in the same publication (Jin et al., 2012) and a series

Please cite this article in press as: David, M.A., et al., *In silico* analysis of the binding of anthelmintics to *Caenorhabditis elegans* P-glycoprotein 1, International Journal for Parasitology: Drugs and Drug Resistance (2016), http://dx.doi.org/10.1016/j.ijpddr.2016.09.001

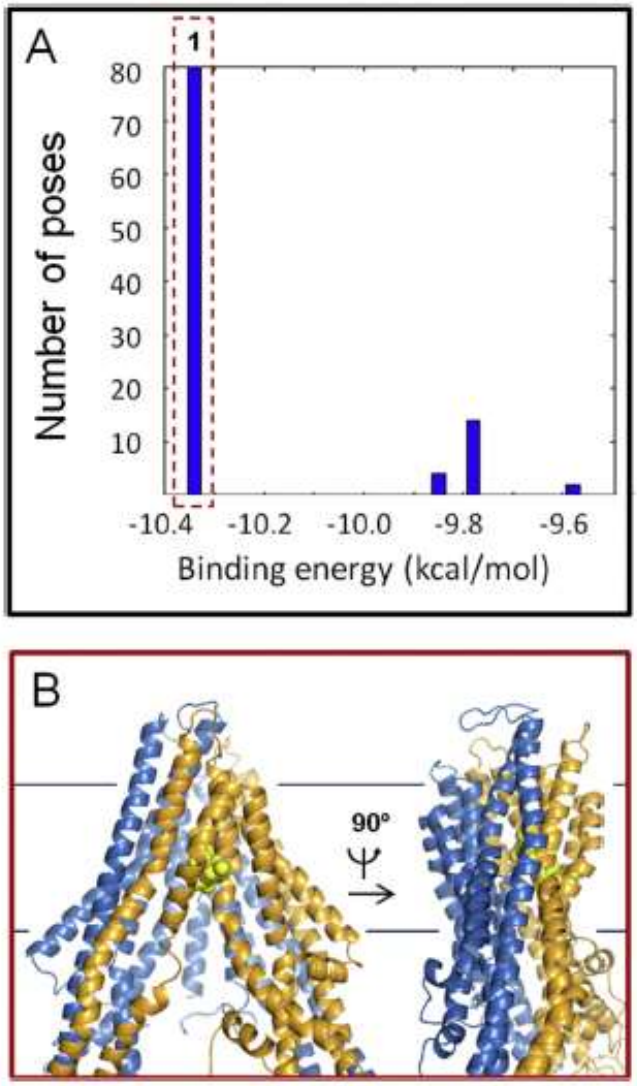


Fig. 5. Closantel-R binding to Cel-Pgp-1. A. Histogram of energy scores, clusterized at RMSD = 2 Å. The lowest (1) energy cluster is framed with red dotted lines. **B**. Binding site of the lowest energy cluster, generated using PyMol. Closantel-R is represented in light yellow spheres, Cel-Pgp-1 in light blue (N-term) and yellow (C-term) ribbon. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

of *in silico* docking calculations performed for a set of 6 molecules that have been shown to stimulate Cel-Pgp-1 ATPase activity, compared to a negative control without effect on this activity. This previous analysis allowed us to draw the conclusion that the inner chamber of Cel-Pgp-1 should be considered as the binding domain competent for recognizing transport substrates, thereby initiating the active transport cycle leading to their transmembrane translocation. We have thus defined a threshold value for binding energies between "ligands" and "weak- or non-ligands", at about −4 to −5 kcal/mol (David et al., work in process).

Finally, since the inner chamber is in direct connection with the lipid phase of the surrounding membrane, it can be expected to be filled with phospholipid molecules when considering the limits of the hydrophobic thickness of the protein and its positioning in the bilayer calculated by the OPM server (http://opm.phar.umich.edu). In a first approximation, the lipid phase can be considered as fluid, so that, similarly to solvent molecules, lipid molecules are not considered as ligands bound to the protein, and thus not included in the grid maps calculated by Autodock. Lipids are rather expected to influence the kinetics of interaction but not the drug binding affinity for the receptor at equilibrium. However, in a multispecific context, the specific binding of lipids cannot be excluded. If such specific lipid binding site exists, with an affinity comparable to the considered drug, possible competitive interactions can occur between drug and lipid, depending on the location of the corresponding binding sites within the multispecific inner chamber. The scope of this study is nevertheless to model and analyze possible AH drugs interactions with Cel-Pgp-1 in a pharmacological perspective.

4.1. Structural characterization of the ML binding site in Cel-Pgp-1

For the first time, our docking data provides strong evidence that Cel-Pgp-1 is involved in ML high affinity recognition and eventual active transport. All the tested ML displayed very low binding energies (between −13.0 and −9.4 kcal/mol), in the range of binding energies found for the set of validation ligands, as discussed elsewhere (David et al., work in process). The whole docking data collected on seven ML revealed a striking homogeneity in the predominant bound conformations found, in favor of a very similar binding mode for all of them within the inner chamber of Cel-Pgp-1. This convergence is indicative of the robustness of the docking results and contributes to give them high confidence.

Remarkably, all ML were bound in the most buried part of the transmembrane inner chamber. More specifically, the macrocycle and the benzofurane moieties, which together constitute the common core of the ML chemical structures, always clearly superimposed in the same sub-sites. In particular, the hotspot residue T1028 was found to establish various H-bonds with benzofurane groups of all tested ML. At variance, the spiroketal moiety, also present in all ML, proved to be able to accommodate two distinct sub-sites in the same area, suggesting possible alternate binding modes for this substituent. More noticeably, all the conformations found for the 7 tested ML involved a high ratio of hotspot residues: indeed, over the 13 considered docking positions, a total of 209 interacting residues were detected by AutoDock, of which 149 (about 70%) were hotspot residues. This emphasizes a possible underlying structural homology with mammalian Pgp for binding large amphiphilic/hydrophobic ligands.

Regarding the disaccharide moiety harbored by four of the tested avermectins (ABA, IVM, EPR, DOR), the comparison with the data obtained with SEL (bearing one monosaccharide), IVA and MOX (no sugar unit), highlighted the effects of a sugar moiety on binding affinity for Cel-Pgp-1. First, SEL displayed binding energies (−12.9/-12.4 kcal/mol) typically within the range of values calculated for the other tested avermectins (−13.0 to −11.1 kcal/mol), whereas IVA and MOX clearly displayed less favorable binding energies (−10.5 to −9.4 kcal/mol). This indicates that the proximal sugar unit plays a major role in establishing energetic interactions with the receptor binding domain, while the distal sugar is less influential in stabilizing the interaction. This is reinforced by the significant number of hotspot residues (13) on Cel-Pgp-1 interacting with the sugar moiety. Second, SEL, IVA and MOX were all found in two overlapping but distinct docking positions, both relevant since they correspond to their two lowest energy clusters. One position overlaid well with that of the four disaccharide-branched avermectins, whereas the other one, although still in the same sub-site, was oriented differently. Although highlighting the remarkable common binding mode of all the tested ML, these data also illustrate the role of the disaccharide moiety in anchoring ML in the summit part of the inner chamber in Cel-Pgp-1, in which it is well fitted due to favorable steric interactions.

Our data are consistent with reports showing interactions between ML and *C. elegans* Pgps, with increased Cel-Pgp-1 expression in IVM-resistant *C. elegans* strains (James and Davey, 2009; Yan

Please cite this article in press as: David, M.A., et al., *In silico* analysis of the binding of anthelmintics to *Caenorhabditis elegans* P-glycoprotein 1, International Journal for Parasitology: Drugs and Drug Resistance (2016), http://dx.doi.org/10.1016/j.ijpddr.2016.09.001

et al., 2012), and increased *C. elegans* sensitivity to IVM (Ardelli and Prichard, 2013; Janssen et al., 2013) and MOX (Bygarski et al., 2014) after disruption of the Cel-Pgp-1 gene. In addition, major differences have been reported on interaction of IVM and MOX with mammalian Pgp. Indeed, the apparent affinity of IVM for Cgr-Pgp calculated *in vitro* was higher than that of MOX (Lespine et al., 2007; Prichard et al., 2012), and disaccharide moities were then suggested to modify the binding property of ML. Similarly, both drugs interact differently with parasitic nematode Pgps (Godoy et al., 2015a, 2015b, 2016; Kaschny et al., 2015; Kerboeuf and Guegnard, 2011; Mani et al., 2016). Thus, our data strengthen the hypothesis that the presence of sugar substituents could partly determine the affinity binding of ML to various mammalian and nematode Pgps.

4.2. Binding of other anthelmintics on Cel-Pgp-1

In silico docking calculations for other AH drugs belonging to various chemical classes brought additional information about the molecular mechanisms of multispecific recognition properties exhibited by Cel-Pgp-1. The chemical diversity of the tested AHs revealed the wide range of ligand-protein interactions accessible in this putative multidrug transporter. In particular, a large majority of docking positions were found within the inner chamber, indicating its pivotal role in the initial step of recognition and binding of substrate drugs. The very negative binding energies of the four hydrophobic AHs, TCZ, CLO, MNP and EMD, indicate that Cel-Pgp-1 might efficiently transport them, favoring the possible contribution of this protein to AH drug resistance. However, two noticeable exceptions were observed for TBZ and LEV, whose best binding energy clusters included poses located at the membrane limit of the transmembrane helix bundle. Such a binding position makes the translocation of these two substrates unlikely. The hypothesis that these drug binding sites are not relevant for active transmembrane drug translocation is thus favored. We cannot, however, exclude that these binding sites can play some functional, allosteric regulatory role, which would be desirable to test *in vitro*. In contrast, the conformations of the second lowest energy clusters for TBZ and LEV appeared to be localized within the inner chamber, but with less favorable binding energy (−5.7 and −6.1 kcal/mol, respectively). It remains to be evaluated, *in vitro*, whether these "weak ligands" actually interact with Cel-Pgp-1 at biochemically/pharmacologically relevant concentrations.

Remarkably, among the total 87 interacting residues encompassed by the docking positions of the 6 tested AHs found in the inner chamber, 57 (i.e. 66%) are hotspot residues. Once again, this high ratio provides an indication of a common structural template for multispecific recognition shared by Cel-Pgp-1 and mammalian Pgp, although their respective drug handling profiles do not fully overlap, possibly due to differences in the balance of hydrophobic/polar residues lining the inner chamber. Interestingly, TCZ, CLO and EMD have been reported to interact with mammalian Pgp (Dupuy et al., 2010; Elmshauser et al., 2014), whereas TBZ and LEV were shown not to interact with mammalian Pgp (Dupuy et al., 2010; Efferth and Volm, 1993; Hayeshi et al., 2006).

4.3. Usefulness and applicability of this structural model of multispecific drug binding

Globally, this *in silico* analysis of ML binding on Cel-Pgp-1 provides a new, detailed vision of the fine molecular mechanisms of multispecific recognition capacities of a putative multidrug transporter. Altogether, our data based on a large selection of potential ligands are convergent to reveal that the inner chamber forms a large multi-specific binding domain composed of some continuously connected sub-sites. This representation can help in understanding or predicting mutual relationships of recognized ligands, depending on the sub-site(s) they potentially bind and the overlaps between sub-sites. In particular, the smallest size ligands, such as TCZ and MNP, can be predicted to bind with a stoichiometry of 2, while some others, such as MOX, display "dual" binding site with two different, partially overlapping positions. These different configurations illustrate the palette of possible interactions offered by all the contact residues lining this chamber, and the large size of the multidrug binding domain. This is clearly at variance with the standard depiction of ligand-receptor interaction typically given by the "lock-and-key" mechanism. Overall, Cel-Pgp-1 and mammalian Pgp appear to share common structural features in their respective inner chambers based on the high proportion of hotspots residues in Cel-Pgp-1, aligned (but often not identical) with residues involved in multidrug recognition in mammalian Pgp.

More relevant pharmacologically is the observation of overlapping docking positions for ML and the AH drugs, TCZ, CLO, EMD and MNP, which permits the prediction that simultaneous presence of a ML and one of these compounds would lead to competitive binding on, and hence transport by, Cel-Pgp-1. This needs of course to be confirmed experimentally, but if so, it could be inferred as advantageous to administer a Cel-Pgp-1-mediated resistant worm a combination of a ML with one of these compounds.

This model can be the basis for establishing similar drug binding models on ABC transporters of parasitic nematodes showing sufficient sequence similarity with Cel-Pgp-1, in order to allow reliable structure rebuilding by homology modelling. This strategy could be powerful to test *in silico* the transporters capacities to exhibit either selective or multiple AH drug binding. In the *C. elegans* genome, 14 Pgp-related ABC genes show differential expression levels during the worm life cycle and various tissue localizations (Zhao et al., 2004) (Wormbase: http://www.wormbase.org/). The closest homologs of Cel-Pgp-1, expressed in the intestine, are the ABC full transporters of the sub-family B and its orthologs in other nematodes. For example, in *H. contortus*, if we restrict the computation of sequence identity to TMD1-TMD2 regions, we observe that Hco-Pgp-1 displays 63% sequence identity with Cel-Pgp-1, while the latter shares only between 22% (Cel-Pgp-11) and 48% (Cel-Pgp-9) of sequence identity with all other Cel-Pgp sequences of the same genome. Thus, Hco-Pgp-1 would be the most direct candidate in this parasite to benefit from a similar *in silico* approach for AH drugs docking on its structural model.

Therefore, based on the Cel-Pgp-1 structural model of multidrug binding, describing the set of interacting residues lining the inner chamber where ML bind specifically on a given nematode Pgp, will be pivotal for the screening or rational design of optimized ligands with high affinity and selectivity. Such molecules would competitively inhibit the transport of AH drugs in nematodes and could be used to increase AH efficacy. For this, it will be important to check for the specificity of the designed inhibitor for nematode Pgps, to avoid toxic effects in hosts. This is especially important in some mammals, such as some collie dogs with defective Pgp transporters in the blood-brain barrier which are highly sensitive to ivermectin, due to excess of drug accumulation into the brain and subsequent neurotoxicity (Roulet et al., 2003; Schinkel et al., 1994). Such data would point to *in vitro* and *in vivo* experiments relevant for investigating the involvement of the ABC transporters in nematode drug transport and resistance, and they could provide a clue for finding specific efficient AH resistance reversing agents.

In conclusion, in this *in silico* work based on a recently released crystal structure of the nematode *C. elegans* ABC transporter Pgp-1, we showed for the first time that most of the AH drugs currently used in veterinary and human medicine to treat helminth infections can be substrates of Cel-Pgp-1. We clearly identified

Please cite this article in press as: David, M.A., et al., *In silico* analysis of the binding of anthelmintics to *Caenorhabditis elegans* P-glycoprotein 1, International Journal for Parasitology: Drugs and Drug Resistance (2016), http://dx.doi.org/10.1016/j.ijpddr.2016.09.001

several hotspots for drug binding, as well as common motifs of different ML interacting with specific sub-sites of the binding pocket of Cel-Pgp-1, which can be targeted to design molecules that would inhibit ML binding to this transporter in the perspective of necessary high host/nematode selectivity. Such information on the key residues, in mammalian Pgp and in nematode Pgp, which interact with the different compounds, should also provide insights for the design of novel anthelmintic analogs which are non-ligands, thus unlikely to be subject to efflux-based mechanisms of drug resistance development in nematodes. The model that we have developed could hence generally be used to assess the possible interactions of new chemical entities with nematode Pgp. This model is a critical insight into the possible role of Pgp homologs in expelling AH drugs out of parasite nematodes, making them resistant to treatment. Furthermore, this is the necessary step to identify in combination with *in silico*-guided *in vitro* experiments, the ABC transporters (at least from the B sub-family) specifically responsible for AH resistance, in order to rationally design or select specific inhibitors that could be novel therapeutic weapons against AH resistant parasites.

Acknowledgements

We deeply acknowledge Chantal Lebrun and Nicolas Loiseau for valuable help with data analysis, presentation and comments on the manuscript. We also thank Fabien Jourdan and Clément Frainay for extensive technical support, and the GenoToul Bioinformatics hardware infrastructure that was used for computing. This work was supported by the Natural Sciences and Engineering Research Council of Canada, the FRQNT Centre for Host-Parasite Interactions, Quebec, and EMIDA ERA-NET project CARES n 11-EMID-003-02.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.ijpddr.2016.09.001.

References

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.

Ardelli, B.F., Prichard, R.K., 2013. Inhibition of P-glycoprotein enhances sensitivity of Caenorhabditis elegans to ivermectin. Vet. Parasitol. 191, 264–275.

Bessadok, A., Garcia, E., Jacquet, H., Martin, S., Garrigues, A., Loiseau, N., Andre, F., Orlowski, S., Vivaudou, M., 2011. Recognition of sulfonylurea receptor (ABCC8/9) ligands by the multidrug resistance transporter P-glycoprotein (ABCB1): functional similarities based on common structural features between two multispecific ABC proteins. J. Biol. Chem. 286, 3552–3569.

Bygarski, E.E., Prichard, R.K., Ardelli, B.F., 2014. Resistance to the macrocyclic lactone moxidectin is mediated in part by membrane transporter P-glycoproteins: implications for control of drug resistant parasitic nematodes. Int. J. Parasitol. Drugs Drug Resist 4, 143–151.

Cully, D.F., Vassilatis, D.K., Liu, K.K., Paress, P.S., Van der Ploeg, L.H., Schaeffer, J.M., Arena, J.P., 1994. Cloning of an avermectin-sensitive glutamate-gated chloride channel from Caenorhabditis elegans. Nature 371, 707–711.

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

Dupuy, J., Alvinerie, M., Menez, C., Lespine, A., 2010. Interaction of anthelmintic drugs with P-glycoprotein in recombinant LLC-PK1-mdr1a cells. Chem. Biol. Interact. 186, 280–286.

Eckford, P.D., Sharom, F.J., 2009. ABC efflux pump-based resistance to chemotherapy drugs. Chem. Rev. 109, 2989–3011.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.

Efferth, T., Volm, M., 1993. Reversal of doxorubicin-resistance in sarcoma 180 tumor cells by inhibition of different resistance mechanisms. Cancer Lett. 70, 197–202.

Elmshauser, S., Straehle, L.C., Kranz, J., Krebber, R., Geyer, J., 2014. Brain penetration of emodepside is increased in P-glycoprotein-deficient mice and leads to neurotoxicosis. J. Vet. Pharmacol. Ther. 38, 74–79.

Forrester, S.G., Prichard, R.K., Dent, J.A., Beech, R.N., 2003. Haemonchus contortus: HcGluCla expressed in Xenopus oocytes forms a glutamate-gated ion channel that is activated by ibotenate and the antiparasitic drug ivermectin. Mol. Biochem. Parasitol. 129, 115–121.

Godoy, P., Che, H., Beech, R.N., Prichard, R.K., 2015a. Characterization of Haemonchus contortus P-glycoprotein-16 and its interaction with the macrocyclic lactone anthelmintics. Mol. Biochem. Parasitol. 204, 11–15.

Godoy, P., Che, H., Beech, R.N., Prichard, R.K., 2016. Characterisation of P-glycoprotein-9.1 in Haemonchus contortus. Parasit. Vectors 9, 52.

Godoy, P., Lian, J., Beech, R.N., Prichard, R.K., 2015b. Haemonchus contortus P-glycoprotein-2: in situ localisation and characterisation of macrocyclic lactone transport. Int. J. Parasitol. 45, 85–93.

Haubertin, D.Y., Madaoui, H., Sanson, A., Guerois, R., Orlowski, S., 2006. Molecular dynamics simulations of E. coli MsbA transmembrane domain: formation of a semipore structure. Biophys. J. 91, 2517–2531.

Hayeshi, R., Masimirembwa, C., Mukanganyama, S., Ungell, A.L., 2006. The potential inhibitory effect of antiparasitic drugs and natural products on P-glycoprotein mediated efflux. Eur. J. Pharm. Sci. 29, 70–81.

Higgins, C.F., 1992. ABC transporters: from microorganisms to man. Annu. Rev. Cell Biol. 8, 67–113.

Jabbar, A., Iqbal, Z., Kerboeuf, D., Muhammad, G., Khan, M.N., Afaq, M., 2006. Anthelmintic resistance: the state of play revisited. Life Sci. 79, 2413–2431.

James, C.E., Davey, M.W., 2009. Increased expression of ABC transport proteins is associated with ivermectin resistance in the model nematode Caenorhabditis elegans. Int. J. Parasitol. 39, 213–220.

Janssen, I.J., Krucken, J., Demeler, J., von Samson-Himmelstjerna, G., 2013. Caenorhabditis elegans: modest increase of susceptibility to ivermectin in individual P-glycoprotein loss-of-function strains. Exp. Parasitol. 134, 171–177.

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

Jones, P.M., George, A.M., 2005. Multidrug resistance in parasites: ABC transporters, P-glycoproteins and molecular modelling. Int. J. Parasitol. 35, 555–566.

Kaplan, R.M., 2004. Drug resistance in nematodes of veterinary importance: a status report. Trends Parasitol. 20, 477–481.

Kaschny, M., Demeler, J., Janssen, I.J., Kuzmina, T.A., Besognet, B., Kanellos, T., Kerboeuf, D., von Samson-Himmelstjerna, G., Krucken, J., 2015. Macrocyclic lactones differ in interaction with recombinant P-glycoprotein 9 of the parasitic nematode Cylicocylus elongatus and ketoconazole in a yeast growth assay. PLoS Pathog. 11, e1004781.

Kerboeuf, D., Guegnard, F., 2011. Anthelmintics are substrates and activators of nematode P glycoprotein. Antimicrob. Agents Chemother. 55, 2224–2232.

Kiki-Mvouaka, S., Menez, C., Borin, C., Lyazrhi, F., Foucaud-Vignault, M., Dupuy, J., Collet, X., Alvinerie, M., Lespine, A., 2010. Role of P-glycoprotein in the disposition of macrocyclic lactones: a comparison between ivermectin, eprinomectin, and moxidectin in mice. Drug Metab. Dispos. 38, 573–580.

Koenderink, J.B., Kavishe, R.A., Rijpma, S.R., Russel, F.G., 2010. The ABCs of multidrug resistance in malaria. Trends Parasitol. 26, 440–446.

Lage, H., 2003. ABC-transporters: implications on drug resistance from microorganisms to human cancers. Int. J. Antimicrob. Agents 22, 188–199.

Laing, R., Hunt, M., Protasio, A.V., Saunders, G., Mungall, K., Laing, S., Jackson, F., Quail, M., Beech, R., Berriman, M., Gilleard, J.S., 2011. Annotation of two large contiguous regions from the Haemonchus contortus genome using RNA-seq and comparative analysis with Caenorhabditis elegans. PLoS One 6, e23216.

Laing, R., Kikuchi, T., Martinelli, A., Tsai, I.J., Beech, R.N., Redman, E., Holroyd, N., Bartley, D.J., Beasley, H., Britton, C., Curran, D., Devaney, E., Gilabert, A., Hunt, M., Jackson, F., Johnston, S.L., Kryukov, I., Li, K., Morrison, A.A., Reid, A.J., Sargison, N., Saunders, G.I., Wasmuth, J.D., Wolstenholme, A., Berriman, M., Gilleard, J.S., Cotton, J.A., 2013. The genome and transcriptome of Haemonchus contortus, a key model parasite for drug and vaccine discovery. Genome Biol. 14, R88.

Leonard, G.D., Fojo, T., Bates, S.E., 2003. The role of ABC transporters in clinical practice. Oncologist 8, 411–424.

Lespine, A., Martin, S., Dupuy, J., Roulet, A., Pineau, T., Orlowski, S., Alvinerie, M., 2007. Interaction of macrocyclic lactones with P-glycoprotein: structure-affinity relationship. Eur. J. Pharm. Sci. 30, 84–94.

Lespine, A., Menez, C., Bourguinat, C., Prichard, R.K., 2012. P-glycoproteins and other multidrug resistance transporters in the pharmacology of anthelmintics: prospects for reversing transport-dependent anthelmintic resistance. Int. J. Parasitol. Drugs Drug Resist. 2, 58–75.

Li, J., Jaimes, K.F., Aller, S.G., 2014. Refined structures of mouse P-glycoprotein. Protein Sci. 23, 34–46.

Loo, T.W., Bartlett, M.C., Clarke, D.M., 2006a. Transmembrane segment 1 of human P-glycoprotein contributes to the drug-binding pocket. Biochem. J. 396, 537–545.

Loo, T.W., Bartlett, M.C., Clarke, D.M., 2006b. Transmembrane segment 7 of human P-glycoprotein forms part of the drug-binding pocket. Biochem. J. 399, 351–359.

Loo, T.W., Clarke, D.M., 2001. Defining the drug-binding site in the human multidrug resistance P-glycoprotein using a methanethiosulfonate analog of verapamil, MTS-verapamil. J. Biol. Chem. 276, 14972–14979.

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

Mani, T., Bourguinat, C., Keller, K., Ashraf, S., Blagburn, B., Prichard, R.K., 2016. Interaction of macrocyclic lactones with a Dirofilaria immitis P-glycoprotein. Int. J. Parasitol. 46, 631–640.

Menez, C., Alberich, M., Kansoh, D., Blanchard, A., Lespine, A., 2016. Acquired tolerance to ivermectin and moxidectin after drug selection pressure in the nematode Caenorhabditis elegans. Antimicrob. Agents Chemother. 60, 4809–4819.

Moreno, Y., Nabhan, J.F., Solomon, J., Mackenzie, C.D., Geary, T.G., 2010. Ivermectin disrupts the function of the excretory-secretory apparatus in microfilariae of Brugia malayi. Proc. Natl. Acad. Sci. U. S. A. 107, 20120–20125.

Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., Olson, A.J., 2009. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J. Comput. Chem. 30, 2785–2791.

Osei-Atweneboana, M.Y., Awadzi, K., Attah, S.K., Boakye, D.A., Gyapong, J.O., Prichard, R.K., 2011. Phenotypic evidence of emerging ivermectin resistance in Onchocerca volvulus. PLoS Negl. Trop. Dis. 5, e998.

Osei-Atweneboana, M.Y., Eng, J.K., Boakye, D.A., Gyapong, J.O., Prichard, R.K., 2007. Prevalence and intensity of Onchocerca volvulus infection and efficacy of ivermectin in endemic communities in Ghana: a two-phase epidemiological study. Lancet 369, 2021–2029.

Prichard, R., Menez, C., Lespine, A., 2012. Moxidectin and the avermectins: consanguinity but not identity. Int. J. Parasitol. Drugs Drug Resist. 2, 134–153.

Roulet, A., Puel, O., Gesta, S., Lepage, J.F., Drag, M., Soll, M., Alvinerie, M., Pineau, T., 2003. MDR1-deficient genotype in Collie dogs hypersensitive to the P-glycoprotein substrate ivermectin. Eur. J. Pharmacol. 460, 85–91.

Schinkel, A.H., 1997. The physiological function of drug-transporting P-glycoproteins. Semin. Cancer Biol. 8, 161–170.

Schinkel, A.H., Smit, J.J., van Tellingen, O., Beijnen, J.H., Wagenaar, E., van Deemter, L., Mol, C.A., van der Valk, M.A., Robanus-Maandag, E.C., te Riele, H.P., et al., 1994. Disruption of the mouse mdr1a P-glycoprotein gene leads to a deficiency in the blood-brain barrier and to increased sensitivity to drugs. Cell 77, 491–502.

Seeger, M.A., van Veen, H.W., 2009. Molecular basis of multidrug transport by ABC transporters. Biochim. Biophys. Acta 1794, 725–737.

Shilling, R.A., Venter, H., Velamakanni, S., Bapna, A., Woebking, B., Shahi, S., van Veen, H.W., 2006. New light on multidrug binding by an ATP-binding-cassette transporter. Trends Pharmacol. Sci. 27, 195–203.

Sousa, S.F., Fernandes, P.A., Ramos, M.J., 2006. Protein-ligand docking: current status and future challenges. Proteins 65, 15–26.

Sousa, S.F., Ribeiro, A.J., Coimbra, J.T., Neves, R.P., Martins, S.A., Moorthy, N.S., Fernandes, P.A., Ramos, M.J., 2013. Protein-ligand docking in the new millennium–a retrospective of 10 years in the field. Curr. Med. Chem. 20, 2296–2314.

Szewczyk, P., Tao, H., McGrath, A.P., Villaluz, M., Rees, S.D., Lee, S.C., Doshi, R., Urbatsch, I.L., Zhang, Q., Chang, G., 2015. Snapshots of ligand entry, malleable binding and induced helical movement in P-glycoprotein. Acta crystallogr. Sect. D Biol. Crystallogr. 71, 732–741.

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

Yan, R., Urdaneta-Marquez, L., Keller, K., James, C.E., Davey, M.W., Prichard, R.K., 2012. The role of several ABC transporter genes in ivermectin resistance in Caenorhabditis elegans. Vet. Parasitol. 190, 519–529.

Zhao, Z., Sheps, J.A., Ling, V., Fang, L.L., Baillie, D.L., 2004. Expression analysis of ABC transporters reveals differential functions of tandemly duplicated genes in Caenorhabditis elegans. J. Mol. Biol. 344, 409–417.

Please cite this article in press as: David, M.A., et al., *In silico* analysis of the binding of anthelmintics to *Caenorhabditis elegans* P-glycoprotein 1, International Journal for Parasitology: Drugs and Drug Resistance (2016), http://dx.doi.org/10.1016/j.ijpddr.2016.09.001

Supplementary Information

In silico analysis of anthelmintics binding to *Caenorhabditis elegans* P-glycoprotein 1

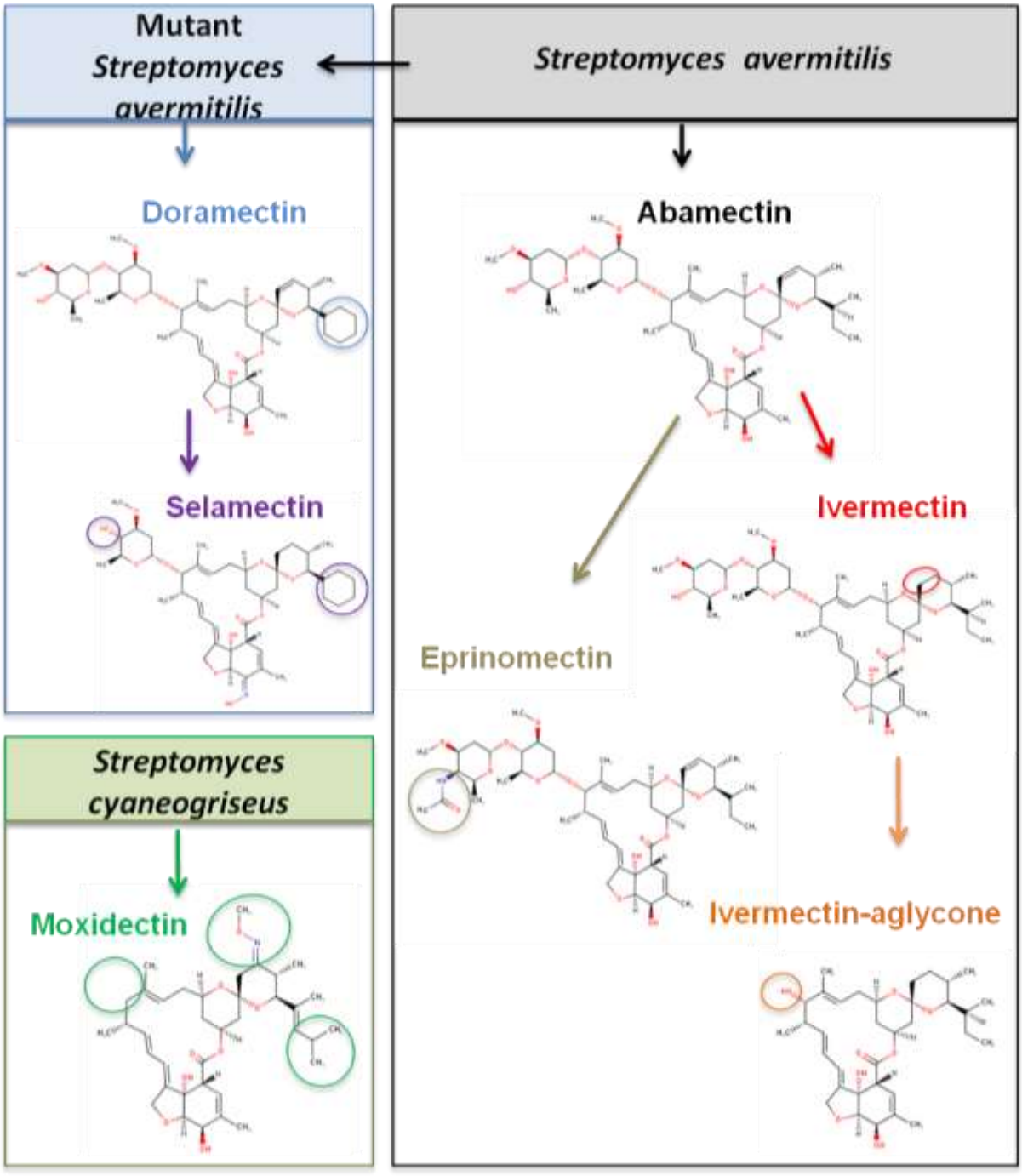
Marion David[a,b], Stéphane Orlowski[c], Roger K. Prichard[b], Shaima Hashem[c], François André[c,#] and Anne Lespine[a,#]

[a] Toxalim (Research Centre in Food Toxicology), Université de Toulouse, INRA, ENVT, INP-Purpan, UPS, Toulouse, France

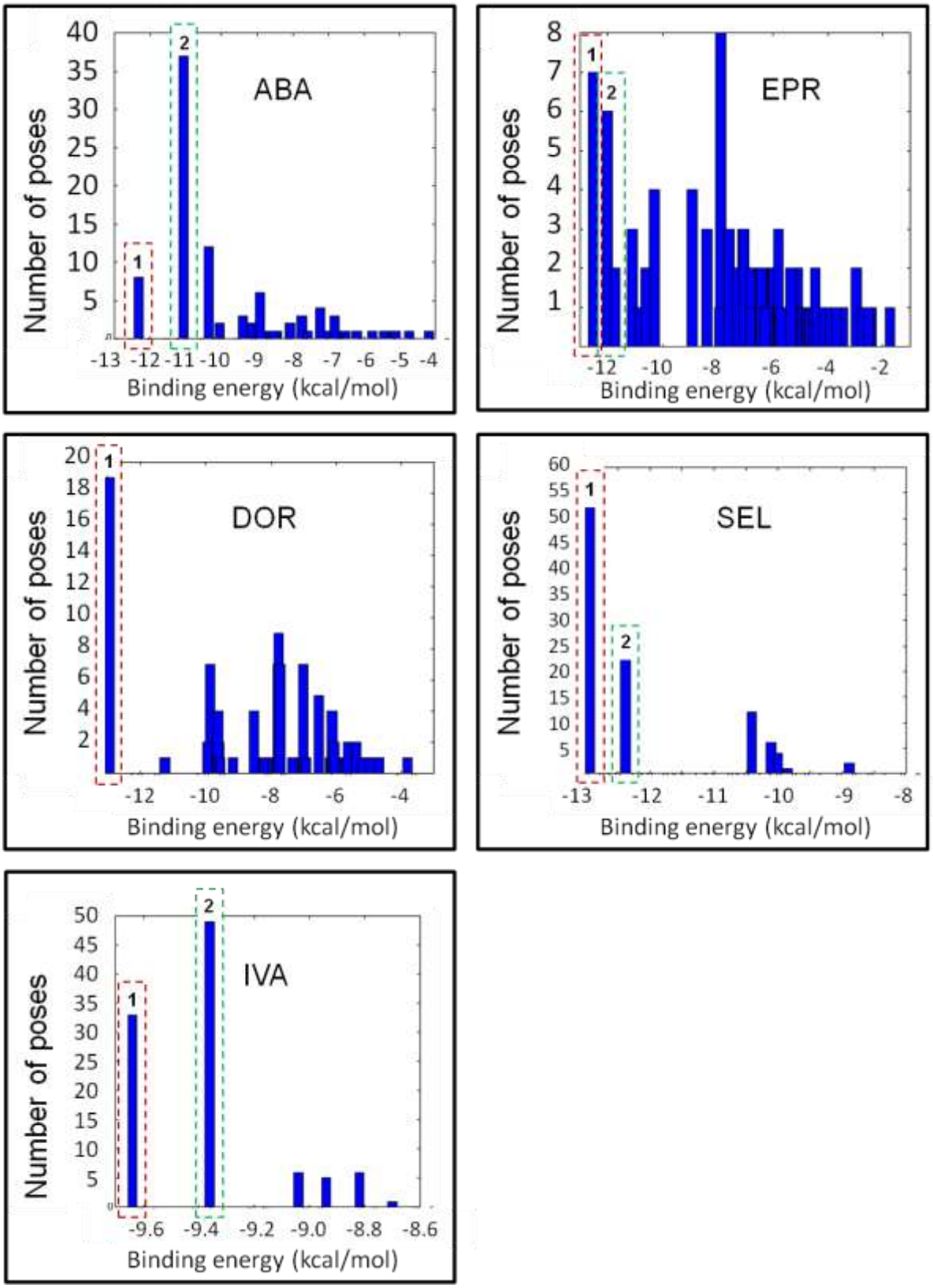
[b] Institute of Parasitology, McGill University, Sainte-Anne-De-Bellevue, Canada

[c] Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS UMR 9198, Univ Paris-Sud, Université Paris-Saclay, 91198, Gif-sur-Yvette cedex, France

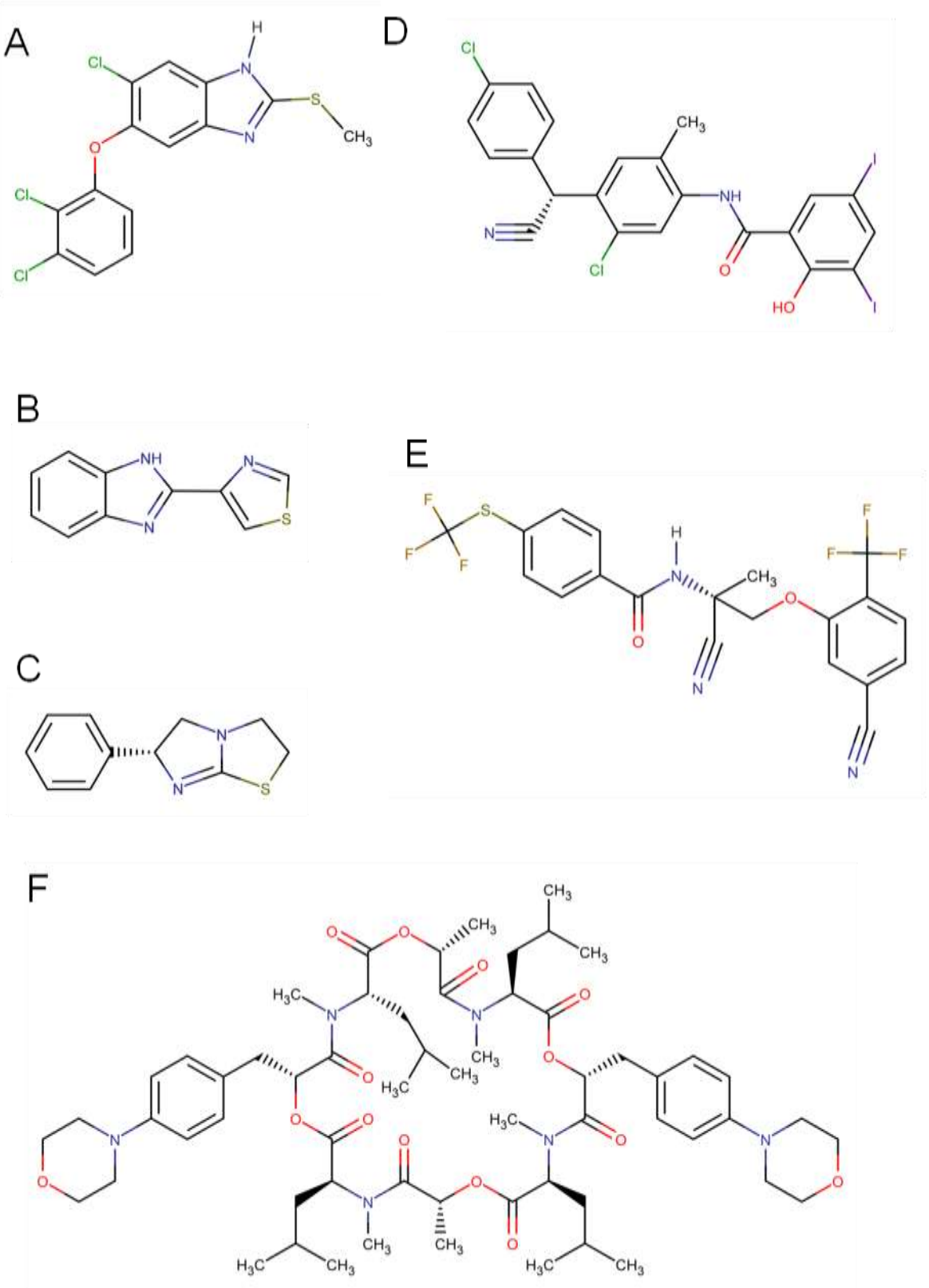
Supplementary Figures



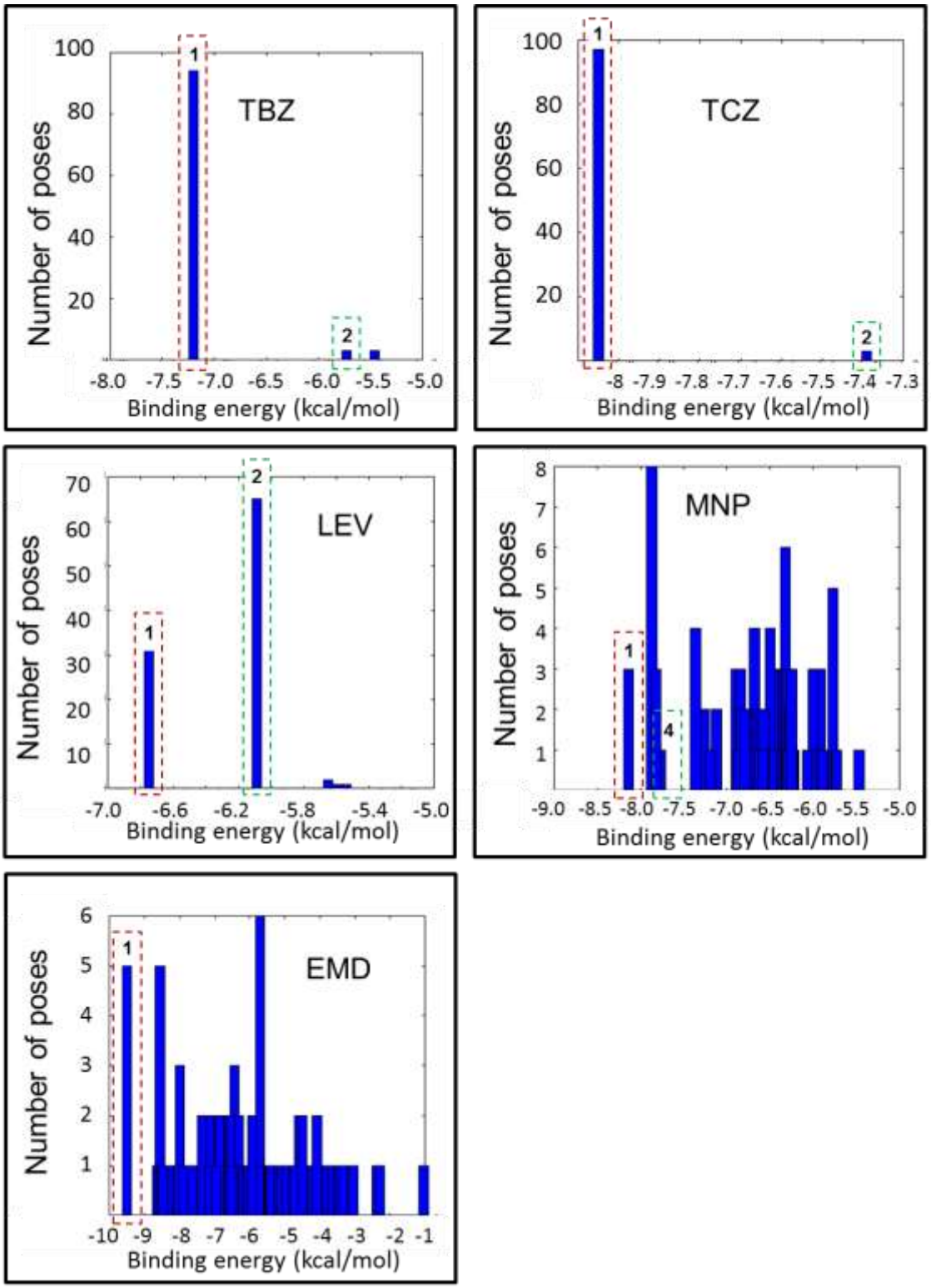
Supplementary Figure S1. Chemical structure and filiation of macrocyclic lactones. Abamectin (M1272), ivermectin (M6566), eprinomectin (M4959), doramectin (M4743), selamectin (M9831) and moxidectin (M7646). Circles indicate key substituents and functions.



Supplementary Figure S2. AutoDock 4.2 binding energy clustering of ML. Abamectin (ABA), eprinomectin (EPR), doramectin (DOR), selamectin (SEL) and ivermectin-aglycone (IVA). The first (1) and second (2) lowest energy clusters are framed with red and green dotted lines, respectively.



Supplementary Figure S3. Chemical structure of anthelmintics. Triclabendazole (CID50248) (A), thiabendazole (DB00730) (B), levamisole-R (DB00848) (C), closantel-R (CID42574) (D), monepantel-S (CID44449087) (E) and emodepside (CID6918632) (F).



Supplementary Figure S4. Binding energy clustering of AHs. Thiabendazole (TBZ), triclabendazole (TCZ), levamisole (LEV) monepantel (MNP) and emodepside (EMD). The first (1) and second (2) lowest energy clusters are framed with red and green dotted lines, respectively.

Supplementary Tables

Supplementary Table S1. List of "hotspot residues" that have been identified as being involved in multispecific substrate recognition in mammalian Pgp.

P-gp	Hsa-Pgp	Mmu-ABCB1a	Cgr-ABCB1	Mmu-ABCB1b	Origin	Cel-Pgp-1
TM1	**H61**	H60	H60	H60	M	T87
	G64	A63	A63	L63	M	G90
	L65	L64	L64	L64	**M/C/SL**	L91
	M68	**M67**	M67	L67	SL	M94
	M69	**M68**	M68	M68	S	S95
	F72	**F71**	F71	F71	SL	Q98
TM2	Y118	**Y114**	Y115	Y117	SL	Y142
TM4	**S222**	S218	S219	S221	C	T246
TM5	I299	**M295**	M296	I298	SL	F323
	F303	**F299**	F300	Y302	SL	Q327
	L304	**L300**	L301	L303	SA	A328
	I306	I302	I303	V305	**M/SL**	N330
	Y307	**Y303**	Y304	Y306	S	F331
	Y310	**Y306**	Y307	Y309	SL	F334
TM6	**F335**	F331	F332	F334	M	F359
	F336	**F332**	F333	F335	S	S360
	V338	V334	V335	I337	M	V362
	L339	**L335**	L336	L338	**C/S**	M363
	I340	**I336**	I337	L339	**C/S**	M364
	G341	G337	G338	G340	M	G365
	A342	A338	A339	T341	C	S366
	F343	**F339**	F340	F342	S	M367
	S351	D347	D348	**D350**	M	Q375
TM7	N721	**N717**	N718	N719	SL	G764
	Q725	**Q721**	Q722	Q723	S	Y768
	F728	**F724**	F725	F726	**C/S**	Y771
	F732	**F728**	F729	F730	S	F775
TM8	L762	**L758**	L759	M760	SA	L803
	F770	**F766**	L767	Y768	SL	S811
TM9	T837	**F833**	T834	T835	SA	I878
	I840	I836	**I837**	V838	M	L881
	A841	A837	A838	A839	C	V882
	N842	N838	**N839**	N840	M	S883
TM10	I864	I860	I861	**I862**	M	I905
	I867	I863	I864	**L865**	M	I908
	I868	I864	I865	I866	C	V909
	I870	I866	I867	**L868**	M	F911
TM11	A935	A931	A932	**A933**	M	A976
	F938	**F934**	F935	F936	M	Q979
	F942	**F938**	F939	F940	**M/C**	Y983
	S943	**S939**	S940	**S941**	M	G984
	T945	**T941**	T942	T943	**M/C**	A986
	Q946	**Q942**	Q943	Q944	M	S987
	M949	**M945**	M946	M947	SL	L990
	Y950	**Y946**	Y947	Y948	M	Y991
	S952	**S948**	S949	S950	M	L993
	Y953	**Y949**	Y950	Y951	**M/S**	N994
	F957	**F953**	F954	F955	**M/SL**	Y998
TM12	**L975**	L971	L972	M973	**M/C/SL**	L1018
	F978	**F974**	F975	F976	**M/S**	M1021

S979	**S975**	S976	S977	S	Y1022
V981	I977	I978	V979	**M/C**	I1024
V982	**V978**	V979	V980	**C/S**	T1025
F983	F979	F980	F981	**M/SL**	I1026
G984	G980	G981	G982	C	S1027
A985	**A981**	A982	A983	SA	T1028
M986	**M982**	M983	M984	S	S1029
A987	**A983**	A984	A985	SL	T1030
G989	**G985**	G986	G987	SA	G1032
Q990	**Q986**	Q987	N988	S	F1033
V991	**V987**	V988	T989	SL	A1034
S993	**S989**	S990	S991	SA	S1036

Hotspots are distributed among the various TM helices, as quoted in the first column. The orthologs: human (Hsa), mouse (Mmu) a and b isoforms and Chinese hamster (Cgr), in which the residue has been indicated, are shown in columns 2-5 by a bold residue symbol. Hotspots were determined, in the literature, by different experimental approaches (point mutagenesis, chemical labelling, co-crystallization, respectively denoted as M, C, S), as quoted in the sixth column; when the residue has been found by at least two different approaches, the symbols are highlighted in bold. M: cytotoxicity assay after Pgp site-directed mutagenesis; C: MTS-mediated chemical labelling by a MTS-derivative of Cys-free mutated Pgp bearing Cys residues at specific sites for scanning mutagenesis; S: contact residues of the co-crystallized ligands QZ59RRR and QZ59SSS in the Pgp crystal structures published in Aller et al. (2009) and Li et al (2014), but presenting few differences (A: residues found in the 2009 structure alone, L: residues found in the 2014 structure alone). See Computational Methods section for the relevant references. Last column: corresponding residues found in Cel-Pgp-1 after sequence alignment with mammalian Pgp.

Supplementary Table S2. List of interacting residues of each transmembrane helix of Cel-Pgp-1 with the first or second lowest energy clusters of ivermectin (IVM2), abamectin (ABA2), eprinomectin (EPR2), selamectin (SEL1), ivermectin-aglycone (IVA1), moxidectin (MOX1).

Molecule	IVM	ABA	EPR	SEL	IVA	MOX
Cluster Rank	2nd	2nd	2nd	1st	1st	1st
TMa-b	D23 K26 K30	E22 K26	E22 K26		E22	E22 D23 K26
TM1		M94 Q98	L91	L91 M94		L91 M94
TM3	Y220					
TM6	M367 G370 L371	M367 L371	M364	M363 M367	M363	S360 M364
TM7				Y771 F775		
TM10	V909 Q913	Q913 R916	V909 Q913 R916	Q913	L906 V909 Q913	L906 V909 Q913
TM11	L990 L993	A986 S987 L990 N994	A986 S987 L990 L993	L990 N994 Y998	L990	L990 L993 N994
TM12	T1025 T1028	T1025 S1029 G1032	T1025 T1028 S1029 L1031 G1032 F1033	M1021 Y1022 T1025 I1026	T1025 I1026 T1028 S1029	Y1022 I1024 T1025 T1028 S1029 G1032

Bold: hotspot residues. Underscored: residues establishing a H-bond.

PART II

Identification, localization and functional characterization of *H. contortus* P-glycoprotein 13

The second main objective was to elucidate the function of Hco-Pgp-13 in this parasite. For this purpose, we first wanted to check the cDNA sequence of the closest homolog of *Cel-pgp-13* in *H. contortus*. We designed primers on the sequence of *Hco-pgp-13* predicted by high-throughput sequencing by the Sanger Institute (Laing et al., 2013), and amplified its cDNA from a susceptible isolate of *H. contortus*. After precision of its full-length cDNA sequence and translation into protein sequence, we showed that Hco-Pgp-13 is a very close homolog to the three proteins Cel-Pgp-12, Cel-Pgp13 and Cel-Pgp-14.

We also found a high degree of homology between the trans-membrane domains of Hco-Pgp-13 and Cel-Pgp-1. As we had previously checked, in the first part of this work, that the crystal structure of *C. elegans* could be used for docking calculations in order to investigate its substrate profile, we modeled the 3D structure of Hco-Pgp-13 by homology with Cel-Pgp-1. We then docked several compounds of interest to screen for putative ligands of Hco-Pgp-13 and found a predicted multispecific property for the binding domain of this parasitic transporter. Various *in vitro* techniques using heterologous systems allowed correlating some of these results, in particular for actinomycin D that was found to stimulate the ATPase activity of Hco-Pgp-13, consistent with its high affinity found *in silico*. More data will be required to precise these results, and in particular to investigate the interaction possibility of ML with Hco-Pgp-13.

Finally, to better understand its role in the worm, we investigated the localization of Hco-Pgp-13 in the L3 larvae and adults. The expression of Hco-Pgp-13 was found in the hypodermis and seam cells of both stages, in the epithelial gonad of at least L3 larvae, in the pharyngeal glands or neurons and in head neurons, possibly the amphids, of adult worms. This wide tissue distribution matching that of Cel-Pgp-12, 13 and 14, and the nature of stained organs, are consistent with an important function for Hco-Pgp-13 in the worm, which correlates well with the predicted multi-recognition capacity of this ABC transporter.

A. Manuscript n°3 in preparation:

Characterization of *Haemonchus contortus* P-glycoprotein 13

Marion David[a,b], Chantal Lebrun[a], Thomas Duguet[b], Franck Talmont[c], Robin Beech[b], Stéphane Orlowski[d], François André[d], Anne Lespine[a] and Roger K. Prichard[b#]

[a] Toxalim, Université de Toulouse, INRA, ENVT, INP-Purpan, UPS, Toulouse, France

[b] Institute of Parasitology, McGill University, Sainte-Anne-De-Bellevue, Canada;

[c] Institute of Pharmacology and Structural Biology, UMR 5089 CNRS, Toulouse, France

[d] CEA, IBiTec-Saclay, SB2SM and UMR9198 CNRS, I2BC; 91191 Gif-sur-Yvette cedex, France.

[#]Corresponding author:

Roger Prichard : roger.prichard@mcgill.ca
Tel : 514-398-7729
Address : Parasitology Building P-108
McGill University, MacDonald Campus
21111, Lakeshore Road
St Anne de Bellevue, QC H9X 3V9, Canada

Abstract

Export of various compounds through ABC transporters has become an increasingly studied process due to the involvement of some of these proteins in drug resistance. P-

glycoproteins (Pgps) have been the main ABC transporters characterized as having multispecific binding capacities in all living kingdoms. In nematodes, several Pgps are found in various parasites and the role of each of them has rarely been studied individually. *Haemonchus contortus* is the most economically important parasite of small ruminants. Deciphering the role of the 10 Pgps expressed in this parasite is thus of major importance to overcoming resistance of this species to various anthelmintic drugs. Here we focused on Hco-Pgp-13 due to the expression in the amphids of its closest ortholog in *Caenorhabditis elegans*. Indeed, the amphids represent a putative entry route of drugs to reach AH targets in the nervous system and have been shown to be linked to AH sensitivity in *C. elegans* and *H. contortus*. In our study, we corrected the previously predicted sequence of *Hco-pgp-13* and found that its translation product was the closest ortholog of Cel-Pgp12, 13 and 14 in *C. elegans*. A 3D model of Hco-Pgp-13 was constructed by homology with the PDB structure 4F4C of Cel-Pgp-1 crystal and indicated a topology matching an ABC transporter. To investigate the drug binding capacities of Hco-Pgp-13, the in silico docking calculation of actinomycin D, a well described substrate of mammalian Pgps and Cel-Pgp-1, was performed on the 3D model of Hco-Pgp-13. A high affinity of actinomycin D was predicted for the transporter and was correlated with *in vitro* stimulation of the ATPase activity of Hco-Pgp-13 by this drug. Finally, the localization of Hco-Pgp-13 in the parasite was found in many tissues, including epithelial cells and pharyngeal neurons, corresponding to the location of Cel-Pgp-12, 13 and 14. Overall, these data are consistent with a possible role for Hco-Pgp-13 in drug transport, which could possibly contribute in drug resistance and survival of this parasite when exposed to toxic exogenous compounds.

Keywords

Nematodes; *Haemonchus contortus*; ABC transporters; P-glycoprotein.

Abbreviations

ABC = ATP-binding cassette; ACD = actinomycin D; AH = anthelmintic; BLASTP = protein Basic Local Alignment Search Tool; Cel = *Caenorhabditis elegans*; Ceg = *Cylicocylus elongatus*; Hco = *Haemonchus contortus*; Hsa = *Homo sapiens*; IVM = ivermectin; MDR = multidrug resistance; ML = macrocyclic lactone(s); Mmu = *Mus musculus*; MOX = moxidectin; NBD = nucleotide binding domain; PDB = Protein data bank; Pgp = P-glycoprotein; QMEAN

= Qualitative Model Energy Analysis; RMSD = root mean square deviation; SNP = single nucleotide polymorphism; TMD = transmembrane domain.

I. INTRODUCTION

ATP-binding cassette (ABC) transporters belong to a large family of ATP hydrolyzing proteins involved in all living kingdoms in the transport of a wide variety of compounds. Among these transporters, the multi-drug resistance (MDR) ABC transporters are more specifically dedicated to the transport of xenotoxics and they have been clearly involved in resistance to drugs in mammals as well as in bacteria and parasites (Jones and George, 2005; Koenderink et al., 2010; Lage, 2003). In mammals, MDR ABC transporters such as P-glycoprotein (MDR1/ABCB1/Pgp) can efflux various, structurally unrelated drugs, and Pgp is responsible for resistance to chemotherapy in mammalian tumor cells (Eckford and Sharom, 2009; Leonard et al., 2003). Noticeably, mammalian Pgp transports macrocyclic lactones (ML), which are today the main anthelmintic medicines used against parasitic nematodes in humans and animals (Lespine et al., 2007; Roulet et al., 2003; Schinkel et al., 1994).

The function of these transporters is tightly linked to their conserved structures. Active MDR exporters of the ABC B sub-family are composed of two nucleotide binding domains (NBDs) and two transmembrane domains (TMDs). NBDs are the most conserved domains, and contain sequences involved in the binding and hydrolysis of ATP, that are hallmarks for belonging to the ABC family across species (Higgins, 1992). Across the lipid bilayer, TMDs are formed of 6 helices, and they combine by two, either in a dimer or in a tandem-like monomer, to form a funnel-like shape that delineates an inner chamber. The latter is either open to the cytoplasmic side of the membrane, where substrates can bind with high affinity to the transmembrane domains, or to the extracellular medium, with lower affinity of the TMDs to the substrates, which allows their release.

Many homologs of Pgps are expressed in nematodes as compared to the single ABCB1 found in human. However, there are three other full-sized members of the B sub-family (B4, B5, B11) in human, which present close sequence homologies but handle different types of substrates. This raises the question of the substrate selectivity, and hence the biological functions, of these various nematode Pgps. The free-living nematode *Caenorhabditis elegans*, a model nematode, and genetically very close to other clade V nematodes such as *Haemonchus contortus*, expresses 60 ABC transporters, among which 14 full-size Pgp homologs are localized in different organs and expressed at various stages of development (Zhao et al., 2004).

H. contortus is among the most prevalent pathogen parasitic nematodes in small ruminant veterinary medicine, and its genome has recently been sequenced and annotated (ftp://ftp.sanger.ac.uk/pub/pathogens/Haemonchus/contortus) (Laing et al., 2011; Laing et al., 2013). In this species, 10 homologs of Pgps were identified, and a few of them have been localized: Hco-Pgp-2 in the pharynx, anterior intestine and head neurons of *H. contortus*, and Hco-Pgp-9.1 in the uterus of females (Godoy et al., 2016; Godoy et al., 2015b).

Today, increasing efforts are made to better understand the respective functions of nematode Pgps and MRPs. In particular, in *C. elegans*, Cel-Pgp-3 has been suggested to be involved in the resistance to chloroquine and colchicine (Broeks et al., 1995), while Cel-Pgp-1 and Cel-MRP-1 were linked to heavy metal sensitivity in this free-living nematode (Broeks et al., 1996). Cel-Pgp-2 was found to be related to lysosome formation and lipid storage in the intestine of *C. elegans* (Nunes et al., 2005; Schroeder et al., 2007). Cel-Pgp-5 was found to be involved in resistance to bacterial infection and heavy metals (Kurz et al., 2007).

In addition, there is clear evidence that some nematode Pgps can transport ML, both in *C. elegans* (Ardelli and Prichard, 2013; Janssen et al., 2013b) and in *Parascaris equorum* (Janssen et al., 2013a; Janssen et al., 2015). Some mammalian Pgp inhibitors alter the transport function of Pgps in *C. elegans* and in the parasitic nematodes *H. contortus*, *Cylicocylus elongatus* and *Dirofilaria immitis* (Godoy et al., 2015a, 2016; Godoy et al., 2015b; Kaschny et al., 2015; Mani et al., 2016). Such inhibitors can also improve ML susceptibility in nematodes (Bartley et al., 2009; James and Davey, 2009; Lespine et al., 2012; Menez et al., 2016), revealing that there could be some substrate selectivity overlapping among species. In addition, increased ML resistance is associated with induction of expression of individual Pgp genes (James and Davey, 2009; Lespine et al., 2012). Recently, the simultaneous release of the first crystal structure of a Pgp of nematode, Cel-Pgp-1, resolved at a resolution of 3.4 Å (Protein Data Bank code: 4F4C) and the demonstration of ATPase activation of purified Cel-Pgp-1 by various mammalian Pgp substrate drugs provide the most compelling evidence of a multidrug transport function for a nematode Pgp (Jin et al., 2012). Furthermore, the high affinity binding of several anthelmintic drugs on Cel-Pgp-1 has been predicted by *in silico* docking calculations performed on a 4F4C structure (manuscript under review, Part I, A).

Because of their strategic tissue expression locations and their putative implication in drug resistance, it is important to gain insight into the characteristics of Pgps of parasitic nematodes. Here we focused on *H. contortus* Pgp-13 (Hco-Pgp-13). Indeed, the closest ortholog of its coding sequence in *C. elegans*, *Cel-pgp-13*, was found to be expressed in the amphids,

which contain chemo-sensitive neurons, playing a role in ML sensitivity (Dent et al., 2000; Freeman et al., 2003; Menez et al., 2016; Urdaneta-Marquez et al., 2014). The *Hco-pgp-13* cDNA sequence and translation into protein sequence was investigated, and its transport capacity was addressed, as well as the localization of expression of Hco-Pgp-13 in the nematode. Based on the predicted cDNA sequence obtained from large scale sequencing (Laing et al., 2013), we synthesized *Hco-pgp-13* cDNA from *H. contortus* RNA and deduced the protein sequence. The 3D conformation of the protein was also modeled in an open inward-facing conformation based on the crystal structure of Cel-Pgp-1 (PDB: 4F4C) (Jin et al., 2012). This model was used for the *in silico* study of the binding on Hco-Pgp-13 of the well-known mammalian Pgp substrate, actinomycin D. Finally, using the nematode Pgp heterologously expressed in *Pichia pastoris*, the capacity of actinomycin D to stimulate the ATPase activity of Hco-Pgp-13 was evidenced *in vitro*. This suggests its possible involvement in the transport of exogenous compounds, even if the additional possibility of lipid handling cannot be excluded.

II. MATERIAL AND METHODS

1. Parasites

The PF23 strain of *H. contortus* used is susceptible to AHs (Ranjan et al., 2002). Worms were originally supplied by Fort Dodge Animal Health, Princeton, NJ, USA and are maintained by our laboratory. Animals and standardized operating procedures used in this research study were approved (Protocol 3845) and subjected to the guidelines from the Animal Care Committee of McGill University, Canada. Worms were obtained from passages consisting of an artificial infection with the larvae from the previous generation in naive lambs, without anthelmintic exposure. They were then collected from the abomasum of the host and incubated in PBS at 37°C before storage at -80°C.

2. RNA extraction and reverse transcription.

Total RNA was extracted from 20 adult *H. contortus* homogenized and extracted with Trizol® reagent according to the instruction of the manufacturer (ThermoFisher, Canada). The RNA was finally dried, and dissolved in 30µL of RNase-free water. RNA concentration was determined with a Nanodrop photometer IMPLEN® at a wavelength of 260nm. The quality of the RNA was assessed by running 2µL of each sample in a 0.8% agarose gel. Extracted RNA of good quality was stored at -80°C.

The reverse transcription to cDNA was performed using the SuperScript® III reverse transcriptase (ThermoFisher, Canada), starting with 1µg RNA and following the instructions of the manufacturer. The cDNA obtained was stored at -20°C for further use.

3. Amplification of *Hco-pgp-13* cDNA sequence

A pair of primers, mda 4 – mda 1 (Suppl. Table S1), was constructed using the Geneious bioinformatics software, version 5.5.6., across the 3' end of the predicted sequence contigs of *Hco-pgp-13* obtained from the Sanger Institute (ftp://ftp.sanger.ac.uk/pub/pathogens/Haemonchus/contortus) (Laing et al., 2013). A first PCR was run using these primers and the reverse transcribed cDNA of whole adult *H. contortus* as template. A fragment of 3488bp was obtained and sequenced twice (Genome Quebec Innovation Centre, McGill University, QC, Canada) using 8 primers, mda 1 to mda 16 (Suppl. Table S1). These had been designed along the gene so that they lead to overlapping fragments. In total, at least 2 different sequences were obtained for each amplification, with one primer along the gene. The full 3488bp sequence obtained was checked to be identical for the majority of fragments obtained at each location.

Four more primers were designed to identify the 5'-end of *Hco-pgp-13* by nested PCR: the nematode spliced leader sequence SL1 (Blaxter and Liu, 1996), a specific forward primer mda 48 (Suppl. Table S1) designed between SL1 and the start codon, and two specific reverse primers mda 19 and mda 31 (Suppl. Table S1) designed on the 5' end of the 3488 bp fragment previously amplified with mda4 - mda1. A nested PCR was performed, first using SL1 and mda 31 on whole *H. contortus* cDNA. The PCR product thus obtained, and migrating at the expected length (around 1 kb), was extracted and subjected to a second PCR with the second inner forward primer mda 48 and one of the two inner reverse primers mda 19 or mda 31. The 887 and 999 bp PCR products thus obtained were sequenced and found to be perfectly identical along their overlap. They were then aligned to the first 3488 bp product with MultAlin software (Corpet, 1988), and the identity of their overlap was confirmed.

The mda 48 – mda 19 fragments as well as small fragments within the 3' sequence amplified with mda 4 – mda 1 were cloned in TOPO 2.1+ vectors (Invitrogen). After transformation of these vectors in *E. coli* TOP10F' competent cells, the bacterial cells were incubated overnight at 37°C on LB agar plates with 50µg/mL ampicillin, 40 µg/mL X-gal and 100mM IPTG. Colonies were screened by PCR and positive ones were sequenced (Genome Quebec Innovation Centre, McGill University, QC, Canada). Once the full-length sequence was

confirmed, it was aligned against the cDNA of the predicted sequence of *Hco-pgp-13*, to check for its identity using MultAlin (Suppl. Figure S1).

4. Prediction of Hco-Pgp-13 protein sequence and phylogenic analysis

The translation of *Hco-pgp-13* cDNA into protein was predicted using ExPASy – Translate tool (http://web.expasy.org/translate/). The parameters of the predicted protein were analyzed with ExPASy – ProtParam. The presence of signature motifs, and N- and O-glycosylation motifs were predicted using ExPASy - ScanProsite tool (http://prosite. expasy.org/) according to homology with proteins in the database. The Walker B motif specific for ABC transporters was not found by this program and was deduced from a Muscle (Edgar, 2004) multiple alignment (Suppl. Figure S2) with Pgps of various organisms including Hsa-Pgp in which all motifs have been previously identified. The prediction of amino-acids located within the transmembrane bilayer was performed with the Protter tool, (Omasits et al., 2014) and the TMHMM Server v. 2.0 (http://www.cbs.dtu.dk/services/ TMHMM/) (Suppl. Table S5). The visualization of the topology of the full length protein sequence across the plasma membrane was represented using the Protter tool. The TM helices and eight N-glycosylation motifs were represented as predicted by this program, N518 being the only one not previously predicted by ExPASy - ScanProsite.

A phylogenetic tree with all Hco-Pgps, Cel-Pgp-12, 13 and 14, Hsa-Pgp and Mmu-ABCB1a protein sequences was constructed using Geneious software with the mammalian Pgps as outgroups, after exclusion of NBDs from all the considered protein sequences. For comparison of protein sequences, a multiple alignment was performed using Muscle (Edgar, 2004) with all Hco-Pgps and Cel-Pgps and the mammalian Pgps: Hsa-Pgp and three Mmu-Pgps (ABCB1a, B1b and B4) (Suppl. Figure S2). The locations of the TMDs of all these proteins, except Mmu-ABCB1b and Mmu-ABCB4, were determined according to this alignment. The first and last amino-acids of TMD1 and TMD2 of each Pgp were identified as those aligned to the first amino-acid of TM1, the last amino-acid of TM6, to the first amino-acid of TM7 and the last amino-acid of TM12 of Cel-Pgp-1, respectively, as given by the 4F4C 3D structure. The similarity and identity percentages of the TMDs of Cel-Pgp-1 with the TMDs of all other Pgps was then determined using BlastP (http://blast.ncbi.nlm. nih.gov/Blast.cgi). TMD1 and TMD2 were considered independently and the mean of the two values is shown in Table 1.

5. Construction of 3D models of Hco-Pgp-13 based on Cel-Pgp-1 4F4C PDB structure as template and *in silico* docking calculations

All Modeller and Autodock calculations were performed using the computing facilities of the CEA-DSV (cluster Gabriel) at Saclay, and of INRA (Genotoul) in Toulouse.

3D models of Hco-Pgp-13 were built using Modeller9v12 (Sali and Blundell, 1993; Webb and Sali, 2014) and the crystal structure of *C. elegans* Pgp (PDB: 4F4C), determined at a resolution of 3.4 Å, as template. As the Cel-Pgp-1 sequence was lacking the following AAs: M1-R3, A52-E54, K666-E715 and G1307-K1321, the full-length sequence of Hco-pgp-13 was modelled without the corresponding residues M1-S3, E668-L725, and G1304-T1317. A multiple alignment was performed with all Cel- and Hco-Pgps as well as 4 mammalian Pgps: Hsa-Pgp and Mmu-ABCB1a/1b/4 to obtain the highest possible precision (Suppl. Figure S2). The pairwise alignment was then deduced (Suppl. Figure S3) and subjected to Modeller to generate 100 models. The best DOPE (Discrete Optimized Protein Energy) score model n°4 (Hco-Pgp-13_04) and the best molecular PDF model n°52 (Hco-Pgp-13_52) were submitted to various online metaservers to assess their quality (Suppl. Table S2). The QMEAN (Qualitative Model Energy Analysis) scoring function (Benkert et al., 2009; Benkert et al., 2008), the ProSA-web (Protein structure analysis) (Sippl, 1993; Wiederstein and Sippl, 2007) and the VADAR (Volume, Area, Dihedral Angle Reporter) (Willard et al., 2003) servers all indicated very close accuracies for the two models according to different parameters, so that both of them were chosen for docking calculations.

The *in silico* docking calculation of actinomycin D (ACD) on Hco-Pgp-13 was performed as described in detail in (manuscript under review, Part I, A). The structure of ACD was extracted from Drugbank n° DB00970 and 10 minimum energy conformers were generated with Marvin Suite under the MMFF94 force field (https://www.chemaxon.com/products/marvin/marvinsketch/). Four representative conformers were chosen as starting conformations for the docking after alignment and calculation of RMSD under PyMOL. Molecular docking experiments were performed using AutoDock 4 (release 4.2.6) in the semi-flexible mode, with the Hco-Pgp13_04 and Hco-Pgp13_52 PDB structures kept rigid, and prepared with AutoDock Tools (Morris et al., 1998; Morris et al., 2009). The grid built by AutoGrid 4 included 95, 120, and 100 points in x, y, and z directions, with a grid spacing of 0.375 Å, to allow a good compromise between resolution of the explored volume and the size of the binding area (box dimensions 35.6x 45 x 37.5 Å, centered in the inner cavity of Hco-Pgp-13, at the point x=23 Å; y=78 Å; z=-2 Å). For each ligand conformer, 100 independent

calculations were performed using the Lamarckian genetic algorithm. All the other parameters were set at the default value. The 100 generated poses were assigned a score calculated by AutoDock that can be considered as an estimated free energy of ligand binding (indicative of binding affinity). They were then clustered as a function of the closeness of their positions and conformations, with RMSD set at 2.0 Å, and finally ranked by their binding score (for the best pose in the cluster). The results are displayed in an energy scores histogram. Different parameters and observables can be used in the interpretation of docking data issued by AutoDock: binding energies (i.e. docking scores), histogram bars energy range, profile of the histogram, and location of calculated positions in the protein structure. We considered only the lowest energy cluster since it contained the maximum of poses for each model of Hco-Pgp-13 and thus was the most representative of all.

For each lowest energy pose of selected clusters, the number and nature of interacting residues were analyzed within the protein. Among these, particular interest was given to residues belonging to the "hotspots for drug binding", described in (manuscript under review, Part I, A). These are a collection of 62 residues, as displayed in (Suppl. Table S3), coming from different experimental approaches that have been conducted for the purpose of determining the key residues responsible for multidrug recognition by mammalian Pgp (Hsa-ABCB1, Mmu-ABCB1a & B1b, Cgr-ABCB1) (Aller et al., 2009; Bessadok et al., 2011; Li et al., 2013; Loo et al., 2006a, b; Loo and Clarke, 2001, 2002; Shilling et al., 2006). All these residues are situated in the transmembrane part of the protein and 14 of these 62 residues are common between at least two different approaches. Altogether, they provide a frame in the inner chamber that offers a set of anchoring points for multi-specific recognition and binding, and eventual translocation, of various transport ligands. Multiple protein sequence alignments have been performed on Hco-Pgp-13, Cel-Pgp-1, human ABCB1, murine ABCB1a and B1b, and Chinese hamster ABCB1, using Muscle software (Edgar, 2004), to identify the corresponding residues in Hco-Pgp-13.

6. Cloning and transfection of *Hco-pgp-13* gene in *Pichia pastoris*

The full-length identified sequence of cDNA sequence of Hco-Pgp-13 was subjected to codon optimization by GenScript (USA). After transformation of the vector in *E. coli* TOP10F' competent cells, *Hco-pgp-13* codon optimized sequence was confirmed by sequencing (Genome Quebec Innovation Centre, McGill University, QC, Canada). *Hco-pgp-13* coding sequence was modified at the 5' end and the 3' end by introducing BstBI and XbaI sites,

respectively, by PCR amplification with the following primers: BstBI-hco-pgp-13: 5'AAAACAACTTAATTATTATTCGAAACGATGACATCAAAACCCGAT'3 (forward) and XbaI-hco-pgp-13: 5'TAGCTAGCTAGCTAGCTAGTGTTCTAGAGGCCCTGTG GTGAGGTCCTGC'3 (reverse). Modified PCR cDNA was digested using BstBI and XbaI enzymes and cloned in the pPICZ-HuMOR-cmyc-his-tag vector (Sarramegna et al., 2005) digested with the same enzymes leading to the creation of pPICZ-Hco-pgp13-cmyc-his-tag vector.

Escherichia coli strain Top10F' was used for the propagation of recombinant plasmids. *E. coli* transformants were selected on low salt LB plates pH 7.5 (0.5% w/v yeast extract, 1% w/v tryptone, 0.5% w/v NaCl, 1.5% w/v bacteriological agar) supplemented with 25 µg zeocin/ml. *Pichia pastoris* SMD1163 (his4, pep4, prB1) strain was used for receptor expression. *P. pastoris* transformants were selected on YPDS plates (1% w/v yeast extract, 2% w/v peptone, 2% w/v dextrose, 1 M sorbitol, and 1.5% w/v bacteriological agar) with 100 µg zeocin/ml. *P. pastoris* growth and induction media were BMGY (1% w/v yeast extract, 2% w/v peptone, 0.1 M phosphate buffer pH 7.5, 1% v/v glycerol) and BMMY (same as BMGY except that glycerol was replaced by 0.5% v/v methanol), respectively. For some experiments growth and recombinant protein expression were realized at the same time in BMMY. Cell cultures were realized at 30° C in shacked flasks.

7. ATPase activity measurement of Hco-Pgp-13 stimulated by actinomycin D

The ATPase activity in crude membranes was measured by the endpoint inorganic phosphate (Pi) release assay (Sarkadi et al., 1992). Briefly, membrane preparations (200 µg/ml) were incubated at 37°C in ATPase assay buffer (10 mM $MgCl_2$, 50 mM Hepes (pH 7.5), 1 mM dithiothreitol, 0,1 mM EGTA, 10 mM sodium azide, 10 mM ouabain) with increasing concentrations of actinomycin D, in the presence or absence of 100 µM sodium orthovanadate. DMSO was used as solvent of the tested drug, and its final concentration was 1% in the assay, as in the control reaction without drug. The reaction was initiated by the addition of 5 mM ATP and terminated with SDS (5% final concentration): After 5 min, 15 min and 30 min of incubation, aliquots (50 µl) were removed and added to SDS-containing wells. The amount of Pi released in each aliquot was measured in triplicate by a modification of the sensitive colorimetric reaction described. The SDS-containing samples (0.1 ml) were supplemented with 0.2 ml of reagent C prepared with 1 vol of reagent A containing 4% ammonium molybdate in 2,5M H_2SO_4 and 4 vol of reagent B containing 10% ascorbic acid freshly prepared. In this way,

the rates of ATPase activity were identified as linear initial rates up to 30 min after incubation with ATP. The vanadate-sensitive ATPase activity was expressed in nmol Pi/min/mg of total protein.

8. Design of specific antibodies against Hco-Pgp-13

The specificity and suitability of various antigens for antibody production against Hco-Pgp-13 were analyzed by GenScript (USA). Among the dozens of potential antigenic determinants identified, two peptides were chosen for their disordered structure and high accessibility in the predicted 3D conformation of the protein: GTADPQRSSETSKK (AA 12-25) and SGRSTLTQSKRSGS (AA 681-694). They were chemically synthesized and used by GenScript to immunize rabbits.

9. Polyacrylamide gel electrophoresis and Western-blot

All operations were carried out at 4°C. After induction of expression, yeast cells were harvested and broken during 30 min with glass beads in a breaking buffer (Tris/HCl 10 mM, pH 7.5) supplemented with protease inhibitors. The cell lysate was then centrifuged at 1000g for 15 min to remove unbroken cells and particulate matter. The supernatant was then centrifuged at 100,000g for 30 min. Resulting pellets were stored at -80°C in the breaking buffer. Membrane Protein contents were determined using the Bradford assay (Bio-Rad) using bovine serum albumin (BSA) as standard. Absorbance was read at 595nm with a spectrofluorometer (Infinite® 200 PRO, Tecan).

An aliquot of 10 µg of membranes was run on 10% sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE) in a Mini Protean® 3 Bio-Rad cell with the Laemmli method (Gallagher, 2006). The gel was then transferred to a nitrocellulose membrane that was blocked in PBS, 0.05% Tween, 5% skimmed milk powder, at 4°C overnight. The membrane was washed three times for 5min with PBS-Tween and incubated with the two primary antibodies designed against each epitope of Hco-Pgp-13 (GenScript, USA) at 1/1000 dilution at 4°C overnight. The next day, the membrane was washed similarly and incubated with a secondary CF^{TM}770 anti-rabbit antibody (Biotium, Canada) at RT for 1 hour. The membrane was again washed three times for 5 min with PBS-Tween and proteins were visualized using the Odyssey® CLx imaging system (LI-COR Biosciences, USA).

10. Immunohistochemistry on larvae and adult *H. contortus* sections

The anti- Hco-Pgp-13 antibodies designed and validated in Western-blot were used for immunohistochemistry detection of Hco-Pgp-13 protein in L3 and adult *H. contortus*. An anti-myosin antibody directed against *C. elegans* myosin heavy chain A was also used to localize muscles (DSHB Hybridoma product 5-6). Fresh worms were fixed in 4% PFA in PBS at 4°C for 16h. After three washes in PBS for 5 min, they were incubated in 30% sucrose in PBS at 4°C for 16h under gentle rocking. Whole worms were then individually placed in a square container and embedded in an optimal cutting temperature compound (OCT) (Thermo Fisher Scientific, USA). They were then quickly frozen to -80°C and stored. Cryosections were performed by slicing 20-30 µm thick transverse sections with a Thermo Shandon cryotome (Thermo Fisher Scientific, USA) and slices collected onto poly-L-lysine coated glass coverslips (Sigma, USA), kept at -80°C before further processing.

Sections were incubated in Antibody Diluent (AbD) (PBS, 0.2 % gelatin fish skin, 0.1% sodium azide, 0.1% (v/v) Triton X-100) at 4°C for 16h. This was followed by incubation with the two primary antibodies: anti-myosin and anti-Hco-Pgp-13 with 1/100 and 1/50 dilutions, respectively, in AbD at 4°C for 16h. Five washes for 5 min with AbD were performed to remove non-specific binding and the secondary fluorescent antibodies were incubated with a 1/500 dilution at 4°C for 16h. The secondary antibodies were Alexa Fluor 488 (Fab) fragment of goat anti-rabbit IgG and Alexa Fluor 635 goat anti -mouse IgG (Invitrogen, USA). Three washes were then performed for 5 min with AbD and three washes for 5 min with PBS. Sections were finally mounted on slides using mounting medium (Sigma, USA) and observed under a fluorescent microscope. The most representative cross-sections were then 3D analyzed under a confocal microscope (Leica SP8 DMI6000, Wetzlar, Germany) at excitation and emission wavelength of 488nm – 520 nm respectively for Alexa Fluor 488, and 635nm – 650nm respectively for AlexaFluor 635.

III. RESULTS

1. Amplification and sequencing of *Hco-pgp-13* cDNA

Hco-Pgp-13 was selected based on the strategic expression of its closest ortholog in *C. elegans* at the level of the amphids. RNA was extracted from adult *H. contortus* males and females, reverse transcribed into cDNA and amplified. Several primers (Suppl. Table S1) were designed along the predicted sequence of *Hco-pgp13* (ftp://ftp.sanger.ac.uk/pub/

pathogens/Haemonchus/contortus) (Laing et al., 2013). A first, partial 3' end fragment of 3488 bp was amplified using two primers encompassing the stop codon of *Hco-pgp13* predicted sequence ending at nucleotide 3954. The forward primer started at position 504 and the reverse primer at position 3992. This fragment was sequenced using eight forward and reverse primers, designed along the corresponding predicted sequence of the *H. contortus* cDNA, the start position of each of them being indicated in (Suppl. Table S1). Amplicons obtained with different primers showed overlapping sequences which allowed alignment of all the PCR products and sequencing of the whole fragment. To amplify the 5' end of *Hco-pgp-13* cDNA, we used four more forward and reverse primers. A forward primer named SL1 was designed based on the conserved sequence of the nematode spliced leader sequence 1 present before the ATG codon of most of genes (Blaxter and Liu, 1996; Blumenthal, 1995). A first PCR was performed on reverse translated RNA with SL1 and a reverse primer starting at position 965. This was followed by a nested PCR on this product with the pair of primers composed of a forward primer starting at position -34 and a reverse primer starting at 853. This allowed amplification of an 887 bp fragment. Once sequenced, the last 349 bp of the 3' end were overlapping and perfectly matching the first 349 bp of the 5' end of the previously amplified fragment. The combination of the two sequences was named *Hco-pgp-13* and corresponded to a 3954 bp coding cDNA from ATG to TGA stop codon (Suppl. Figure S1).

By sequencing the cDNA directly obtained from reverse transcription of *H. contortus* RNA, we were able to obtain a cDNA fragment which is representative of the transcript RNA present in the parasitic worms, and we could compare it with the gene sequence previously published (Laing et al., 2013). Interestingly, the alignment data revealed a deletion of 30 nucleotides in the amplified cDNA sequence strictly matching a 30-nucleotide repeat present on the *Hco-pgp-13* predicted sequence (Figure 1). 90 SNPs were found along the whole amplified cDNA relative to the one predicted by large scale sequencing, which represents 2.3% of the nucleotide sequence (Suppl. Figure S1).

2. Translation product amino-acid sequence and topology of amplified *Hco-pgp-13*

Translation of *Hco-pgp-13* cDNA from the first ATG to the next stop codon in the same reading frame yielded a 1317 amino-acid (AA) protein. Its expected molecular weight was 145.9 kDa, without taking account of possible post-translational modifications. The 30 nucleotide deletion in the cDNA sequence as compared to the predicted sequence translated into a 10 AA deletion in the protein sequence, thus correcting the sequence starting at Gly 804

until Phe 813. In addition, among the 90 SNPs found in the cDNA sequence between the predicted and amplified *Hco-pgp-13*, 8 translated into a different AA in the protein sequence, spread in the whole protein, as listed in Suppl. Table S4. This represents 0.6% of the protein sequence and no amino-acid change was located in the TM segments in a position or orientation that could modify the walls of the inner chamber, as compared to the predicted sequence. Eight N-glycosylation motifs were identified on Asn 191, 507, 518, 601, 704, 747, 984 and 1033 (Figure 2), while no putative O-glycosylation site and no peptide signal found.

The prediction of signature motifs, according to homology with proteins containing well-described domains, revealed the presence of 2 NBDs with the domain arrangement NBD1 (AA 418-654) and NBD2 (AA 1076-1312), matching the conserved arrangement of ABC B full transporters (Figure 2). Consistently, a Walker A (/P-loop) motif: GHSGCGKS (AA 453 – 460) and GPSGSGKS (AA 1111 – 1118) as well as a Walker B motif: VLLLDE (AA 577 – 582) and ILLLDE (AA 1235 – 1240) were found in each of the NBDs. An ABC transporter family signature motif (Hewitt and Lehner, 2003) was also found in each of these domains: LSGGQKQRI (AA 557 – 565) in NBD1 and LSGGQKQRI (AA 1215 – 1223) in NBD2, as well as an aromatic residue (A-loop) located 25 AA upstream of each Walker A motif: Y427 and Y1085. All these conserved motifs, involved in the binding and hydrolysis of ATP, are consistent with a functional primary active transporter able to export substrates.

In addition, 2 TMDs were also identified based on homology with proteins containing such domains: TMD1 (AA 95 – 383) and TMD2 (AA 756 – 1043), that signed the transmembrane-type topology of this transporter protein. Within TMDs, a more detailed prediction of TM helices was found, as shown in Figure 2B. It was mainly identical when searched with two different servers (Suppl. Table S5) based on homology with well-known transmembrane proteins, and led to the prediction of 12 TM helices, starting and ending on the cytoplasmic side.

3. Phylogenetic analysis of Hco-Pgp-13

A phylogenetic tree was constructed after exclusion of NBDs of all Hco-Pgps, Cel-Pgp-12, 13 and 14, Hsa-Pgp and Msm-ABCB1a and is represented in Figure 3. It shows that Hco-Pgp-13 has a similarly very high degree of homology to the three proteins: Cel-Pgp-12, Cel-Pgp-13 and Cel-Pgp-14. They all appear to be orthologs of the Hco-Pgp-13 sequence, and have no closer ortholog in *H. contortus* than Hco-Pgp-13.

Then, a multiple alignment was performed with all Hco-Pgps and Cel-Pgps and the mammalian Pgps: Hsa-Pgp and three Mmu-Pgps (ABCB1a, B1b and B4). Interestingly, N-terminal regions show a length between 30 and 50 AA in 24 proteins out of 32 aligned (until the first AA aligned to the start of TM1 in Cel-Pgp-1) whereas the N-terminal sequence of Hco-Pgp-13 as well as Hco-Pgp-10, Cel-Pgp-1, 10, 11, 12, 13 and 14 extends to 70-90 AA (Suppl. Figure S2). In the crystal structure of Cel-Pgp-1, the first one of a nematode Pgp to be elucidated and released in the Protein Data Bank (4F4C) (Jin et al., 2012), the N-terminal region of Cel-Pgp-1 was found to form a hairpin in the TMDs formed by two supplementary helices called TMa and TMb. Thus, other Pgps harboring a long N-terminal region might also form this hairpin structure, especially if AAs in this region are in majority aromatic or hydrophobic, which is the case for Hco-Pgp-13 (Suppl. Figure S3). Short N-terminal regions, on the other hand, are expected to show a cytosolic location.

The homologies of the TMDs of Cel-Pgp-1 with the TMDs of all Hco-Pgps, other Cel-Pgps, and mammalian Pgps: Hsa-Pgp, Mmu-ABCB1a, were calculated to identify the most similar putative binding domains to Cel-Pgp-1 (Table 1). As expected, Hco-Pgp-1 presented the highest homology for TMDs to Cel-Pgp-1, with 63% of identity and 79% of similarity of AAs, followed by the Pgp-9 of the two nematodes. Surprisingly, the TMDs of the 3 mammalian Pgps were closer to the TMDs of Cel-Pgp-1, around 35% identity and 57% similarity of AAs, than all other nematode Pgps investigated. Of particular interest, Hco-Pgp-13 showed the highest degree of homology to Cel-Pgp-1 TMDs among Pgps showing a long N-terminal helix, with 33% identity and 54% similarity of AAs.

4. Homology modelling of Hco-Pgp-13 on Cel-Pgp-1

This high degree of homology of sequence and of predicted structure led us to use the Cel-Pgp-1 4F4C structure as a template for building a 3D homology model of Hco-Pgp13, according to the pairwise alignment shown in Suppl. Figure S3, deduced from the multiple alignments, partly shown in Suppl. Figure S2. Among the 100 possibilities of 3D structures calculated by Modeller, the structure n°04 (Hco-Pgp-13_04) showed the lowest Discrete Optimized Protein Energy (DOPE) score (-151276) (molecular PDF score: 5791) and the model n° 52 (Hco-Pgp-13_52) had the lowest molecular PDF score (5292) (DOPE score: -150298). The DOPE score relies on the measure of the energy of the protein model, which is widely considered as the best parameter for choosing the most accurate model, but it has a tendency to pick the most compactly packed model, whereas the molPDF score calculates the sum of restraint violations. The Qualitative Model Energy Analysis (QMEAN) score was also

calculated as it takes into account four parameters: local geometry, long distance interactions, solvation energy and torsion, with the closer the score to 1 the better. A QMEAN score of 0.582 was found for Hco-Pgp-13_04 and 0.592 for Hco-Pgp-13_52, against 0.566 for Cel-Pgp-1, and other parameters evaluating the quality of the models were also very close between the two models (Suppl. Table S2), so that both models were considered to show the highest accuracy possible by homology modelling.

The two model proteins showed a mostly superimposable conformation consistent with an ABC transporter (Figure 4). The alignment matched well with a putative arrangement of 12 alpha helices across a hydrophobic membrane forming the TMDs, and NBDs composed of alpha helices and beta leaflets (Figure 4A). When looking closer at TMDs of both models, TM helices aligned well with Cel-Pgp-1 helices, with a slight shift of their backbones, and their amino-acid composition was consistent with what was predicted by Protter and TMHMM softwares (Suppl. Table S5). The extracellular loop 1 (ECL1) linking TM1 and TM2, as well as TM1 extension towards the extracellular part of the protein, were found to be shorter in Hco-Pgp-13 than in Cel-Pgp-1, as indicated by a black arrow in Figure 4B. This main difference between the two ABC transporters also appears in the pairwise alignment shown in Suppl. Fig. S3 and it is consistent with the absence of N-glycosylation motifs found on ECL1 of Hco-Pgp-13. The tertiary structures of the two Hco-Pgp-13 models are very similar, with a single variation in the middle of TM11 where a rupture of the alpha helix is found in Hco-Pgp-13_04, as compared to Hco-Pgp-13_52 and Cel-pgp-1. This is due to this helix containing 1amino-acid less in Hco-pgp-13 than in Cel-Pgp-1, leading to a torsion constraint (Figure 4C).

At a more precise level, the orientation of the side chains of amino acids composing the TMDs of Hco-Pgp-13 varied to different degrees between the two models, from being almost superimposed, to pointing towards opposite directions (Figure 4D). Residues aligned to mammalian Pgp "hotspots", i.e., residues found to interact with substrates experimentally (Suppl. Table S3) were found for the majority (50/62) of those recorded in the mammalian transporter, and point towards the inner pocket, formed by the 12 TM helices at the level of the drug binding site, being favorable for interaction with substrates.

5. *In silico* docking of actinomycin D on Hco-Pgp-13

Cel-Pgp-1 was proven to be a multispecific transporter of various human Pgp substrates both *in vitro* (Jin et al., 2012) and *in silico*, and in particular of AHs *in silico* (manuscripts under review, Part I, A and B). Actinomycin D (ACD) is a well-known mammalian Pgp substrate that was able to stimulate the ATPase activity of Cel-Pgp-1 to the highest extent *in vitro* and was

also found to interact with the Cel-Pgp-1 binding pocket *in silico*, with the highest affinity. To test the ability of Hco-Pgp-13 to bind drugs, we thus performed *in silico* docking of ACD on Hco-Pgp-13_04 and Hco-Pgp-13_52.

The clustering histograms of poses resulting from each calculation had a similar profile, with a lowest energy cluster containing the majority of poses, 21 and 42 respectively, at a very low binding energy, -16.0 and -14.4 kcal/mol respectively for Hco-Pgp-13_04 and Hco-Pgp-13_52 (Figure 5A and B and Table 2). The binding location of ACD was very close in the two models, occupying the inner pocket of Cel-Pgp-1 from its cytoplasmic opening to its inner core (Figure 5C). When looked at closely, one of the cycles of the molecule (cycle 1, Figure 5D) was almost superimposed between the two poses, deep in the core of the inner pocket, but with opposite orientations. The center of the molecule and the second cycle (cycle 2, Figure 5D) mainly occupied the same space while crossing each other, thus leading to locations of ACD on the two models, which overall were very similar, but oppositely oriented. These two possibilities of binding of ACD showed a similar number of interacting residues: 20 with Hco-Pgp-13_04 (including 11 hotspots) and 19 with Hco-Pgp-13_52 (including 9 hotspots). Twelve predicted interacting residues, among which 7 hotspots, were common between the two models and were either similarly oriented (e.g. L919) or pointing towards very different orientations (e.g. Q108) (Figure 5E, 5F and Table 3). Two H-bonds were formed in both cases, and one of them, formed with N1033, was common between the two models (Figure 5E, 5F and Table 3). Another H-bond with Q108 was only found with Hco-Pgp-13_04, whereas an H-bond with Y369 was only found with Hco-pgp-13_52 (Figure 5E, 5F and Table 3). All these three residues were hotspots, underlining their importance in substrate binding across species. The role of the small hairpin formed by TMa-b was also significant as three residues lining TMb: E21, K25 and L29 participated in the stability of ACD on this site in both models. These results indicate a very high affinity of ACD for Hco-Pgp-13 with a binding site not very dependent on the orientation of partially flexible AAs composing the TMDs; which increases the confidence in this model.

6. Expression of Hco-Pgp-13 in *Pichia pastoris* cells and stimulation of its ATPase activity by actinomycin D

Cells of *Pichia pastoris* were transfected with *Hco-pgp-13*. Pgp-13 could be readily detected by SDS-PAGE of crude membrane fractions followed by Coomassie blue staining, as well as Western-blot with an anti-His tag antibody (data not shown).

To confirm the possible interaction of Hco-Pgp-13 with ACD, we performed ATPase assays on membranes vesicles from *P. pastoris* and on Hco-Pgp-13 expressing *P. pastoris* cells. The vanadate-sensitive basal ATPase activity was around 110 nmol/min/mg in WT and transfected cells (Figure 6). It rose with increasing concentrations of ACD added to the *P. pastoris* membranes expressing Hco-Pgp-13, until approximately 180 nmol/min/mg protein under full stimulation with 10 µM of ACD. The concentration required for 50% stimulation was in the range 0.1 - 1 µM. This result indicated an interaction of ACD with Hco-Pgp-13, and combined with the predicted high affinity binding site of ACD within Hco-Pgp-13 binding site, it suggests that ACD could be a transport substrate for Hco-Pgp-13.

7. Immunolocalization of Hco-Pgp-13 protein in larvae and adult parasites

In order to identify the expression sites of Hco-Pgp-13 in different stages of *H. contortus*, immunofluorescence assays were performed on transverse cryosections of the larvae and adult parasites with a mixture of antibodies targeting two epitopes of Hco-Pgp-13.

To design these antibodies, two regions of interest were identified on the protein sequence of Hco-Pgp-13 as being the least conserved between different Pgps: the N-terminal region of the protein, before TMD1, and the linker region between NBD1 and TMD2. Two antigenic peptides identified to be specific for Hco-Pgp-13 were chosen for antibody production on each of these regions. The antigenic peptide n°1: GTADPQRSSETSKK (AA 12-25) was localized in the N-terminal region and the antigenic peptide n°2: SGRSTLTQSKRSGS (AA 681-694) in the linker region (Figure 2). Thus, prediction tools indicated an intra-cytoplasmic location for both of them, where the structure of Pgps is generally disordered and where antibodies are expected to easily reach their target peptide. The antigenic peptides were chemically synthesized and used independently to immunize two rabbits by GenScript (USA Inc.). The two purified antibodies were then checked for their specificity.

Membranes of *P. pastoris* were separately incubated with anti-epitope 1 (Figure 7A) or anti-epitope-2 (Figure 7B) antibody. A protein of about 130 kDa was detected in both cases, confirming the expected size of Hco-Pgp-13 protein, and showing that both of the predicted epitopes were successful antigenic peptides. As both antibodies appeared highly specific against Hco-Pgp-13 protein, a mixture of these two antibodies could then be used for immuno-fluorescence assays. An antibody directed against an epitope present on *C. elegans* myosin, and

very conserved in *H. contortus*, was used in parallel as a positive control and as an indication of muscle location.

. In larvae, the Hco-Pgp-13 expression was found in the seam cell between muscle quadrants and more generally in the hypodermis around muscles, including dense bodies between myosin filaments (grey and white arrows, Figure 8). Significant staining was also observed at the membrane of epithelial cells of the internal organs in development; corresponding to the gonad or intestine that cannot be readily distinguished at this stage (yellow arrow, Figure 8).

. In the adult *H. contortus*, Hco-Pgp-13 signal was also found at the level of seam cells between muscle quadrants and in the hypodermis surrounding them (grey and white arrows, Figure 9), as well as at the surface of the male gonad (Figure 9D, yellow arrow). In this stage, some staining also appeared in the procorpus of the pharynx at the level of pharyngeal nerve cords which contain cell bodies of epithelial cells, longitudinal extensions of neurons and gland cell processes. The staining observed in the metacorpus and terminal bulb also matches the extensions or cells bodies of pharyngeal neurons and gland cells (purple arrows, Figure 9B and 9C). Additional staining appeared in the head sections between the pharynx and seam cells (blue arrows, Figure 9B and 9C), and within the structure localized at the level of seam cells in other sections (Figure 9A). This corresponds to the location of neuronal structures and it likely matches with the amphidial neurons. Indeed, their dendrites run parallel to the pharynx from the nerve ring to the tip of the head, where seam cells end, and amphids are located at the corresponding place.

III. DISCUSSION

Nematode Pgps need to be individually studied in order to better understand their respective functions and possible implication in xenobiotics export and AH resistance. We focused our interest on the P-glycoprotein 13 of *H. contortus*, one of the most economically important parasites in small ruminants. By combining molecular biology, biochemistry and *in silico* approaches, we more accurately identified the cDNA of *Hco-pgp-13*, and characterized the localization and function of its translation product.

1. The *Hco-pgp-13* corrected cDNA sequence encodes a protein matching the topology of an ABC transporter

We amplified a full length 3954 bp *Hco-pgp-13* cDNA. When aligned with the *Hco-pgp-13* predicted cDNA sequence (Laing et al., 2013), the sequences were almost identical except for a 30 nucleotide repeat on the *Hco-pgp-13* published cDNA sequence. We found that this was due to a misalignment of contigs in the large scale sequencing. In addition, 90 SNPs were found in our sequence, compared with the published genome (Laing et al., 2013), consistent with polymorphism expected between different *H. contortus* populations.

The translated sequence, corrected for the 10 amino-acids duplication, showed a high degree of identity with the protein sequence previously published (Laing et al., 2013). Only eight isolated amino acids were found to vary between the Hco-Pgp-13 protein sequence predicted from the genome sequencing and that predicted from amplified cDNA (Suppl. Table S3). The 92 silent nucleotide changes between the predicted and amplified cDNA sequences suggest that the protein sequence is very conserved in order to maintain its function. Consistent with this, the 8 amino-acids that differed between the two translated sequences were either, not located in important domains for the function of the protein, or not in a critical location or orientation for the structure and supposed function of an ABC exporter, according to our homology models. In addition, for most of these residues the two possibilities given by the predicted and amplified sequence in the amino-acid were found in Pgps from other species at the corresponding location, according to our alignment (partly shown in Suppl. Fig S2). Thus, polymorphisms at these amino-acid locations are not expected to alter the function of the ABC exporter.

The full-length amino-acid sequence translated from the amplified cDNA revealed a protein of 1317 amino-acids. This was consistent with the Western-blot showing a band around 130 kDa, which matched the expected size of Hco-Pgp-13 without post-translational modification. Interestingly, no N-glycosylation motif was found on the first extra-cellular loop (ECL1) of Hco-Pgp-13. Hsa-Pgp shows 3 N-glycosylation motifs on this loop, and an absence of glycosylation of ECL1decreases its stability at the plasma membrane by making it undergo degradation in the proteasome, but does not alter the transport capacity of Pgp when present at the plasma membrane (Schinkel et al., 1993). Cel-Pgp-1, whose crystal structure was recently released, also showed an N-glysosylation on ECL1 (PDB code: 4F4C) (Jin et al., 2012). Hco-Pgp-13 is thus the first ABC B transporter to be described as showing no N-glycosylation motif on ECL1.

A transmembrane domain composed of 12 α-helices and two nucleotide-binding domains could be predicted by homology with other ABC transporters using topology prediction tools and homology modelling. At the level of the NBDs, all the motifs required for ATP binding and hydrolysis could be identified and indicate the possibility for these domains to provide the energy required for Hco-Pgp-13 to perform substrate translocation. Plus, 80% of residues aligned to hotspots, for drug binding identified in mammalian Pgps, were found located within TMDs and pointing towards the inner pocket of Hco-Pgp-13. However, only 19% of them were conserved in terms of nature in Hco-Pgp-13, which is less than the 26% that were found conserved in Cel-Pgp-1 (manuscript under review and writing, Part I, A and B). These data indicate that the general shape of the inner pocket might be critical for the function of these proteins, and thus very conserved across species, whereas the amino-acid composition of the pocket might vary to offer some specificities of binding for each transporter.

2. Hco-Pgp-13 can interact with actinomycin D

Since the modeled protein inherently presented the same conformation as the structural template, i.e. the so-called open inward-facing conformation, in the absence of bound nucleotide, it is well suited for substrate recognition and binding, as the first step of the eventual transmembrane transport catalytic cycle, and thus is amenable for *in silico* calculations of drug docking. To investigate the interaction profile of Hco-Pgp-13, a preliminary study of the *in silico* binding of ACD on the two homology models n°04 and n°52 of Hco-Pgp-13 was performed. In this context, a semi-flexible docking strategy for each model, i.e., flexible ligand and rigid protein, was similar to some extent to performing flexible docking, as the backbone of Hco-Pgp-13 only slightly varied between the two models. On the other hand, some residue side chains positions were found in very different orientations (Fig. 4D, 5E and F). ACD showed a predicted high affinity to Hco-Pgp-13, with a minimum binding energy below -16 kcal/mol for one model and -14.4 kcal/mol for the other, which is commonly considered to indicate a very high affinity of binding (Morris et al., 1998; Morris et al., 2009). This is very similar to what we found previously with Cel-Pgp-1, for which the two lowest energy clusters appeared at -17.0 and -14.9 kcal/mol (David et al., 2016). Interestingly, the location of the lowest energy cluster of ACD on Cel-Pgp-1 mostly overlapped that of the pose found at -14.4 kcal/mol on Hco-Pgp-13 (data not shown), with a similar shape elongating from the cytoplasmic entrance of the binding pocket until its core. On the two models of Hco-Pgp-13, about half of the putative interacting residues of the chosen poses of ACD were hotspot residues, revealing its binding within the large drug-binding domain of Pgps conserved between

nematodes and mammals. This proportion of hotspots among interacting residues was similar to what was found for ACD on Cel-Pgp-1. When examined more closely, 12 residues among those interacting with ACD, in each of the two models of Hco-Pgp-13, were found aligned with residues of Cel-Pgp-1 interacting with ACD1 or ACD2, among 19 and 20 in total, respectively. Seven hotspot residues were found in a total of 16 common interacting residues between the binding sites of ACD in Hco-Pgp-13 and Cel-Pgp-1. The 2 H-bonds formed by ACD with N1033 and either Q108 or Y369 of Hco-Pgp-13, depending on the model observed, were however very different from those formed with Q913 and R916 of Cel-Pgp-1. This is consistent with the medium level of conservation of residues in the TMDs of the respective proteins that leads to some specificity for the nature of interactions of compounds with each protein. Again, this correlates well with a conservation of the general binding domain of Pgps more in terms of shape and biochemical properties of various sub-domains than in terms of the nature of each residue composing it. In line with this, among the 6 hotspot residues that have been evidenced on mammalian Pgp to be specifically involved in ACD recognition (Tab. S3), only 1 (M995) was found to be involved in ACD binding on Hco-Pgp-13.

We also studied the stimulation of the ATPase activity of Hco-Pgp-13 by ACD, and we found a maximum efficacy at 10 µM of ACD. The activation of the vanadate sensitive ATPase activity at this concentration was around 70 nmol Pi/min/mg of protein. This represented about 1.6-fold activation as compared to the basal ATPase activity of membranes from Hco-Pgp-13-expressing transfected cells measured in the absence of drug. The half-activating concentration (EC50) was around 0.1 – 1 µM, significantly higher than the EC50 found for ACD on Cel-Pgp-1, previously determined of about 0.05 µM (Jin et al., 2012). This indicates a superior affinity for ACD to Cel-Pgp-1 than to Hco-Pgp-13, which is consistent with the slight difference of energy binding found *in silico* for ACD between Cel-Pgp-1 and Hco-Pgp-13. Otherwise, Jin et al. (2012) found a 2.5-fold stimulation of Cel-Pgp-1 ATPase activity by ACD as compared to its basal activity in the absence of drugs, and this suggests a faster rate of transport of ACD by Cel-Pgp-1 than by Hco-Pgp-13. The effect of ACD was also studied on the ATPase activity of chinese hamster (*Cricetulus griseus*) Pgp, showing an apparent affinity of about 1µM (Garrigues et al., 2002). This compound thus appears to be a substrate of several Pgps. However, the observation, in the absence of added exogenous drug, of a basal ATPase activity for Hco-Pgp-13 is an indication of the presence of an endogenous transport substrate in the membranes of this heterogenous protein expression system. This raises the question of the possible involvement of some lipids in the physiological function of Hco-Pgp-13 besides its

possible role of drug handling (which remains to be confirmed and extended for various other drugs: manuscript in preparation).

3. Hco-Pgp-13 sequence and localization are very close to those of Cel-Pgp-12, Cel-Pgp-13 and Cel-Pgp-14

The phylogenetic tree constructed with the protein sequence translated from the amplified cDNA of *Hco-pgp-13* confirmed the high degree of homology of Hco-Pgp-13 with Cel-Pgp-13 after removal of the much conserved NBDs. Hco-Pgp-13 also showed a high homology of protein sequence with Cel-Pgp-14 and Cel-Pgp-12, which have no closer ortholog in *H. contortus* than Hco-Pgp-13. Interestingly, (Zhao et al., 2004) had shown that the three genes encoding these proteins, together with the pseudogene *Cel-pgp-15*, form a cluster of tandemly duplicated genes on chromosome X of *C. elegans* genome. *Cel-pgp-12* is followed by *Cel-pgp-13* in one orientation, with *Cel-pgp-14* and *Cel-pgp-15* following with the opposite orientation. Genes arranged in the same orientation are supposed to be recently duplicated paralogues and can be under a common regulatory control (Zhao et al., 2004). They have very close sequences compared to other genes of the same family and are thus expected to show very similar functions. Zhao et al. reported no significant phenotype for single RNAi for each of the clustered genes of *C. elegans*, consistent with a redundant function of the encoded proteins. Plus, little selection pressure is generally thought to occur on duplicated paralogs, which is supposed to lead these genes to be not functional over evolution, as appears to be the case for Cel-Pgp-15 (http://www.wormbase.org/species/ c_elegans/pseudogene/F22E10.4#0--10).

However, in this cluster, all genes were found to be expressed in different tissues: Cel-Pgp-12 in the excretory cell at all stages, Cel-Pgp-13 in the posterior intestine and amphids of adults, Cel-Pgp-14 in the anterior region and first bulb of pharynx in the adult and larvae and Cel-Pgp-15 in the adult head and tail neurons and in the embryo (Zhao et al., 2004). This suggests a specific function for each of them after the expansion of the cluster. In our study, the expression of Hco-Pgp-13 was localized in the hypodermis surrounding muscles and lining the body wall of L3 and adult *H. contortus*, as well as in the seam cells within which run the excretory canals. Plus, larvae showed expression in gonad epithelial cells and staining was observed in pharyngeal structures of the adult, which could be the pharyngeal glands, neurons, and/or epithelial cells, and in head neurons, possibly the amphids. Thus, Hco-Pgp13 expression appears to match the localization of several of its closest orthologs in *C. elegans*. It can be hypothesized that the function of all the three Pgps of *C. elegans* (Pgp-15 being a pseudogene) are very close and that Hco-Pgp-13 alone has a function corresponding to all of them due to its

wider tissue expression. Zhao et al. (2004) suggested that the presence of similar paralogs functionally very close, rather than one protein, could be useful for a more effective protection against xenobiotics by spatially and temporally differential expression, and possibly higher overall expression level. This might explain why the duplication of a common ancestor gene to Hco-Pgp-13 could have happened in *C. elegans* to optimize the efficiency of their close function. These hypotheses are consistent with the presence of a single ortholog in *H. contortus* corresponding to the cluster in *C. elegans*.

Inactivation of each Cel-Pgp individually was shown to lead to an increased sensitivity to MLs, especially for several Pgps including *Cel-pgp-12* and *Cel-pgp-13* (Ardelli and Prichard, 2013; Bygarski et al., 2014; Janssen et al., 2013b). Of interest, the global location of Hco-Pgp-13 and its closest orthologs in *C. elegans* are mainly found in the digestive apparatus that can be a way of entry of xenobiotics, in the excretory cell that can be involved in their detoxication and the amphids which may play a role in the entry of ML; neurons matching the amphids were stained with anti-Hco-Pgp-13 antibodies. Moreover, the pharyngeal structures stained might be the neurons innervating them, and the major effect of MLs on worms is starvation by paralysis of the pharyngeal muscles, so that Pgps expressed in pharyngeal neurons might be critical to protecting the worm from this effect. The expression of Hco-Pgp-13 in the epithelial cells of the uterus resembled the finding of Hco-Pgp-9.1 expression in this tissue (Godoy et al., 2016), and an involvement of these proteins in xenobiotics detoxication could be a way of avoiding toxic compounds reaching the embryos, as was proven to occur in mammalian placental tissue (Kolwankar et al., 2005; Nakamura et al., 1997). Thus, the function of such proteins can be expected to be not only the transport of endogenous substrates but also of exogenous compounds, possibly toxic for the worms.

Cel-Pgp-1 was shown to be a multispecific substrate transporter (Jin et al., 2012), and was predicted to have high affinities for MLs (Manuscripts under review and writing, Part I, A and B). The high homology between the sequence and 3D structure of this Pgp and Hco-Ppg-13 suggests that these two proteins might have related functions. In particular, the hairpin TMa-b, never described in mammalian Pgp, appears to close the binding site of Cel-Pgp-1 and Hco-Pgp-13 on one side, thus possibly playing a key role in their interaction with substrates, and making the putative structure of Hco-Pgp-13 closer to that of Cel-Pgp-1, than to other Pgps studied so far. Interestingly, the long N-terminal sequence that forms this hairpin was also found in the closest orthologs of Hco-Pgp-13 in *C. elegans*: Cel-Pgp-12, 13 and 14, which correlates well with a high homology between these proteins.

In conclusion, this work allowed the correction of a 3954 bp coding cDNA of *Hco-pgp-13* and the prediction of its translation into a protein of 1317 AA with very close homology to the multispecific active efflux transporter Cel-Pgp-1 both in terms of sequence and 3D conformation. *In silico* docking experiments showed a putative high affinity binding site for ACD on Hco-Pgp-13, and *in vitro* experiments indicated the possibility for this molecule to stimulate Hco-Pgp-13 activity. Antibodies detecting Hco-Pgp-13 were successfully designed and Hco-Pgp-13 expression was found in the worm in the digestive, excretory and neuronal systems, and matching the localizations of its three closest orthologs in *C. elegans*: Cel-Pgp-12, 13 and 14. Altogether, these protein characteristics, similar to those of other multidrug transporters, and the wide location of its expression, indicate a possible important function for Hco-Pgp-13 in the transport of many substrates. This could be critical for the export of toxic compounds. Further experiments studying the interaction of AHs with Hco-Pgp-13 could indicate a possible role in AH resistance.

Acknowledgements

We thank Kathy Keller and Hua Che for valuable support and advice for experiments performed at McGill. We also acknowledge Shaima Hashem for technical help with *in silico* calculations. We thank the computing facilities of the CEA-DSV at Saclay and the GenoToul bioinformatics hardware infrastructure that were used for computing. This work was supported by the Natural Sciences and Engineering Research Council of Canada, the FRQNT Centre for Host-Parasite Interactions, Quebec, and EMIDA ERA-NET project CARES n° 11-EMID-003-02.

References

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.

Ardelli, B.F., Prichard, R.K., 2013, Inhibition of P-glycoprotein enhances sensitivity of Caenorhabditis elegans to ivermectin. Vet Parasitol 191, 264-275.

Bartley, D.J., McAllister, H., Bartley, Y., Dupuy, J., Menez, C., Alvinerie, M., Jackson, F., Lespine, A., 2009, P-glycoprotein interfering agents potentiate ivermectin susceptibility in ivermectin sensitive and resistant isolates of Teladorsagia circumcincta and Haemonchus contortus. Parasitology 136, 1081-1088.

Benkert, P., Kunzli, M., Schwede, T., 2009, QMEAN server for protein model quality estimation. Nucleic Acids Res 37, W510-514.

Benkert, P., Tosatto, S.C., Schomburg, D., 2008, QMEAN: A comprehensive scoring function for model quality assessment. Proteins 71, 261-277.

Bessadok, A., Garcia, E., Jacquet, H., Martin, S., Garrigues, A., Loiseau, N., Andre, F., Orlowski, S., Vivaudou, M., 2011, Recognition of sulfonylurea receptor (ABCC8/9) ligands by the multidrug resistance transporter P-glycoprotein (ABCB1): functional similarities based on common structural features between two multispecific ABC proteins. J Biol Chem 286, 3552-3569.

Blaxter, M., Liu, L., 1996, Nematode spliced leaders--ubiquity, evolution and utility. International journal for parasitology 26, 1025-1033.

Blumenthal, T., 1995, Trans-splicing and polycistronic transcription in Caenorhabditis elegans. Trends Genet 11, 132-136.

Broeks, A., Gerrard, B., Allikmets, R., Dean, M., Plasterk, R.H., 1996, Homologues of the human multidrug resistance genes MRP and MDR contribute to heavy metal resistance in the soil nematode Caenorhabditis elegans. EMBO J 15, 6132-6143.

Broeks, A., Janssen, H.W., Calafat, J., Plasterk, R.H., 1995, A P-glycoprotein protects Caenorhabditis elegans against natural toxins. EMBO J 14, 1858-1866.

Bygarski, E.E., Prichard, R.K., Ardelli, B.F., 2014, Resistance to the macrocyclic lactone moxidectin is mediated in part by membrane transporter P-glycoproteins: Implications for control of drug resistant parasitic nematodes. Int J Parasitol Drugs Drug Resist 4, 143-151.

Corpet, F., 1988, Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res 16, 10881-10890.

Dent, J.A., Smith, M.M., Vassilatis, D.K., Avery, L., 2000, The genetics of ivermectin resistance in Caenorhabditis elegans. Proceedings of the National Academy of Sciences of the United States of America 97, 2674-2679.

Eckford, P.D., Sharom, F.J., 2009, ABC efflux pump-based resistance to chemotherapy drugs. Chem Rev 109, 2989-3011.

Edgar, R.C., 2004, MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32, 1792-1797.

Freeman, A.S., Nghiem, C., Li, J., Ashton, F.T., Guerrero, J., Shoop, W.L., Schad, G.A., 2003, Amphidial structure of ivermectin-resistant and susceptible laboratory and field strains of Haemonchus contortus. Vet Parasitol 110, 217-226.

Gallagher, S.R., 2006, One-dimensional SDS gel electrophoresis of proteins. Curr Protoc Immunol Chapter 8, Unit 8 4.

Garrigues, A., Loiseau, N., Delaforge, M., Ferte, J., Garrigos, M., Andre, F., Orlowski, S., 2002, Characterization of two pharmacophores on the multidrug transporter P-glycoprotein. Mol Pharmacol 62, 1288-1298.

Godoy, P., Che, H., Beech, R.N., Prichard, R.K., 2015a, Characterization of Haemonchus contortus P-glycoprotein-16 and its interaction with the macrocyclic lactone anthelmintics. Mol Biochem Parasitol 204, 11-15.

Godoy, P., Che, H., Beech, R.N., Prichard, R.K., 2016, Characterisation of P-glycoprotein-9.1 in Haemonchus contortus. Parasit Vectors 9, 52.

Godoy, P., Lian, J., Beech, R.N., Prichard, R.K., 2015b, Haemonchus contortus P-glycoprotein-2: in situ localisation and characterisation of macrocyclic lactone transport. International journal for parasitology 45, 85-93.

Hewitt, E.W., Lehner, P.J., 2003, The ABC-transporter signature motif is required for peptide translocation but not peptide binding by TAP. Eur J Immunol 33, 422-427.

Higgins, C.F., 1992, ABC transporters: from microorganisms to man. Annu Rev Cell Biol 8, 67-113.

James, C.E., Davey, M.W., 2009, Increased expression of ABC transport proteins is associated with ivermectin resistance in the model nematode Caenorhabditis elegans. International journal for parasitology 39, 213-220.

Janssen, I.J., Krucken, J., Demeler, J., Basiaga, M., Kornas, S., von Samson-Himmelstjerna, G., 2013a, Genetic variants and increased expression of Parascaris equorum P-

glycoprotein-11 in populations with decreased ivermectin susceptibility. PLoS One 8, e61635.

Janssen, I.J., Krucken, J., Demeler, J., von Samson-Himmelstjerna, G., 2013b, Caenorhabditis elegans: modest increase of susceptibility to ivermectin in individual P-glycoprotein loss-of-function strains. Exp Parasitol 134, 171-177.

Janssen, I.J., Krucken, J., Demeler, J., von Samson-Himmelstjerna, G., 2015, Transgenically expressed Parascaris P-glycoprotein-11 can modulate ivermectin susceptibility in Caenorhabditis elegans. Int J Parasitol Drugs Drug Resist 5, 44-47.

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

Jones, P.M., George, A.M., 2005, Multidrug resistance in parasites: ABC transporters, P-glycoproteins and molecular modelling. International journal for parasitology 35, 555-566.

Kaschny, M., Demeler, J., Janssen, I.J., Kuzmina, T.A., Besognet, B., Kanellos, T., Kerboeuf, D., von Samson-Himmelstjerna, G., Krucken, J., 2015, Macrocyclic lactones differ in interaction with recombinant P-glycoprotein 9 of the parasitic nematode Cylicocylus elongatus and ketoconazole in a yeast growth assay. PLoS Pathog 11, e1004781.

Koenderink, J.B., Kavishe, R.A., Rijpma, S.R., Russel, F.G., 2010, The ABCs of multidrug resistance in malaria. Trends Parasitol 26, 440-446.

Kolwankar, D., Glover, D.D., Ware, J.A., Tracy, T.S., 2005, Expression and function of ABCB1 and ABCG2 in human placental tissue. Drug Metab Dispos 33, 524-529.

Kurz, C.L., Shapira, M., Chen, K., Baillie, D.L., Tan, M.W., 2007, Caenorhabditis elegans pgp-5 is involved in resistance to bacterial infection and heavy metal and its regulation requires TIR-1 and a p38 map kinase cascade. Biochem Biophys Res Commun 363, 438-443.

Lage, H., 2003, ABC-transporters: implications on drug resistance from microorganisms to human cancers. Int J Antimicrob Agents 22, 188-199.

Laing, R., Hunt, M., Protasio, A.V., Saunders, G., Mungall, K., Laing, S., Jackson, F., Quail, M., Beech, R., Berriman, M., Gilleard, J.S., 2011, Annotation of two large contiguous regions from the Haemonchus contortus genome using RNA-seq and comparative analysis with Caenorhabditis elegans. PLoS One 6, e23216.

Laing, R., Kikuchi, T., Martinelli, A., Tsai, I.J., Beech, R.N., Redman, E., Holroyd, N., Bartley, D.J., Beasley, H., Britton, C., Curran, D., Devaney, E., Gilabert, A., Hunt, M., Jackson, F., Johnston, S.L., Kryukov, I., Li, K., Morrison, A.A., Reid, A.J., Sargison, N.,

Saunders, G.I., Wasmuth, J.D., Wolstenholme, A., Berriman, M., Gilleard, J.S., Cotton, J.A., 2013, The genome and transcriptome of Haemonchus contortus, a key model parasite for drug and vaccine discovery. Genome Biol 14, R88.

Leonard, G.D., Fojo, T., Bates, S.E., 2003, The role of ABC transporters in clinical practice. Oncologist 8, 411-424.

Lespine, A., Martin, S., Dupuy, J., Roulet, A., Pineau, T., Orlowski, S., Alvinerie, M., 2007, Interaction of macrocyclic lactones with P-glycoprotein: structure-affinity relationship. European journal of pharmaceutical sciences : official journal of the European Federation for Pharmaceutical Sciences 30, 84-94.

Lespine, A., Menez, C., Bourguinat, C., Prichard, R.K., 2012, P-glycoproteins and other multidrug resistance transporters in the pharmacology of anthelmintics: Prospects for reversing transport-dependent anthelmintic resistance. Int J Parasitol Drugs Drug Resist 2, 58-75.

Li, J., Jaimes, K.F., Aller, S.G., 2013, Refined structures of mouse P-glycoprotein. Protein Sci 23, 34-46.

Loo, T.W., Bartlett, M.C., Clarke, D.M., 2006a, Transmembrane segment 1 of human P-glycoprotein contributes to the drug-binding pocket. Biochem J 396, 537-545.

Loo, T.W., Bartlett, M.C., Clarke, D.M., 2006b, Transmembrane segment 7 of human P-glycoprotein forms part of the drug-binding pocket. Biochem J 399, 351-359.

Loo, T.W., Clarke, D.M., 2001, Defining the drug-binding site in the human multidrug resistance P-glycoprotein using a methanethiosulfonate analog of verapamil, MTS-verapamil. J Biol Chem 276, 14972-14979.

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

Mani, T., Bourguinat, C., Keller, K., Ashraf, S., Blagburn, B., Prichard, R.K., 2016, Interaction of macrocyclic lactones with a Dirofilaria immitis P-glycoprotein. International journal for parasitology.

Menez, C., Alberich, M., Kansoh, D., Blanchard, A., Lespine, A., 2016, Acquired tolerance to ivermectin and moxidectin after drug selection pressure in the nematode Caenorhabditis elegans. Antimicrob Agents Chemother.

Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., Olson, A.J., 1998, Automated docking using lamarckian genetic algorithm and an empirical binding free energy function. J Comput Chem 19, 1639-1662.

Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., Olson, A.J., 2009, AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J Comput Chem 30, 2785-2791.

Nakamura, Y., Ikeda, S., Furukawa, T., Sumizawa, T., Tani, A., Akiyama, S., Nagata, Y., 1997, Function of P-glycoprotein expressed in placenta and mole. Biochem Biophys Res Commun 235, 849-853.

Nunes, F., Wolf, M., Hartmann, J., Paul, R.J., 2005, The ABC transporter PGP-2 from Caenorhabditis elegans is expressed in the sensory neuron pair AWA and contributes to lysosome formation and lipid storage within the intestine. Biochem Biophys Res Commun 338, 862-871.

Omasits, U., Ahrens, C.H., Muller, S., Wollscheid, B., 2014, Protter: interactive protein feature visualization and integration with experimental proteomic data. Bioinformatics 30, 884-886.

Ranjan, S., Wang, G.T., Hirschlein, C., Simkins, K.L., 2002, Selection for resistance to macrocyclic lactones by Haemonchus contortus in sheep. Vet Parasitol 103, 109-117.

Roulet, A., Puel, O., Gesta, S., Lepage, J.F., Drag, M., Soll, M., Alvinerie, M., Pineau, T., 2003, MDR1-deficient genotype in Collie dogs hypersensitive to the P-glycoprotein substrate ivermectin. Eur J Pharmacol 460, 85-91.

Sali, A., Blundell, T.L., 1993, Comparative protein modelling by satisfaction of spatial restraints. Journal of molecular biology 234, 779-815.

Sarramegna, V., Muller, I., Mousseau, G., Froment, C., Monsarrat, B., Milon, A., Talmont, F., 2005, Solubilization, purification, and mass spectrometry analysis of the human mu-opioid receptor expressed in Pichia pastoris. Protein Expr Purif 43, 85-93.

Schinkel, A.H., Kemp, S., Dolle, M., Rudenko, G., Wagenaar, E., 1993, N-glycosylation and deletion mutants of the human MDR1 P-glycoprotein. J Biol Chem 268, 7474-7481.

Schinkel, A.H., Smit, J.J., van Tellingen, O., Beijnen, J.H., Wagenaar, E., van Deemter, L., Mol, C.A., van der Valk, M.A., Robanus-Maandag, E.C., te Riele, H.P., et al., 1994, Disruption of the mouse mdr1a P-glycoprotein gene leads to a deficiency in the blood-brain barrier and to increased sensitivity to drugs. Cell 77, 491-502.

Schroeder, L.K., Kremer, S., Kramer, M.J., Currie, E., Kwan, E., Watts, J.L., Lawrenson, A.L., Hermann, G.J., 2007, Function of the Caenorhabditis elegans ABC transporter PGP-2 in the biogenesis of a lysosome-related fat storage organelle. Mol Biol Cell 18, 995-1008.

Shilling, R.A., Venter, H., Velamakanni, S., Bapna, A., Woebking, B., Shahi, S., van Veen, H.W., 2006, New light on multidrug binding by an ATP-binding-cassette transporter. Trends Pharmacol Sci 27, 195-203.

Sippl, M.J., 1993, Recognition of errors in three-dimensional structures of proteins. Proteins 17, 355-362.

Urdaneta-Marquez, L., Bae, S.H., Janukavicius, P., Beech, R., Dent, J., Prichard, R., 2014, A dyf-7 haplotype causes sensory neuron defects and is associated with macrocyclic lactone resistance worldwide in the nematode parasite Haemonchus contortus. International journal for parasitology 44, 1063-1071.

Webb, B., Sali, A., 2014, Comparative Protein Structure Modeling Using MODELLER. Curr Protoc Bioinformatics 47, 5 6 1-32.

Wiederstein, M., Sippl, M.J., 2007, ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35, W407-410.

Willard, L., Ranjan, A., Zhang, H., Monzavi, H., Boyko, R.F., Sykes, B.D., Wishart, D.S., 2003, VADAR: a web server for quantitative evaluation of protein structure quality. Nucleic Acids Res 31, 3316-3319.

Zhao, Z., Sheps, J.A., Ling, V., Fang, L.L., Baillie, D.L., 2004, Expression analysis of ABC transporters reveals differential functions of tandemly duplicated genes in Caenorhabditis elegans. Journal of molecular biology 344, 409-417.

Figures

```
                1       10        20        30        40        50
                |       |         |         |         |         |
Amplified 2301 GGAACTTCCAGCCTTGTCACTTGTTTTTGCCTATGTCTTTGAAGCTTTTC
Predicted 2301 GGAACTTCCAGCCTTGTCTCTTGTTTTTGCCTATGTCTTTGAAGCTTTTC

Amplified 2351 AAATGGTCCCATGGGGAGCCGATATGATGCACAGATTATGCATGGCTGTC
Predicted 2351 AAATGGTCCCATGGGGAGCCGATATGATGCACAGATTATGTATGGCTGTC

Amplified 2401 ATCATTTTCGGTTCCATTGGTGTTGGTGTCGTCATCTTC----------
Predicted 2401 ATCATTTTCGGTTCCATTGGTGTTGGTGTCGTCATCTTCGGTTCCATTGG

Amplified 2440 ------------------CAGCTCCTTATCAGTGTGTTCTTCGCAATTG
Predicted 2451 TGTTGGTGTCGTCATCTTCCAGCTCCTTAGCAGTGTGTTCTTCGCAATCG

Amplified 2501 TGTCATACAATTTGGCAATGCGATTTCGAGTGGAATCCTTCAAAAATCTA
Predicted 2531 TGTCAGAGAATTTGGCGATGCGATTTCGAGTGGAATCCTTCAAAAATCTA

Amplified 2551 CTCTACCAGGATGCTTCGTATTTTGACAATCCTGCCCATACACCTGGCAA
Predicted 2581 CTCTACCAGGATGCTTCGTATTTCGACAATCCTGCCCATACACCTGGCAA

Amplified 2601 GCTCATAACTCGCTTGGCTAGTGACGCACCAAATATCAAAGCAGTTGTCG
Predicted 2631 GCTTATAACTCGTTTGGCTAGTGATGCACCGAATATTAAAGCAGTTGTCG

Amplified 2651 ATGGTCGTGCACTTCAAGTTATCTACGCAATGACGGCCGTAATCGCATGT
Predicted 2681 ATGGTCGTGCACTTCAAGTTATCTACGCAATGACGGCTGTAATCGCATGT
```

Figure 1. Alignment of *Hco-pgp-13* amplified cDNA sequence from AA 2301 to 2651 with predicted sequence of Laing et al. (2013) available at (ftp://ftp.sanger.ac.uk/pub/pathogens/Haemonchus/contortus) from AA 2301 to 2681. Common sequences are represented as red characters. The 30 nucleotide repeat is highlighted in grey. SNPs are indicated as blue characters.

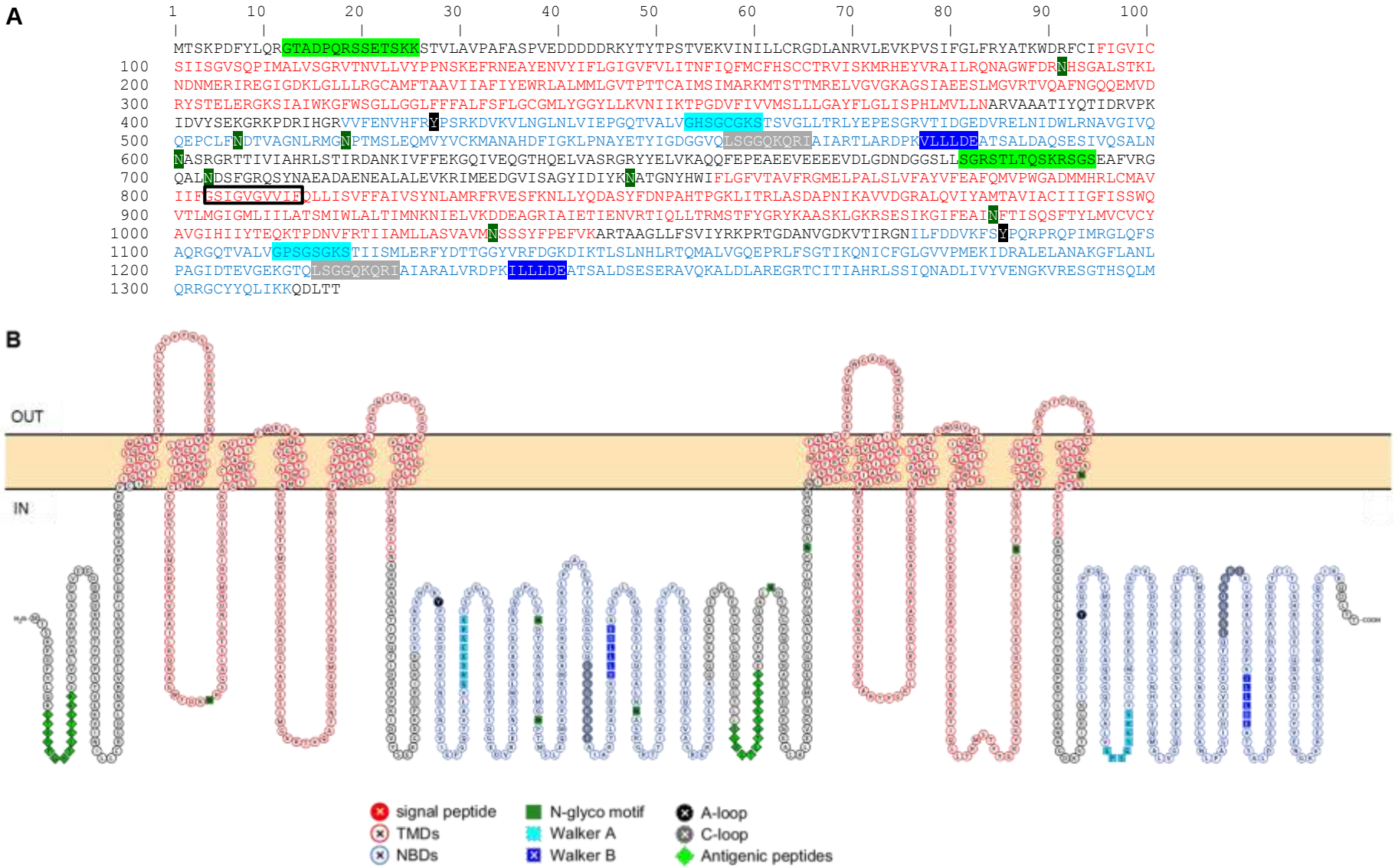

A
1 10 20 30 40 50 60 70 80 90 100
MTSKPDFYLQRGTADPQRSSETSKKSTVLAVPAFASPVEDDDDDRKYTYTPSTVEKVINILLCRGDLANRVLEVKPVSIFGLFRYATKWDRFCIFIGVIC
100 SIISGVSQPIMALVSGRVTNVLLVYPPNSKEFRNEAYENVYIFLGIGVFVLITNFIQFMCFHSCCTRVISKMRHEYVRAILRQNAGWFDRNHSGALSTKL
200 NDNMERIREGIGDKLGLLLRGCAMFTAAVIIAFIYEWRLALMMLGVTPTTCAIMSIMARKMTSTTMRELVGVGKAGSIAEESLMGVRTVQAFNGQQEMVD
300 RYSTELERGKSIAIWKGFWSGLLGGLFFFALFSFLGCGMLYGGYLLKVNIIKTPGDVFIVVMSLLLGAYFLGLISPHLMVLLNARVAAATIYQTIDRVPK
400 IDVYSEKGRKPDRIHGRVVFENVHFRYPSRKDVKVLNGLNLVIEPGQTVALVGHSGCGKSTSVGLLTRLYEPESGRVTIDGEDVRELNIDWLRNAVGIVQ
500 QEPCLFNDTVAGNLRMGNPTMSLEQMVYVCKMANAHDFIGKLPNAYETYIGDGGVQLSGGQKQRIAIARTLARDPKVLLLDEATSALDAQSESIVQSALN
600 NASRGRTTIVIAHRLSTIRDANKIVFFEKGQIVEQGTHQELVASRGRYYELVKAQQFEPEAEEVEEEEVDLGDNDGGSLLSGRSTLTQSKRSGSEAFVRG
700 QALNDSFGRQSYNAEADAENEALALEVKRIMEEDGVISAGYIDIYKNATGNYHWIFLGFVTAVFRGMELPALSLVFAYVFEAFQMVPWGADMMHRLCMAV
800 IIFGSIGVGVVIFQLLISVFFAIVSYNLAMRFRVESFKNLLYQDASYFDNPAHTPGKLITRLASDAPNIKAVVDGRALQVIYAMTAVIACIIIGFISSWQ
900 VTLMGIGMLIILATSMIWLALTIMNKNIELVKDDEAGRIAIETIENVRTIQLLTRMSTFYGRYKAASKLGKRSESIKGIFEAINFTISQSFTYLMVCVCY
1000 AVGIHIIYTEQKTPDNVFRTIIAMLLASVAVMNSSSYFPEFVKARTAAGLLFSVIYRKPRTGDANVGDKVTIRGNILFDDVKFSYPQRPRQPIMRGLQFS
1100 AQRGQTVALVGPSGSGKSTIISMLERFYDTTGGYVRFDGKDIKTLSLNHLRTQMALVGQEPRLFSGTIKQNICFGLGVVPMEKIDRALELANAKGFLANL
1200 PAGIDTEVGEKGTQLSGGQKQRIAIARALVRDPKILLLDEATSALDSESERAVQKALDLAREGRTCITIAHRLSSIQNADLIVYVENGKVRESGTHSQLM
1300 QRRGCYYQLIKKQDLTT
B
OUT
IN
signal peptide
TMDs
NBDs
N-glyco motif
Walker A
Walker B
A-loop
C-loop
Antigenic peptides

Figure 2. Predicted protein amino acid sequence (ExPASy – Translate) **(A), and topology** (ExPASy – ScanProsite, Protter) **(B).** The sequence of transmembrane domains (TMDs) and nucleotide binding domains (NBDs) are represented in red and blue characters, respectively (A), or as cellular topology as red and blue circles, respectively (B). Within each NBD, the Walker A domain sequence is highlighted in light blue and the Walker B motif in dark blue (A and B), in square characters (B). The tyrosine residue forming the A-loop of each domain is highlighted in black (A and B) and the C-loop/ABC transporter signature motif is highlighted in grey (A and B); both are circled in (B). Putative N-glycosylation motifs are highlighted in dark green (A and B) and squares (B). Protein regions chosen as antigenic determinant n°1 (N-term region) and n°2 (between NBD1 and TMD2) are both highlighted in light green (A and B) and represented inside diamonds (B). The sequence within the black box (A) indicates the duplicated 10 AAs in the predicted sequence published by Laing et al. (2012).

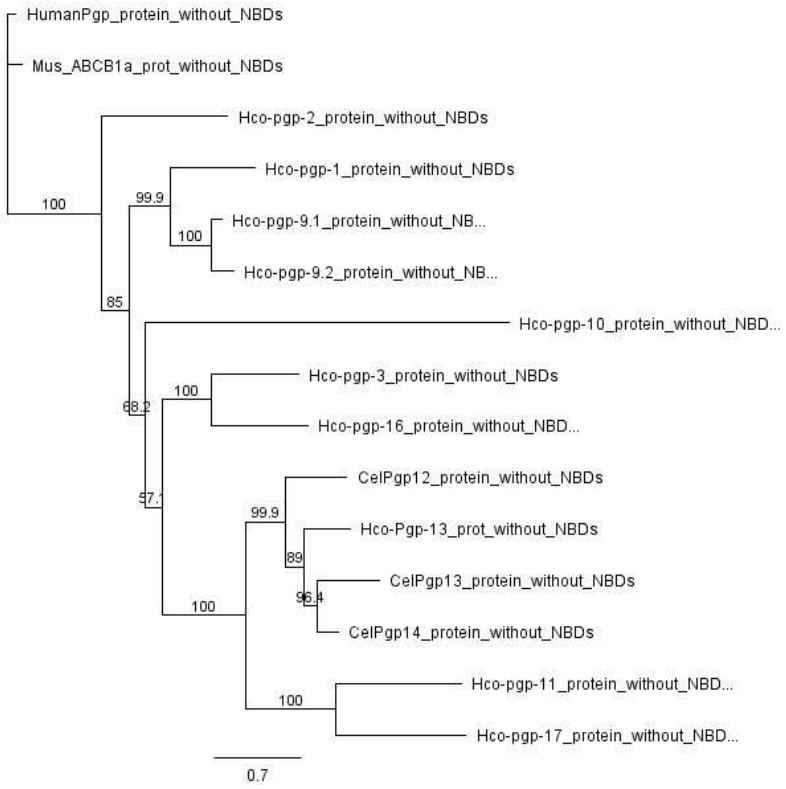


Figure 3. Comparison of Hco-Pgp-13 protein sequence with Hco-Pgps and Cel-Pgps. The Neighbor-Joining phylogenetic tree was built using Geneious software after NBD sequence removal, Blosum 62 alignment, and selection of Hsa-Pgp and Mmu-ABCB1a as outgroups.

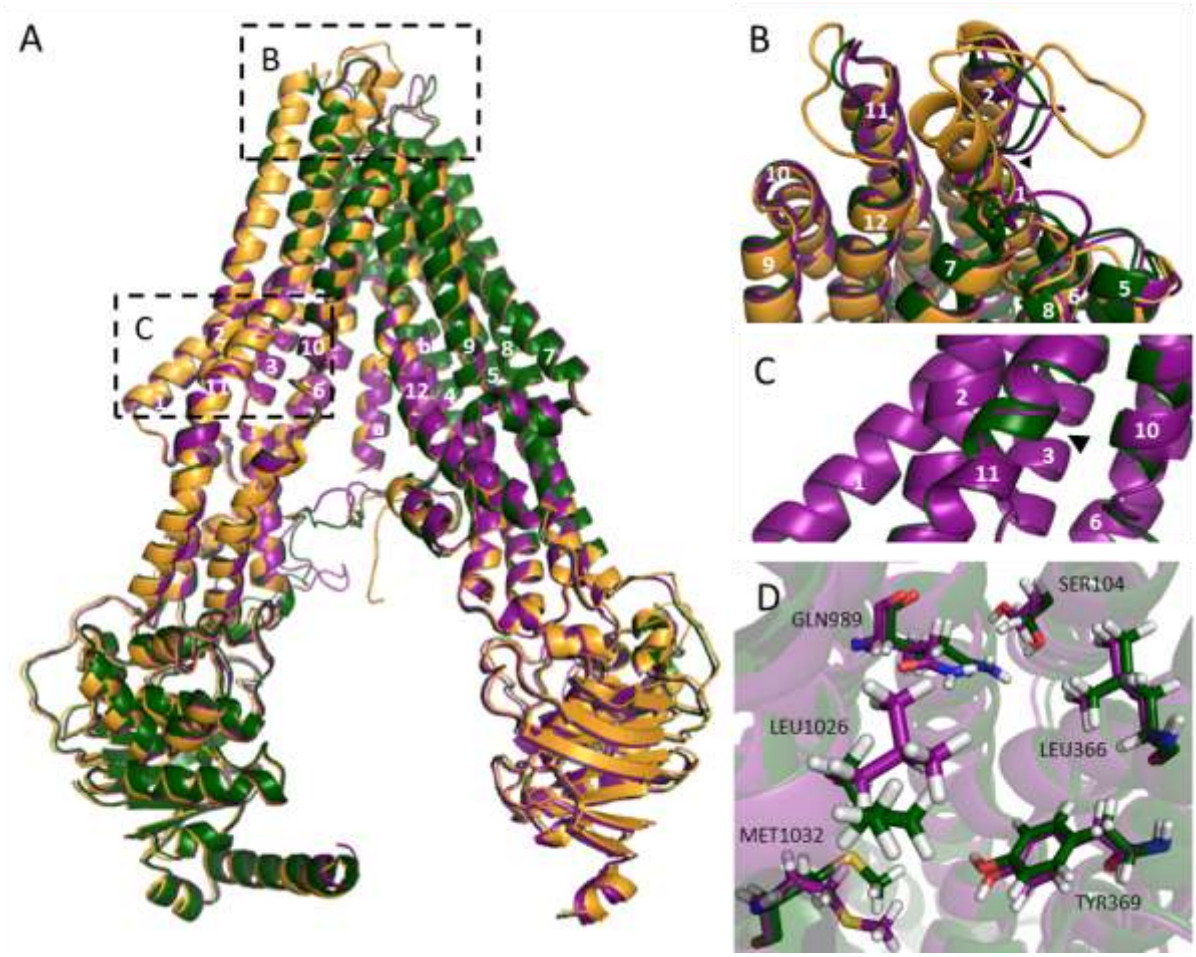


Figure 4. Alignment of Hco-Pgp-13 homology modeled structures with Cel-Pgp-1 crystal. A. Whole view of the aligned 3D structures of Cel-Pgp-1, Hco-Pgp-13-04 and Hco-Pgp-13-52. B. Zoom on the end of TM1 on the extracellular side of the membrane, indicated by a black arrow, and ECL1 linking TM1 and TM2. C. Zoom on the middle of TM11 of Hco-Pgp-13-04 and Hco-Pgp-13-52. D. Zoom on the side chains of several amino acids composing the TMDs of Hco-Pgp-13-04 and Hco-Pgp-13-52 and identified as hotspots. Cel-Pgp-1 is represented as gold ribbon, Hco-Pgp-13-04 is represented as purple ribbon and Hco-Pgp-13-52 is represented as green ribbon. TM identities are indicated as white numbers. AA side chains are represented as sticks colored by element nature (backbone color for C, red for O, blue for N, yellow for S, white for H) and labeled accordingly. All images were generated using PyMol.

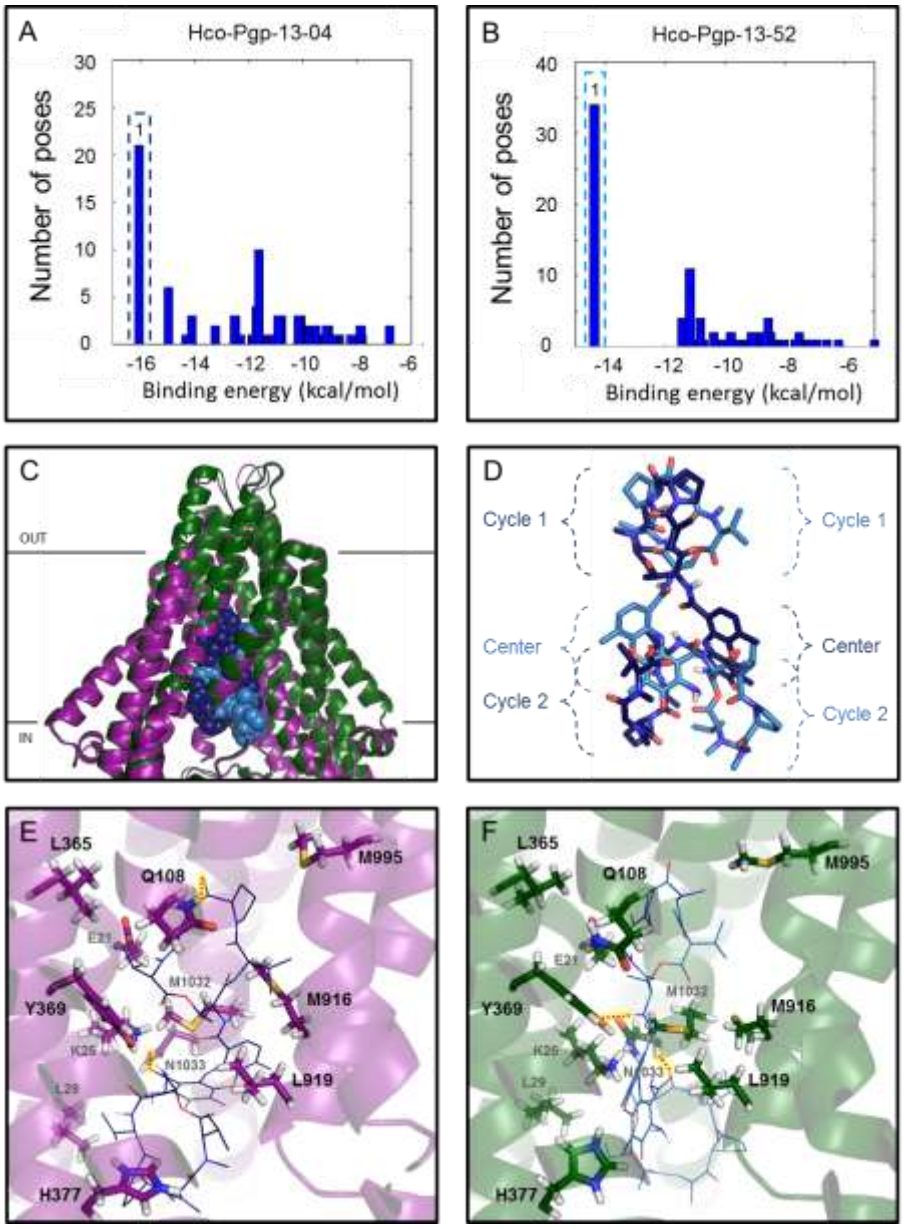


Figure 5. Actinomycin D (ACD) binding to Hco-Pgp-13-04 (A, C, D and E) and -52 (B, C, D and F). A and B. Energy clustering histograms of ACD docked to Hco-Pgp-13-04 (A) and Hco-Pgp-13-52 (B). The lowest energy cluster (ACD1) of each docking is squared in dark blue (for HcoPgp13-04) or light blue (for HcoPgp13-52). C and D. binding sites of the 1st lowest energy clusters of ACD (ACD1) on each protein represented within the two superimposed models of Hco-Pgp-13 (C) or without the proteins, for better comparison of ACD locations (D). E and F. Zoom on the common interacting residues of ACD1 on Hco-Pgp-13-04 (E) and Hco-Pgp-13-52 (F). Hco-Pgp-13-04 is represented in purple ribbon and Hco-Pgp-13-52 in green ribbon, using PyMol (C, E and F). ACD1 is represented in dark blue spheres for docking on Hco-Pgp-13-04 and in light blue spheres for docking on Hco-Pgp-13-52 (C). Alternatively, ACD1 is represented in sticks (D) or lines (E and F) with colored elements (red for O, blue for N, white for H). Interacting residues of Hco-Pgp-13 are represented in sticks with colored elements **(**C of the same color as the ribbon, red for O, blue for N, yellow for S, white for H**).** All images were generated using PyMol.

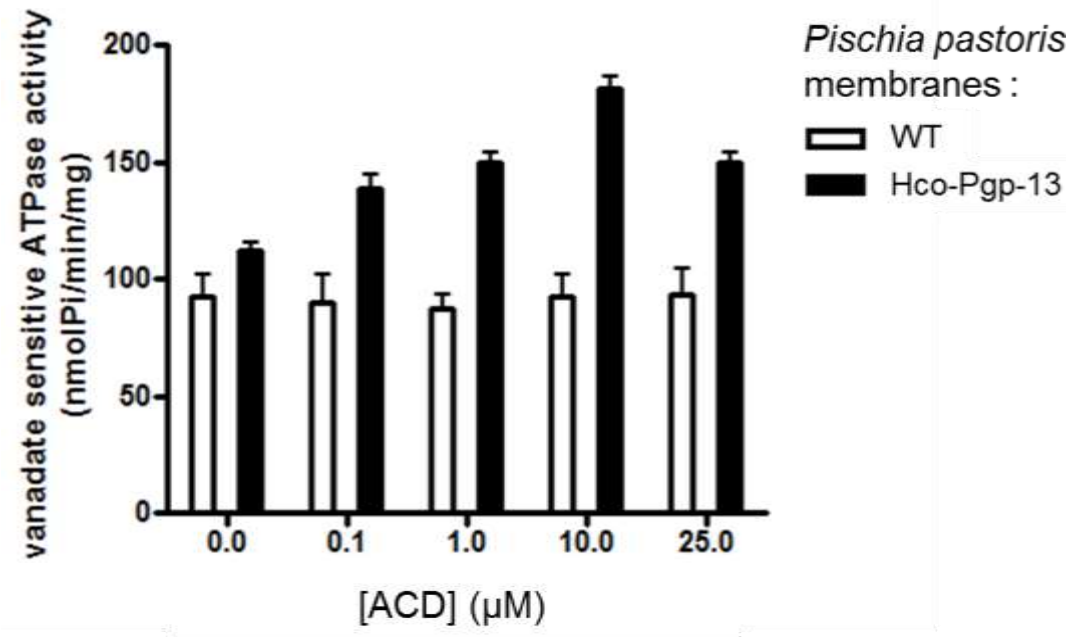


Figure 6. Stimulation of the ATPase activity of Hco-Pgp-13 expressed in *Pichia pastoris* membranes by actinomycin D. The vanadate-sensitive ATPase activity is represented in nmol of Pi /min /mg of protein in membranes from control, untransfected *P. pastoris* cells (white bars) or membranes expressing Hco-Pgp-13 (black bars), as a function of actinomycin D concentration in µM. Error bars have been calculated on triplicates within one experiment, which has been replicated three times, and one representative experiment is shown.

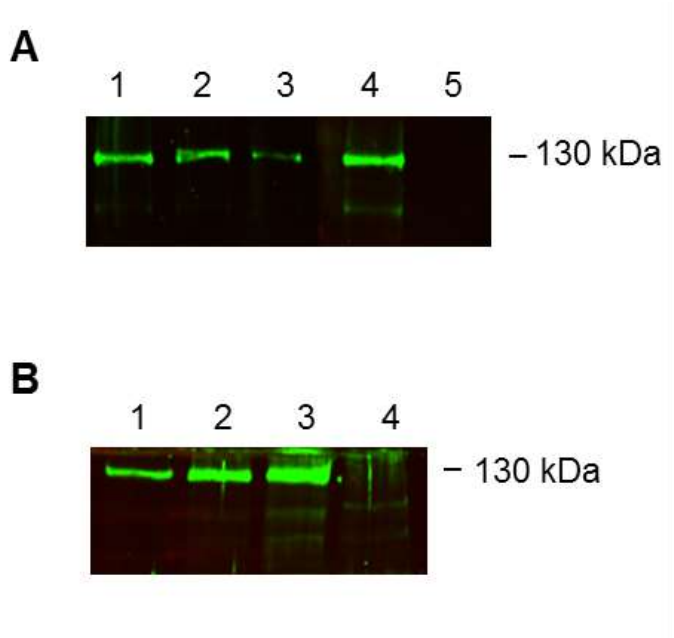


Figure 7. Expression of Hco-Pgp-13 in *Pichia pastoris* membranes. Gel electrophoresis of proteins was performed with *P. pastoris* membranes expressing Hco-Pgp-13 and Western blots were performed with the two antibodies raised against Hco-Pgp-13: anti-epitope 1 (**A**) and anti-epitope 2 (**B**) incubated separately at 1/1000 dilution. **A.** Antibody anti-epitope 1 of Hco-Pgp-13 was incubated with 1.5 µg (Lane 1), 1.0 µg (Lane 2) or 0.5 µg (Lane 3) of *P. pastoris* membranes expressing Hco-Pgp-13. As a comparison, the antibody anti-epitope 2 of Hco-Pgp-13 was incubated with 2.5 µg (Lane 4) of *P. pastoris* membranes expressing Hco-Pgp-13. As a negative control, the antibody anti-epitope 1 of Hco-Pgp-13 was incubated with 10 µg of membranes from WT *P. pastoris* (Lane 5). **B.** Antibody anti-epitope 2 of Hco-Pgp-13 was incubated with 2.5 µg (Lane 1), 5 µg (Lane 2) or 10 µg (Lane 3) of *P. pastoris* membranes expressing Hco-Pgp-13, or with 10 µg of membranes from WT *P. pastoris* (Lane 4). Hco-Pgp-13 protein was detected in all cases around 130 kDa.

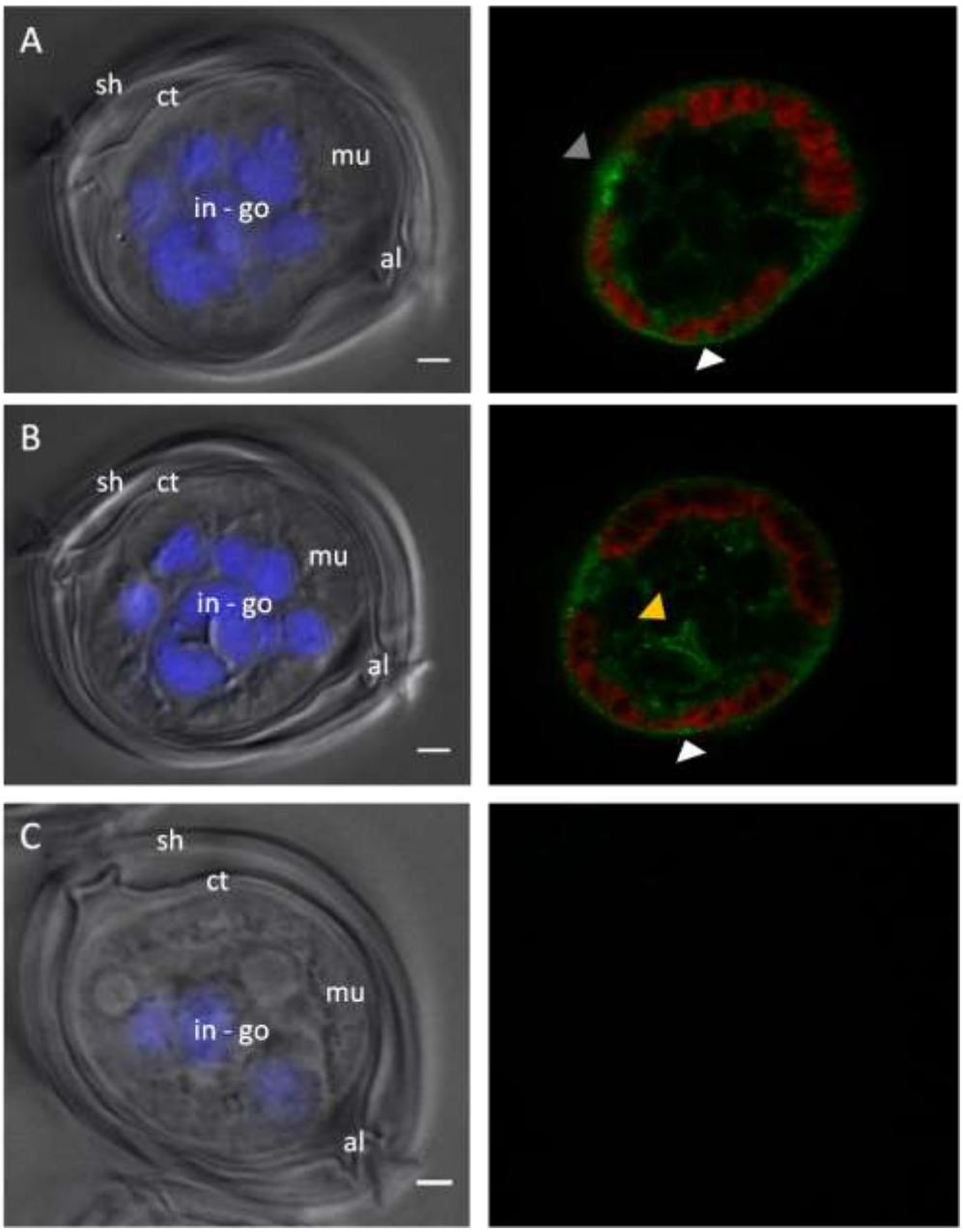


Figure 8. Immuno-localization of Hco-Pgp-13 in the L3 *Haemonchus contortus* larvae. Left panel: differential interference contrast (DIC) image and DAPI signal superimposed, right panel: myosin and Hco-Pgp-13 staining superimposed. A. and B. Different Z-stacks of one slice in the mid-body observed after incubation of primary and secondary antibodies. C. Slice in the mid-body observed with no primary antibody incubation. All slices were incubated with DAPI and numerous nuclei are observed in the gonad and intestine in development and not well distinguishable. sh = supplementary sheath of the L3 stage larvae, ct = cuticle, al = alae, in = intestine, go = gonad, mu = muscle. White arrow: hypodermis, grey arrow: seam cell, yellow arrow: epithelial cells of the gonad or intestine. Scale bar = 2 µm.

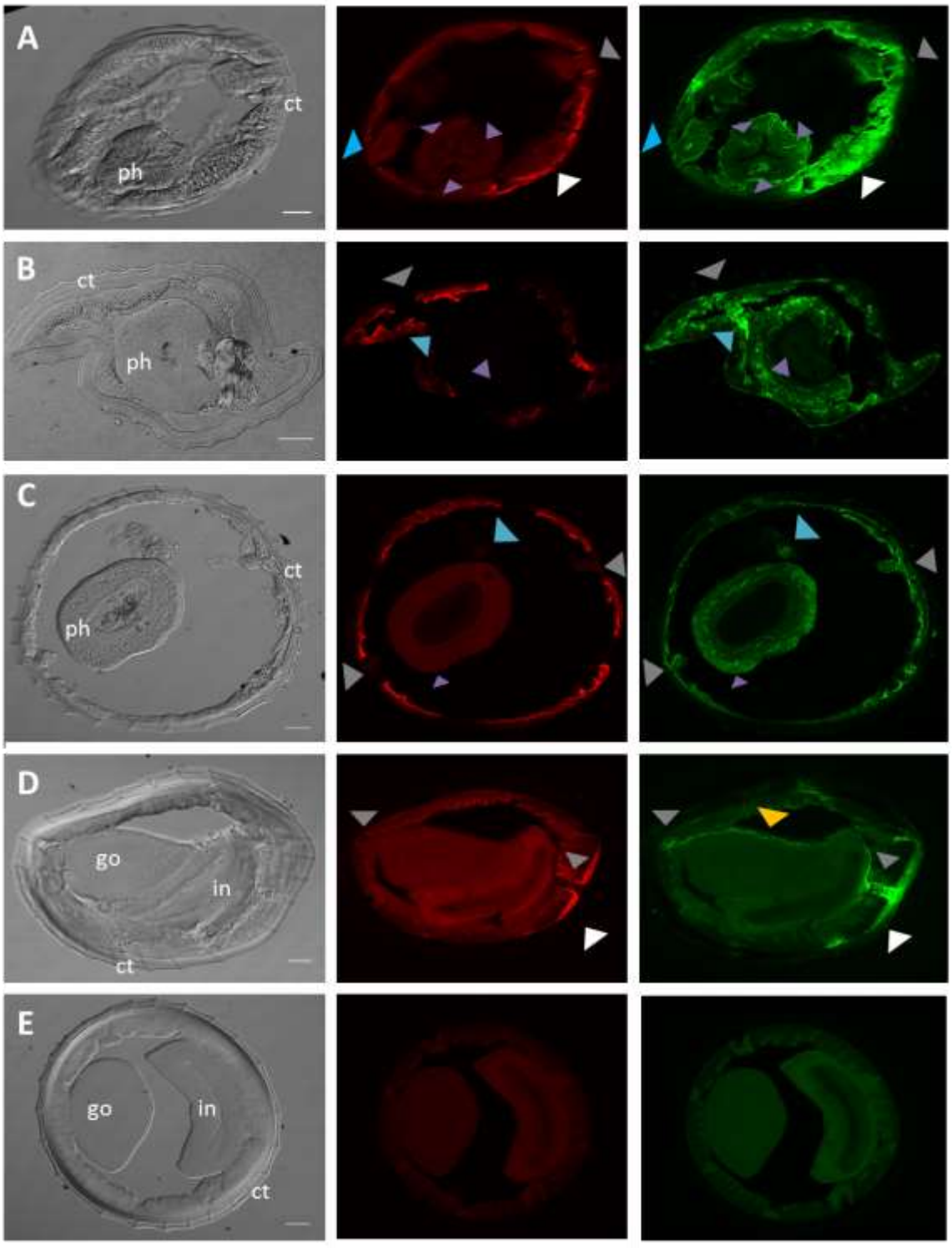


Figure 9. Immuno-localization of Hco-Pgp-13 in adult *H. contortus*. Left panel, DIC image; middle panel, myosin staining; right panel, Hco-Pgp-13 staining. A. Slice in the anterior region of the pharynx (procorpus). B. Section in the mid-region of the pharynx (metacorpus). C. Section in the posterior region of the pharynx (terminal bulb). D, E. Slices in the mid-body of a male parasite. A – D: incubation of primary and secondary antibodies. E: no primary antibody incubation. ct = cuticle, ph = pharynx, go = gonad, in = intestine. White arrow: hypodermis, grey arrow: seam cell, blue arrow: neuronal structures, purple arrow: pharyngeal glands, yellow arrow: gonad. Scale bar = 20 µm.

Tables

Table 1. Percentage of amino acid identity and similarity of mean (TMD1-TMD2) protein sequences, as determined by BlastP. TMDs were defined by the multiple sequence alignment with Muscle partly shown in (Suppl. Fig. S2).

Cel-Pgp-1	% Identity	% Similarity
Hco-Pgp-1	63	79
Hco-Pgp-9.1	50	69
Cel-Pgp-9	48	67
Hsa-Pgp	36	58
Cel-Pgp-2	35	57
Mmu-Abcb1a	35	56
Hco-Pgp-2	34	56
Hco-Pgp-13	**33**	**54**
Hco-Pgp-16	32	52
Cel-Pgp-3	30	52
Hco-Pgp-3	31	51
Cel-Pgp-4	30	51
Cel-Pgp-8	29	51
Cel-Pgp-12	**28**	**51**
Hco-Pgp-17	25	51
Cel-Pgp-5	30	50
Cel-Pgp-14	**29**	**50**
Cel-Pgp-13	**27**	**50**
Cel-Pgp-7	29	49
Cel-Pgp-6a	29	49
Hco-Pgp-11	25	46
Cel-Pgp-11	**22**	**45**
Hco-Pgp-10	**22**	**42**
Cel-Pgp-10	**ND**	**ND**

For each Pgp, TMD1 was considered as starting at the amino-acid aligned with the first helix of TM1 of Cel-Pgp-1 as given by the 4F4C crystal, without taking account for TMa-b, and ending at the amino-acid aligned with the last of TM6 of Cel-Pgp-1. The same was done with TMD2 starting at TM7 and ending at TM12 of Cel-Pgp-1 4F4C structure. Each TMD homology was calculated independently and the mean of the two values is indicated. Bold: Pgps showing a long N-terminal sequence (>70AAs). The list starts with the highest percentage of similarity until the lowest percentage of similarity. ND: Cel-pgp-10 showed a longer TMD1 than other Pgps, which prevented the software from calculating homology percentages.

Table 2. Docking characteristics of actinomycin D on Hco-Pgp-13.

Protein model	Hco-Pgp-13_04	Hco-Pgp-13_52
Binding energy (kcal/mol)	-16.02	-14.42
Number of poses	21	42
Number of interacting residues	20	19
Number of hotspot residues	11	9
Number of H-bonds	2	2

Various characteristics listed in the 1st column are indicated for the lowest energy cluster of actinomycin D on Hco-Pgp-13-04 (2nd column) and Hco-Pgp-13-52 (3rd column).

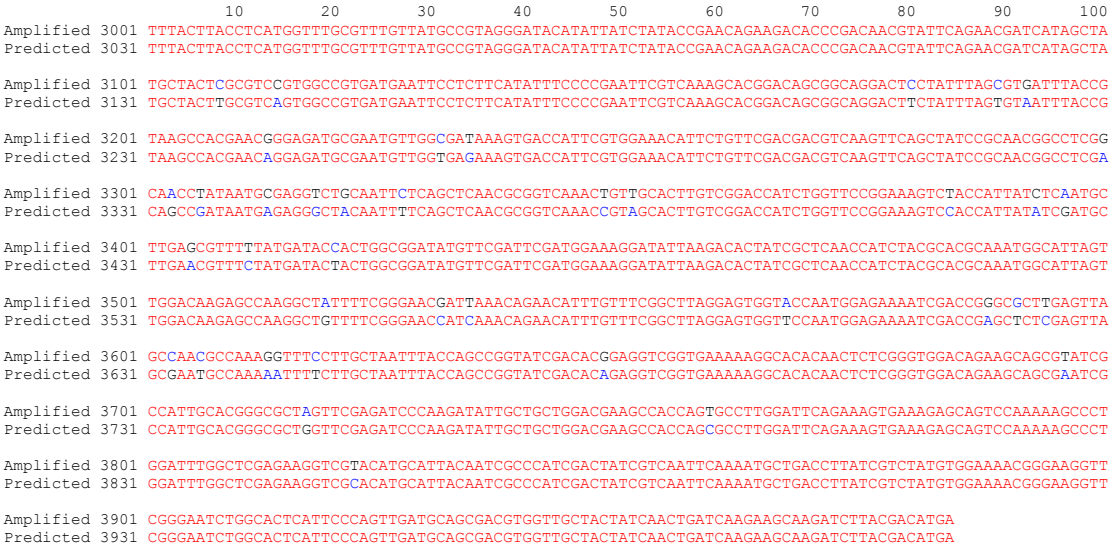
Table 3. List of interacting residues of each transmembrane helix (listed in the 1st column) of Hco-Pgp-13-04 (2nd column) or Hco-pgp-13-52 (3rd column) with the lowest energy cluster of actinomycin D found for each of these models.

Molecule	ACD	
Hco-Pgp-13_	04	52
Cluster rank	1	1
TMa-b	R11	
	S19	
	E21	E21
	K25	K25
	L29	L29
TM1	**Q108**	**Q108**
TM3	L217	
		R220
TM6	**L365**	**L365**
		L366
	Y369	**Y369**
		L373
	H377	**H377**
TM10	**L912**	
	M916	M916
	L919	L919
		A920
ICL4	L921	
TM11	**Q989**	
	T992	
	M995	**M995**
TM12		**V1029**
	M1032	**M1032**
	N1033	**N1033**
		S1035
	S1036	
Nter		P1039

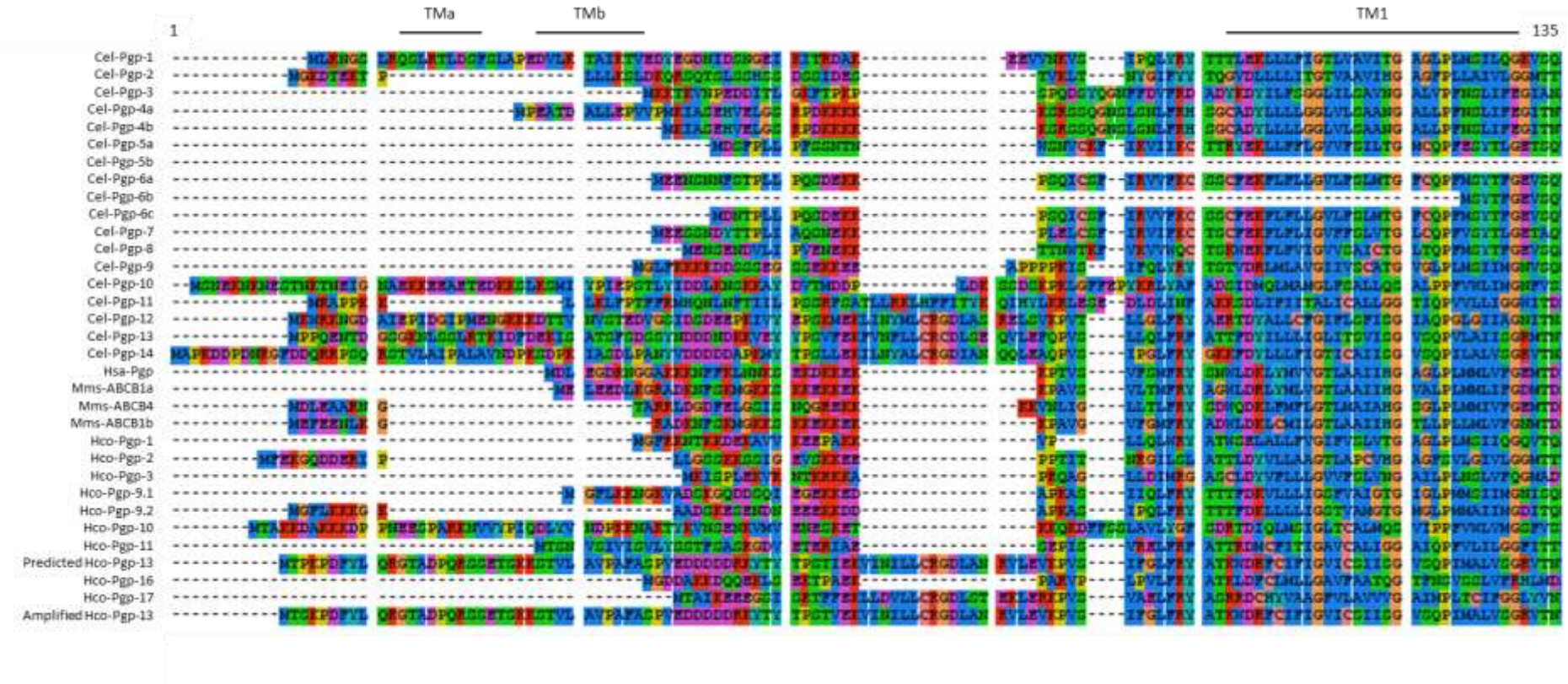
Bold: hotspot residues. Underscored: residues establishing a H-bond.

Supplementary Information

Supplementary Figures



Supplementary Figure S1. Alignment of full length *Hco-pgp-13* amplified cDNA sequence with full length predicted sequence of Laing et al. (2013). Common sequences are represented as red characters. The 30 nucleotide repeat is highlighted in grey. SNPs are indicated as blue or black characters.



Supplementary Figure S2. Comparison of Hco-Pgp-13 protein N-terminal sequence with all Hco-Pgps, Cel-Pgps and four mammalian Pgps: Hsa-Pgp, Mmu-ABCB1a, Mmu-ABCB1b and Mmu-ABCB4. Multiple sequence alignment with Muscle, representation using Clustal X Default Colouring. Amino-acid number and helices annotations of Cel-Pgp-1 are indicated according to 4F4C PDB structure (Jin et al., 2012).

Supplementary Figure S3. Alignment of Cel-Pgp-1 protein sequence from 4F4C PDB entry with Hco-Pgp-13 protein sequence translated from the amplified cDNA for homology modelling of its 3D structure. The sequence of the 4F4C crystal structure of Cel-Pgp-1 lacks AA M1-R3, A52-E54, K666-E715 and G1307-K1321 and the sequence of Hco-Pgp-13 lacks M1-S3, E668-L725, and G1304-T1317. Helices annotations of Cel-Pgp-1 are indicated according to 4F4C PDB structure (Jin et al., 2012). Representation using Clustal X Default Colouring.

Supplementary Tables

Supplementary Table S1. Primers used for amplification or sequencing of *Hco-pgp-13* cDNA sequence.

F/R	Name	Sequence	Start
R	mda 1	AACTAGGATTCATTTCTCCACAAATACG	3992
F	mda 4	AATTTCGAAAATGCGTCACGAATATGTT	504
F	mda 5	TAAAAACACCTGGCGATGTGTTTATT	1055
F	mda 6	ACCATTCGAGATGCCAATAAAATTGTG	1851
R	mda 7	TCATTTTCGGTTCCATTGGTGTTGG	2405
R	mda 8	ATGGTGTACGTATGCAAAATGGCAAAT	1605
F	mda 15	TTCCATTGGTGTTGGTGTCG	2415
R	mda 16	AACAGAATGTTTCCACGAATGG	3215
F	SL1	GGTTTAATTACCCAAGTTTGAG	- ?
R	mda 31	AGAATCCTTTCCAAATCG	965
F	mda 48	GCGAAAAAAGTGATGGTG	-34
R	mda 19	CCATAAGCGATTCTTCAGC	853

Supplementary Table S2. Parameters of evaluation of the best DOPE score (Hco-Pgp-13_04) and best molpdf score (Hco-Pgp-13_52) among the 100 Hco-Pgp-13 3D models generated by Modeller using Cel-Pgp-1 4F4C as template.

	Hco-Pgp-13_04	Hco-Pgp-13_52	Cel-Pgp-1 4F4C
Modeller			
DOPE score	-151276	-150298	
molpdf score	5791	5292	
Qmean score	0.582	0.592	0.566
C_beta interaction energy	60.80 (Z-score: -2.41)	69.12 (Z-score: -2.46)	-208.94 (Z-score: -1.11)
All-atom pairwise energy	-12804.05 (Z-score: -2.09)	-13154.19 (Z-score: -2.05)	-22101.92 (Z-score: -1.00)
Solvation energy	-74.27 (Z-score: -0.99)	-74.11 (Z-score: -0.99)	-88.11 (Z-score: -0.64)
Torsion angle energy	-105.95 (Z-score: -3.00)	-124.74 (Z-score: -2.76)	-118.29 (Z-score: -2.85)
Secondary structure agreement	86.8% (Z-score: 1.60)	86.1% (Z-score: 1.46)	83.6% (Z-score: 0.96)
Solvent accessibility agreement	73.7% (Z-score: -1.36)	74.1% (Z-score: -1.28)	72.8% (Z-score: -1.52)
ProSA-web			
Overall model quality Z-Score:	-12.29	-12.33	-14.35
VADAR Ramachandran plot (%)			
Most favored	92.8	93.1	91.3
Allowed	6.0	6.2	7.4
Generously allowed	0.8	0.5	1.2
Disallowed	0.4	0.2	0.1

Supplementary Table S3. List of "hotspot residues" that have been identified as being involved in multispecific substrate recognition in mammalian Pgp, and aligned residues in Hco-Pgp-13.

Hsa-Pgp	Mmu-ABCB1a	Cgr-ABCB1	Mmu-ABCB1b	Origin	Hco-Pgp-13
H61	H60	H60	H60	Mv	S104
G64	A63	A63	L63	Mv	S107
L65	L64	L64	L64	**Mv/CV/SL**	Q108
M68	**M67**	M67	L67	SL	M111
M69	**M68**	M68	M68	S	A112
F72	**F71**	F71	F71	SL	S115
Y118	**Y114**	Y115	Y117	SL	F143
S222	S218	S219	S221	CV	T247
I299	**M295**	M296	I298	SL	G324
F303	**F299**	F300	Y302	SL	F328
L304	**L300**	L301	L303	SA	F329
I306	I302	I303	V305	**M/SL**	L331
Y307	**Y303**	Y304	Y306	S	F332
Y310	**Y306**	Y307	Y309	SL	L335
F335	F331	F332	F334	Mva	V361
F336	**F332**	F333	F335	S	M362
V338	V334	V335	I337	Mv	L364
L339	**L335**	L336	L338	**CV/S**	L365
I340	**I336**	I337	L339	**CR/S**	L366
G341	G337	G338	G340	M	G367
A342	A338	A339	T341	CV	A368
F343	**F339**	F340	F342	S	Y369
S351	D347	D348	**D350**	M	H377
N721	**N717**	N718	N719	SL	L769
Q725	**Q721**	Q722	Q723	S	R765
F728	**F724**	F725	F726	**CV/S**	L772
F732	**F728**	F729	F730	S	F776
L762	**L758**	L759	M760	SA	I806
F770	**F766**	L767	Y768	SL	Q814
T837	**F833**	T834	T835	SA	I881
I840	I836	**I837**	V838	M	M884
A841	A837	A838	A839	CR	T885
N842	N838	**N839**	N840	M	A886
I864	I860	I861	**I862**	M	M908
I867	I863	I864	**L865**	M	I911
I868	I864	I865	I866	CV	L912
I870	I866	I867	**L868**	M	T914
A935	A931	A932	**A933**	M	G978
F938	**F934**	F935	F936	M	E981
F942	**F938**	F939	F940	**M/CV**	F985

S943	**S939**	S940	**S941**	Mva	T986
T945	**T941**	T942	T943	**Mv/CV**	S988
Q946	**Q942**	Q943	Q944	M	Q989
M949	**M945**	M946	M947	SL	T992
Y950	**Y946**	Y947	Y948	Mva	Y993
S952	**S948**	S949	S950	Ma	M995
Y953	**Y949**	Y950	Y951	**Ma/S**	V996
F957	**F953**	F954	F955	**Mva/SL**	Y1000
L975	L971	L972	M973	**M/CR/SL**	F1018
F978	**F974**	F975	F976	**M/S**	I1021
S979	**S975**	S976	S977	S	I1022
V981	I977	I978	V979	**M/CR**	M1024
V982	**V978**	V979	V980	**CR/S**	L1025
F983	F979	F980	F981	**M/SL**	L1026
G984	G980	G981	G982	CV	A1027
A985	**A981**	A982	A983	SA	S1028
M986	**M982**	M983	M984	S	V1029
A987	**A983**	A984	A985	SL	A1030
G989	**G985**	G986	G987	SA	M1032
Q990	**Q986**	Q987	N988	S	N1033
V991	**V987**	V988	T989	SL	S1034
S993	**S989**	S990	S991	SA	S1036

Hotspots are distributed among the various TM helices, as quoted in the first column. The orthologs: human (Hsa), mouse (Mmu) a and b isoforms and Chinese hamster (Cgr), in which the residue has been indicated, are shown in columns 2-5 by a bold residue symbol. Hotspots were determined, in the literature, by different experimental approaches (point mutagenesis, chemical labelling, co-crystallization, respectively denoted as M, C, S), as quoted in the sixth column; when the residue has been found by at least two different approaches, the symbols are highlighted in bold. M: cytotoxicity assay after Pgp site-directed mutagenesis using, among other drugs, VBL (v) or ACD (a); C : MTS-mediated chemical labelling by a MTS-derivative of verapamil (V) or rhodamine (R) of Cys-free Pgp specifically bearing Cys residues obtained by scanning mutagenesis; S: contact residues of the co-crystallized ligands QZ59RRR and QZ59SSS in the Pgp crystal structures published in Aller et al. (2009) and Li et al (2014), but presenting few differences (A: residues found in the 2009 structure alone, L: residues found in the 2014 structure alone). See Computational Methods section for the relevant references. Last column: corresponding residues found in Hco-Pgp-13 after multiple sequence alignment with mammalian Pgps (partly shown in Suppl. Figure S2).

Supplementary Table S4. Amino acid differences between the predicted and amplified sequences of Hco-Pgp-13.

Region	Predicted Hco-Pgp-13	Amplified Hco-Pgp-13
N-terminal	Pro 3	Ser 3
TMD1	Lys 135	Glu 135
linker	Lys 729	Arg 729
TMD2	Gly 814 - Phe 823	Deletion
TMD2	Ser 827	Ile 817
TMD2	Glu 836	Tyr 826
TMD2	Val 921	Ile 911
Between TMD2 and NBD2	Glu 1078	Asp 1068
NBD2	Asn 1205	Gly 1195

The region of the change is indicated in the first column, with the nature of each amino acid and its number in the sequence of predicted Hco-Pgp-13 in the second column, and of amplified Hco-Pgp-13 in the third column.

Supplementary Table S5. Location (indicated as first and last amino acid number) of TM helices, which number appear in the 1st column, as predicted by Protter (2nd column), TMHMM (3rd column) or as found in the homology model built on Cel-Pgp-1 4F4C PDB structure (4th column).

TM	Protter	TMHMM	Homology Model
1	92-114	92-114	88-116
2	140-164	140-162	139-162
3	215-235	216-235	194-236
4	241-258	239-258	237-261
5	318-345	322-344	314-339
6	357-375	359-381	358-378
7	754-780	754-776	752-777
8	800-829	801-823	800-824
9	880-897	878-897	873-898
10	903-923	901-923	899-920
11	991-1008	985-1007	978-1002
12	1020-1038	1017-1039	1018-1038

B. Supplementary experiments for the localization of *Hco-pgp-13* mRNA in the parasite

With the objective to localize the expression of the transcription product of *Hco-pgp-13* gene in *H. contortus*, we first attempted to detect its mRNA in larvae using fluorescence *in situ* hybridization on larvae permeabilized by the "freeze and crack" technique. The results are reported here, and show that this strategy proved to be inadapted for studying of the localization of putatively low expressed mRNAs. This led to our choice of detection of Hco-Pgp-13 protein by worms cryosectionning followed by immuno-histochemistry, as previously described in Part II, A.

I. MATERIAL AND METHODS

1. DNA Probes amplification

The most specific region to Pgps is known to be the linker domain separating the 1st NBD and the 2nd TMD of the protein. That of HcoPgp13 was determined using ExPASy ProSite prediction tool as covering amino acids 655 – 756. Their corresponding nucleotides were identified as nt 1966 – 2268 using ExPASy Translate tool. Their specificity was checked by alignment with other Pgps and it was found to be maintained until nt 2400. Two Digoxigenin (Dig)-labeled probes were thus designed on two separate sequences within this region. For this purpose, a first PCR amplification of DNA probes was carried out on *H. contortus* adult cDNA using 2 sets of primers (sequences in Table 1) amplifying a 183bp and a 144 bp non-overlapping fragments, as illustrated in Fig. 1. In addition, a 290 bp region of the beta-tubulin gene of *H. contortus* was amplified with two other primers to design a positive control probe (sequence in Table 1).

Each of the amplification products were run on a 1.5% agarose gel, and the expected size fragments were extracted and purified. They were then cloned in a TOPO-TA vector and cultured in TOP10 F' *Escherichia coli* overnight. Positive clones with a blue/white screening were purified and vectors were sequenced to check for their identity. *E. coli* cultures containing the plasmids with confirmed sequences were stored at -80°C. The promoter for the T7 RNA polymerase was then inserted in the cDNA probes by amplification with one of the

F/R	Name	Sequence
F	mda20	AAGTTGAAGAAGAAGAGGTC
R	mda25	CAAACGAGTCATTAAGAGC
F	mda27	ACAGGAAATTACCACTGG
R	mda22	ATAATCTGTGCATCATATCG
F	Btub F	GACGCATTCACTTGGAGGAG
R	Btub R	CATAGGTTGGATTTGTGAGTT
F	T7mda20	**TAATACGACTCACTATAGGG**AAGTTGAAGAAGAAGAGGTC
F	T7mda25	**TAATACGACTCACTATAGGG**CAAACGAGTCATTAAGAGC
R	T7mda27	**TAATACGACTCACTATAGGG**ACAGGAAATTACCACTGG
F	T7mda22	**TAATACGACTCACTATAGGG**ATAATCTGTGCATCATATCG
R	sT7BtubF	**TAATACGACTCACTATAGGG**ATTCACTTGG
F	sT7BtubR	**TAATACGACTCACTATAGGG**TTGGATTTGTGT
R	Lmda27	ACAGGAAATTACCACTGGATATTCCTTGGC

Table 1. Primers used for amplification of DIG-labeled probes. The sequence of the T7 promoter is indicated in bold and underlined characters. F = forward; R = reverse.

forward or reverse primers starting with the T7 sequence (TAATACGACTCACTATAGGG): the reverse primer to translate into the anti-sense RNA probe, and the forward primer to translate into the sense negative control probe (Table 1). For the second fragment, a longer forward primer Lmda27 was designed to amplify the anti-sense probe with T7mda22 for homogenization of the Tm of the set of primers. This was not required for the first fragment. The Beta-tubulin fragment was amplified using shorter T7 primers than others to keep a Tm close to previously designed primers. Their specificity was checked in every case using Nucleotide Blast (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The PCR products were cloned and sequenced as previously to check for their identity.

2. RNA probes amplification and validation

The transcription reaction was performed following the instructions given by the manufacturer (DIG RNA labeling kit, Roche) with 100-200 ng of purified template DNA diluted into RNAse-free water up to 13 µL; 2 µL of NTP labeling mixture containing 10x 10 mM each of ATP, CTP, GTP, 6.5 mM UTP and 3.5 mM DIG-11-UTP, pH 7.5; 2 µL of 10X Transcription buffer; and 2 µL of RNA polymerase T7. This mix was incubated at 37°C for 2h, 2 µL of DNase I was added to remove the template DNA and incubated at 37°C for 15min. The reaction was then stopped with 2 µL of EDTA 0.2 M, pH 8. The concentration of

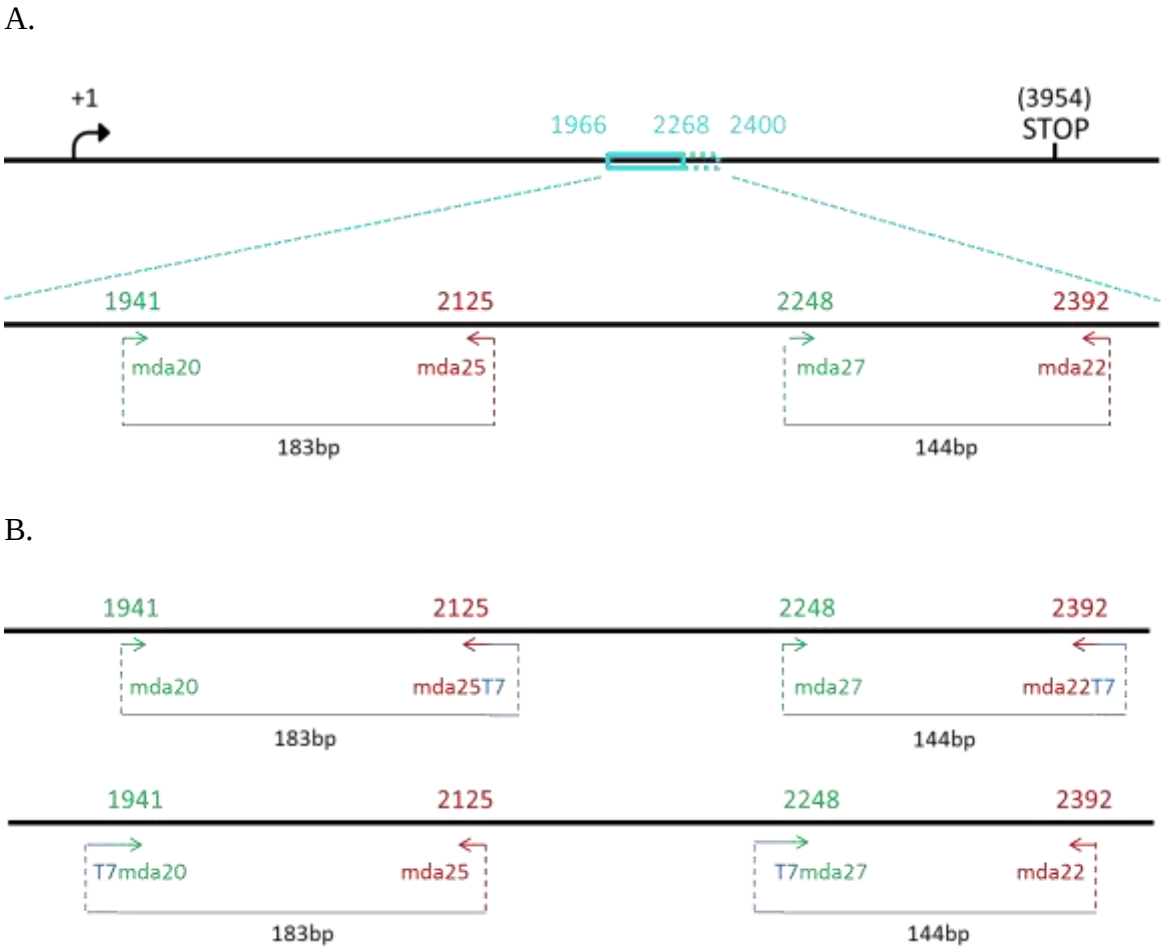


Figure 1. Schematic representation of the first two steps of amplification of DIG-labelled RNA probes. A. Amplification of cDNA probes. **B.** Amplification of T7promoter – cDNA probes.

transcribed RNA was measured with a Nanodrop and its 260/280 ratio was checked to be above 1.8 for good purity.

The labeling efficiency of the probes was assessed by a series of dilutions of DIG-labeled probes from 0.01 pg/µL to 10 ng/µL, as described in the DIG Northern starter kit (Roche). A positively-charged nylon membrane was loaded with 1 µL of each of these different concentrations of probe. Fixation of RNA to the membrane was done by crosslinking with UV light. The membrane was then successively incubated into 20 mL of washing buffer for 2min, in 10 mL blocking solution for 30min, in 10 mL antibody solution containing the anti-digoxigenin antibody coupled to alkaline phosphatase (anti-DIG-AP) for 30min, in 20 mL washing buffer 2 x 15min, and in 10 mL detection buffer for 2-5min. CDP-Star was applied on the membrane incubated at room temperature (RT) for 5min in a development folder. Exposure

to the imaging device at room temperature for 5-25min revealed until which dilution some signal could be observed, thus indicating the labeling efficiency.

3. Fluorescence *in situ* Hybridization

The "freeze&crack" method was used to permeabilize L2 stage larvae. 20 µL of PBS containing concentrated fresh larvae were compressed between 2 slides and immediately incubated on dry ice for at least 30min. Alternatively, they were conserved at -80°C until the next step. Slides were cracked and 200 µL PFA 4% was immediately added on each slide. Larvae were then detached from slides, pooled in a 10mL glass tube and incubated for 30min at 4°C. They were subsequently pelleted by centrifugation at 2000 rpm for 1min at 4°C. Fixation was carried out by incubation of larvae in ice cold methanol for 1 min, centrifugation for 1min at 2000 rpm at 4°C and removal of the supernatant. Larvae were then incubated in ice cold acetone for 1 min, centrifugation for 1min at 2000 rpm at 4°C and removal of the supernatant. Finally, larvae were washed twice in PBS, by successive centrifugation and removal of the supernatant.

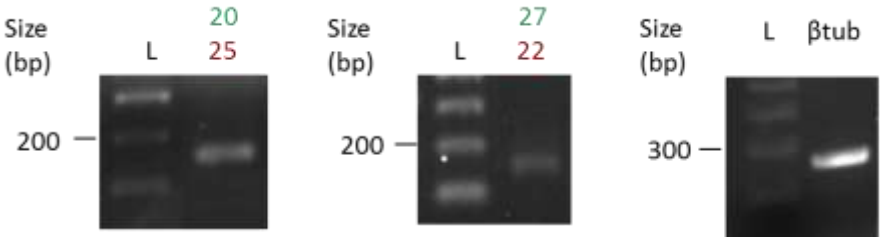
For the FISH experiment, larvae were incubated in formamide hybridization buffer (KPL) + 1/100 salmon sperm DNA and spread evenly in eppendorf tubes that would contain the different probes. These were incubated at hybridization temperature for 1h under agitation. Probes were denatured at 68°C for 10min and diluted in hybridization buffer containing larvae to 1 µg/mL final concentration. This mixture was incubated at 46°C hybridizing temperature O/N under agitation. Larvae were washed three times in maleic acid buffer (MAB) at RT for 5min, and blocked in blocking buffer at RT for 1h. The anti-DIG FITC-conjugated antibody (Roche) was then incubated at 1/100 dilution for 4h at RT. Three washes were performed in MAB at RT for 5min. Larvae were mounted on slides in 1 drop of mounting medium + 0.2 µL DAPI 5 mg/mL. Visualization was performed by fluorescence microscopy.

II. RESULTS

1. Probes amplification and validation

After amplification of probes and running of the PCR products on a 1.5% agarose gel, the obtained fragments appeared around the expected lengths for each probe: 183bp for mda20-mda25, 144bp for mda27-mda22 and 290bp for β-tubulin (Fig. 2A). They were each

A.



B.

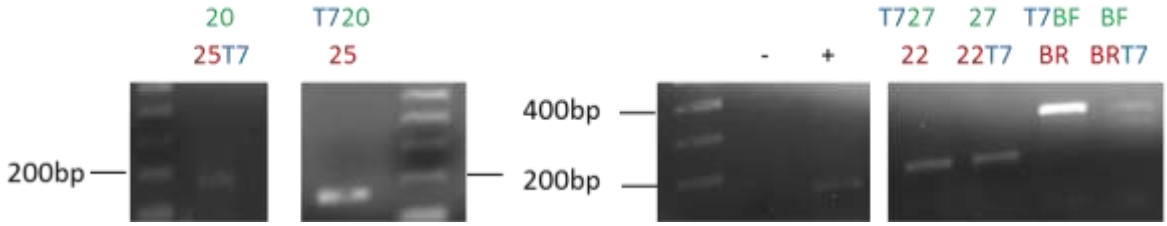


Figure 2. Result of the running on a 1.5% agarose gel of amplified cDNA probes before (A) and after (B) T7 promoter incorporation.

cloned in TOP10 F' *E. coli* (see M&M) and between 3 and 5 colonies were chosen for plasmid extraction and sequencing. The sequences obtained were aligned to the expected sequence with MultAlin (Corpet, 1988) and checked to be 100% identical.

The T7 promoter incorporation led again to PCR products of the expected size with fragments of Hco-Pgp-13 between 100 and 200bp; and fragment of β-tubulin around 400bp (Fig. 2B). As before, 4 colonies transformed with each of these fragments were plasmid extracted and their sequence was checked with Multalin (Corpet, 1988) to be 100% identical to the expected sequence.

After transcription of these probes into RNA with the T7 RNA polymerase, according to the Northern Starter kit, the various fragments concentration and A260/A280 ratio were checked. All the concentrations were above 100ng/ µL, and A260/280 ratios were above 1.8 (Table 2), indicating suitable quantity and purity of RNA fragments for performing *in situ* hybridization.

Fragment	[ ] (ng/µL)	A260/A280

+	420	1.909
LβF-βRsT7	328	1.952
20-25T7	230	1.949
T720-25	268	1.861
sT7βF-LβR	179	1.956
T727-22	177	1.991
L27-22T7	159	1.847

Table 2. Concentration of RNA fragments and purity ratio A260/A280.

Finally, these probes were confirmed to label *H. contortus* adult mRNA after dilution up to 0.3 pg/µL or 0.1 pg/µL in Dot-Blot (Fig. 3). Each set of sense and anti-sense probes showed similar detection levels; which is a required condition for the control of specificity of the antisense probe during the *in situ* hybridization, by co-incubating the two probes.

2. Fluorescence *in situ* hybridization

Finally, each of the specific probes were used to perform fluorescent *in situ* hybridization using different conditions of washing and hybridization temperature. A signal was found with the anti-sense probe targeted against β-tubulin with the minimum of washes (i.e. without the recommended SSC 2X washes) after hybridization at 46°C (Fig. 4). The signal was present in all the organs of the worm, as expected for this probe, and was significantly stronger than the signal found in larvae incubated with the sense probe. No signal could be detected, however, in worms that were not correctly permeabilized. Plus, the signal was constrained in other worms around the permeabilized zone. Consequently, stained worms had highly damage morphology, with organs that would hardly be recognizable using probes that target specifically a few of them. On the other hand, no significant staining was observed under these conditions in worms incubated with the probes recognizing various regions of the linker domain of *hco-pgp-13*.

III. DISCUSSION

The major issue of the FISH technique is that the probe is labile, because sensitive to RNases, which brings many difficulties regarding the high number of steps in this experiment.

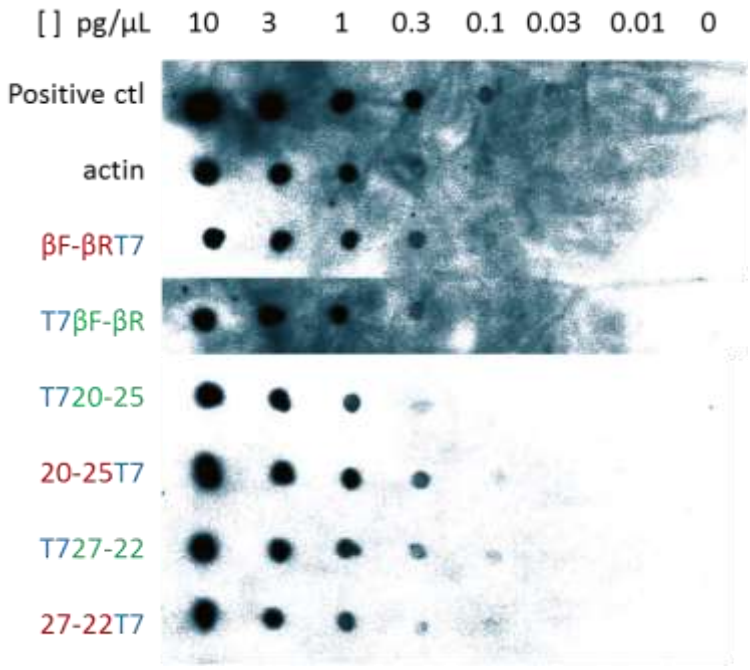


Figure 3. Dot-Blot with designed probes and internal controls (positive control of transcription, and DIG-labelled actin provided as positive control of the dot-blot).

Indeed, all the incubations and washing must be done in a RNase-free environment that can easily lead to absence of signal if this is not the case. However, because we found a signal with the anti-β-tubulin probes, these conditions seem to have been respected and should not explain why no signal was found with probes targeting *Hco-Pgp-13*.

The optimal hybridization temperature, found to be 46°C for the β-tubulin probe, could be different for *Hco-pgp-13* probes. Indeed, these are of about half the length of the positive control probe, and the size is base pair is one of the determining factors of the hybridization temperature, so that other temperatures would need to be tested. On the other hand, probes longer than 250bp are recommended for optimal sensitivity and specificity of FISH experiment, but this length of probes could not be successfully amplified from *Hco-pgp-13* cDNA sequence. This might explain why the anti-β-tubulin 290bp probe led to signal when anti-*Hco-pgp-13* probes did not. Plus, the labeling efficiency of regular FISH protocols can be insufficient in the case where the target is a low-expressed gene. The β-tubulin gene being a housekeeping gene present in all cells, the staining of its RNA could have been visible when that of the *pgp-13* was present but too low.

In this case, some techniques such as the Tyramide signal amplification allow amplifying the DIG signal. This method uses a HRP conjugated anti-DIG antibody that leads

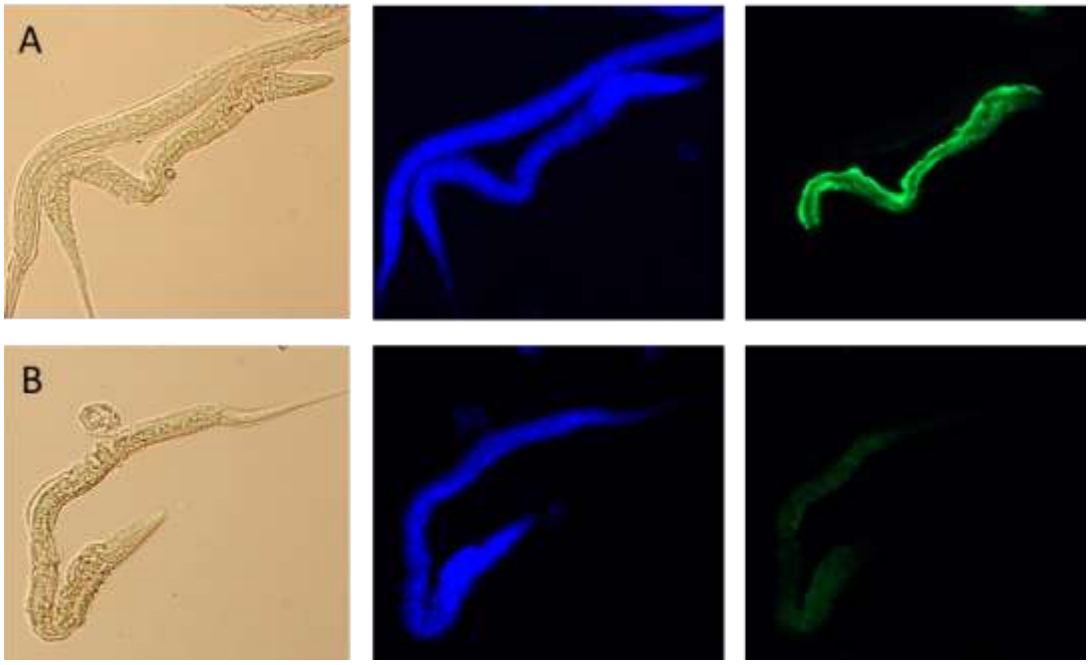


Figure 4. Fluorescence in situ hybridization on L2 stage with anti- β-tubulin probe coupled to DIG and detected by anti- DIG antibody coupled to FITC. DIC image (left panel), staining of nuclei with DAPI (middle panel) and FITC staining of the probe (right panel) after incubation of the anti-sense probe (A) or the sense probe (B).

to a peroxidase-tyramide reaction in presence of H_2O_2 and biotinylated tyramides (Speel et al., 2006). The numerous activated tyramides deposit close to the antibody and can be detected by different methods, for example by fluorescence microscopy if they are labelled with a fluorochrome. This leads to a mean of 5-50 fold increase of sensitivity. On the other hand, the stellaris method uses around 50 different very short singly-labeled sequences of a single RNA to stain its whole length and thus lead to increased signal (Raj et al., 2008).

However, the organs of the worms appeared much damaged when the permeabilization step was sufficient for probes to enter organs. Thus, we would probably not be able to visualize any precise staining in organs expressing *Hco-pgp-13* even after optimizing the conditions of the method by testing other hybridization temperature or using techniques to amplify the signal. Plus, the low number of washes required to be able to visualize some signal even with the anti-β-tubulin gene showed that the RNA probes could easily detach from the tissue and that even after implementation, the FISH method would still present many drawbacks.

Thus, it was decided to design antibodies against specific regions of Hco-Pgp-13 protein to look for the localization of the protein rather than the mRNA, with an immuno-histochemistry method technically more feasible than the FISH. The permeabilization

procedure was also changed from “freeze and crack” to cryosectioning in order to better maintain the morphology of the worm. Finally, this allowed looking for the expression of Hco-Pgp-13 both in larvae and adult stages, whereas the “freeze and crack” technique was unsuccessful with the adult worms due to their larger size that did not allow proper permeabilization without breaking them into species. These results are presented in the manuscript under writing, in Part II, A.

C. Supplementary experiments for the characterization of Hco-Pgp-13 function

In order to characterize the function of Hco-Pgp-13, we first optimized the codons of its cDNA coding sequence for heterologous expression in LLCPK1 mammalian cells, which express a low endogenous level of Pgps, and are commonly used for study of the function of nematode parasite Pgps (Godoy et al., 2015a, 2016; Godoy et al., 2015b; Mani et al., 2016). We performed several functional tests reported here and faced the limitation of the loss of expression of the protein at the membrane of cells along passages. This led us to transfect *Hco-pgp-13* cDNA in *Pichia pastoris* cells, which are known to generally lead to high level expression of proteins encoded by transgenes, and allow studying the ATPase activity of ABC transporters in order to test their substrate profile (Jin et al., 2012). Preliminary data show that PCT might be a ligand of Hco-Pgp-13, in addition to ACD, as shown in Part II, A. Before more *in vitro* experiments are performed, *in silico* docking calculations were run with compounds of interest on the two 3D models of this transporter, constructed as described in Part II, A, to screen for putative substrates of Hco-Pgp-13. These results indicate that the compounds previously predicted to bind with high affinity to Cel-Pgp-1 in Part I, A and B of this manuscript, could also bind with high affinity to Hco-Pgp-13, suggesting a putative multidrug binding capacity for this parasitic nematode transporter. These results will need to be completed by *in vitro* assays.

I. MATERIAL AND METHODS

1. Codon optimization and transfection of *Hco-pgp-13* in LLCPK1 cells

The full-length identified sequence of *Hco-pgp-13* was subjected to codon optimization for pig cells by GenScript (USA). The codon adaptation index was thus changed from 60% to 83%, with 80% the minimum required for high expression level. A Kozak sequence was introduced around the ATG for optimal translation in mammalian system (Kozak, 1987), together with a BamHI restriction site starting before the Kozak sequence and a Xba I restriction site after the stop codon. A histidine tag composed of six consecutive histidine codons was also added just before the stop codon following a Gly-Pro motif for better accessibility of anti- His

tag antibodies. The optimized sequence was then subcloned into the mammalian expression pcDNA3.1(+) vector by site-directed ligation using BamHI and XbaI restriction enzymes (New England Biolabs Inc.). After transformation of this vector in *E. coli* TOP10F' competent cells, *Hco-pgp-13* codon optimized sequence was confirmed by sequencing (Genome Quebec Innovation Centre, McGill University, QC, Canada).

The pig kidney epithelial cell line LLCPK-1 was used as a heterologous system for over-expression of *Hco-pgp-13*, as it shows a low endogenous level of multidrug ABC transporters (Lespine et al., 2007). These cells were grown on a 24-wells plate in 199 medium supplemented with 10% fetal bovine serum (FBS), 100 units/mL penicillin and 100 µg/mL streptomycin. Once the cells reached full confluence, they were cultured with 199 medium alone for 24 hours before being transfected with 37.5 µg of linearized plasmid DNA using Lipofectamine 2000® (Invitrogen). One day later, the cells were transferred back to supplemented medium for 24 hours, after which they were be diluted and grown for 4 weeks in the supplemented medium with 400 µg/mL G418 (Geneticin, Gibco) for selection of the transfected cells. The vector pcDNA3.1(+)/*cat* containing the gene coding for the chloramphenicol acetyl transferase (CAT) was transfected in cells selected with G418, as a positive control. The parental untransfected cells also underwent selection, as a negative control. Resistant colonies were then transferred individually using cloning cylinders (Corning, Life Sciences) into a new flask and screened to select those that were the most resistant to G418, up to 1 mg/mL, that grew the fastest, had the best morphology, and expressed the highest level of *Hco-pgp-13* transcript and protein.

2. Characterization of mRNA expression

Total RNA from cultured cells and tissues was extracted using Trizol Reagent according to the manufacturer's instructions. Total RNA was quantified using a nanodrop ND-1000 spectrophotometer (Nanodrop Technologies Inc., USA). RNA purity was checked by measurement of the A260/280 nm ratio, which was routinely in the range of 1.8–2.0. cDNA was synthesized from 2 mg of total RNA using the High-Capacity cDNA Reverse Transcription kit (Applied Biosystems – Life Technologies, Courtaboeuf, France). Finally, the synthesized cDNA was used as template for real-time quantitative PCR (qRT-PCR) using ABI Prism 7300 Sequence Detection System instrument and software (Applied Biosystems, Courtaboeuf, France). Gene-specific primers for SYBR Green assays are described in Supplementary Table 1 for *Hco-pgp-13*, pig mdr1 (*Ssc-pgp*) (GenBank Accession number: XM_003130205.3) and pig GAPDH that was used as a housekeeping gene for normalisation (GenBank Accession number: NM_001206359.1). All primers were designed using Primer Express software version

2.0 (Applied Biosystems) and synthesized by Invitrogen (Cergy- Pontoise, France). All primers were entered into the NCBI Blast program to ensure specificity. Results were expressed using the comparative Ct method as described in User Bulletin 2 (AppliedBiosystem). Briefly, the ΔCt values were calculated in every sample for each gene of interest as following:

$$\Delta Ct = Ct_{\text{gene of interest}} - Ct_{\text{reference gene}},$$

with GAPDH as the reference gene. The fold change in the level of target mRNA between treated and untreated samples was then expressed as $2^{-\Delta Ct}$ with ΔCt ± S.D. where S.D.is the standard deviation of the mean of the ΔCt value for the three replicates. A dissociation curve allowed us to verify the specificity of the amplification.

3. Characterization of protein expression

Membranes were prepared from transfected cells by collecting the cell culture on ice in a lysis buffer composed of PBS, 1 M DTT and protease inhibitors. The cell extract was then centrifuged at 1000 g at 4°C for 10 min, and the supernatant was discarded to remove debris. The pellet was resuspended in a volume 5 times its weight, and after 15 min of incubation on ice, cells were lysed with a potter and centrifuged at 1000 g at 4°C for 10min. This was done twice to discard the pellet containing debris and collect the supernatant containing the cell homogenate. The supernatant was subjected to an ultra-centrifugation at 100000 g for 40 min at 4°C. Pellets containing membranes were collected in around 200 µL lysis buffer. The quantity of proteins present in the samples was determined by the Bradford assay (Bio-Rad), with bovine serum albumin (BSA) as the standard, and absorbance at 595 nm was read with a spectrometer (Infinite® 200 Pro, Tecan). An aliquot of 10 µg of membranes was then run on a 10% sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE) in a Mini Protean® 3 Bio-Rad cell with the Laemmli method (Gallagher, 2006). The gel was then transferred to a nitrocellulose membrane that was blocked in PBS, 0.05% Tween, 5% milk, at 4°C overnight. The membrane was washed three times for 5 min with PBS-Tween and incubated with the primary antibodies: either 6x-His epitope tag monoclonal mouse antibody (MA121315, Fisher Scientific, France) or rabbit polyclonal antibodies designed against two epitopes of Hco-Pgp-13 (GenScript, USA) (as described in the manuscript in Part II, A) at 1/1000 dilution at 4°C overnight. The next day, the membrane was washed similarly and incubated with a secondary CF^{TM}770 anti-rabbit antibody (Biotium) at RT for 1 hour. The membrane was again washed three times for 5min with PBS-Tween and proteins were visualized using the Odyssey® CLx imaging system (LI-COR Biosciences, USA).

4. Transport assays

The transport activity of Hco-Pgp13 was studied by measuring in the transfected LLCPK1 the accumulation over time of rhodamine 123 (RHO) or doxorubicin (DXR), two fluorescent substrates of mammalian Pgps commonly used for cell transport assay (Lespine et al., 2006b). This was performed in the presence or absence of valspodar (VSP), a potent inhibitor of some ABC transporters, especially evidenced on mammalian Pgp. For these experiments, LLCPK1 cells expressing the protein of interest were cultured in 96-wells plates in supplemented 199 medium without G418, the positive control being a LLCPK1 cell line expressing a murine Pgp (LLCPK1-MDR1); whereas cells transfected with the *cat* gene and untransfected LLCPK1 cells were both used as negative controls. At time t=0, 10 µM RHO or 10 µM DXR in HBSS/DMSO 50/50 v/v were added to the medium at 37°C, with or without 20 µM VSP, expected to assess the maximum inhibition of the transporter. Within a time range of 30 min, the cells were then washed in PBS and lysed in PBS 0.5% SDS. Intracellular fluorescence of Rho123 or DXR was measured in the cell lysates with a spectrofluorometer (Infinite® 200 Pro, Tecan) with excitation and emission wavelengths of 507 nm and 529 nm respectively for RHO, or 470 nm and 600 nm for DXR. Fluorescence emission intensity was normalized to the protein content of cells measured by the Bradford method, using bovine serum albumin as the protein standard (Bio-Rad Protein Assay). For each parameter, mean and standard error (S.E.M.) were calculated.

5. Cytotoxicity assays

Cell viability of the stable transformants were assayed with the CellTiter96® Aqueous One solution Cell Proliferation Assay (MTS) (Promega, France), that contains a tetrazolium compound: [3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium, inner salt; MTS] and an electron coupling reagent (phenazine ethosulfate; PES). The MTS tetrazolium compound is reduced by a mitochondrial enzyme and produces formazan in metabolically active cells, which production can be quantified by measurement of absorbance at 490 nm. Cells transfected with *Hco-pgp-13* or *cat* gene were grown in 96-wells plates at 37°C, 5% CO_2 in 200 µL medium containing 0.005 µM to 0.5 µM ACD or no ACD. All ACD solutions were prepared to obtain 1% DMSO final, as well as in the control medium without ACD. The cultures were started from 5x 10^4 cells and 4x 10^4 cells, respectively, in each well, until the cells reach about 90% confluency. 48 hours later, MTS was added for 3 hours at 37°C. Reading of absorbance at 490 nm was performed with a spectrophotometer (Infinite® 200 Pro,

Tecan). This allowed the determination of the percentage of living cells. The *cat*-transfected LLCPK1 cells were used as negative control cells.

6. ATPase activity assays

The ATPase activity assay was performed as previously detailed in the manuscript in Part II, A.

7. *In silico* docking of various molecules on the two Hco-Pgp-13 3D structural models

The *in silico* docking calculations of paclitaxel (PCT), verapamil (VRP), rhodamine 123 (RHO), doxorubicin (DXR), valspodar (VSP), cholesterol (CLS), ivermectin (IVM) and moxidectin (MOX) on Hco-Pgp-13 were performed as described in detail in the manuscript in Part II, A.

II. RESULTS AND DISCUSSION

1. Expression of Hco-Pgp-13 in LLCPK1

Pig cells transfected with codon optimized *Hco-pgp-13* cDNA showed a high *Hco-pgp-13* mRNA expression level, reflected in the more than three orders of magnitude higher of *Hco-pgp-13* transcript compared with endogenous pig mdr1 (Fig. 1A).

The Western blot of the cytosol and membrane extracts from the Hco-Pgp-13 transfected LLCPK1 cells, following SDS–PAGE and incubation with the specific anti-His tag antibody (Fig. 1B), showed a band of around 130 kDa in the membrane extract. This corresponds to the expected form of the ABC transporter without post-translational modification, thus a MW smaller than Hsa-Pgp that is known to be glycosylated (about 30 kDa of glycosylation residues).

We also used this expression system to check the specificity of our antibodies designed against two epitopes of Hco-Pgp-13. Cells transfected with Hco-Pgp-2 showed no signal with these antibodies (Fig. 2A), whereas the antibodies specific for Hco-Pgp-2 did not show any signal when incubated with membranes of LLCPK1 cells transfected with Hco-Pgp-13 (Fig. 2B). Each of these antibodies thus appears specific for the protein they were raised against.

2. Functional characterization of Hco-Pgp-13 expressed in LLCPK1 cells

LLCPK1 cells were then used to characterize the function of Hco-Pgp-13 with accumulation assays of intracellular rhodamine 123 (RHO) and doxorubicin (DXR) in the presence or absence of valspodar (VSP), an inhibitor of mammalian Pgp-mediated drug transport. The results are presented in Fig. 3 as accumulation of RHO or DXR fluorescence / mg of proteins. The accumulation kinetics showed a similar pattern in untransfected and transfected cells. The intracellular accumulation of fluorescent dyes depends both on passive transport across the cell membrane and on transport by endogenous influx and efflux proteins, as well as by Hco-Pgp-13 in cells transfected with its encoding gene. Thus, only the level of accumulation reached at the steady state, resulting from all these transport phenomena, were compared. Without VSP, cells expressing Hco-Pgp-13 showed a lower level of accumulation of RHO at the steady state than untransfected cells (or cells transfected with the control gene coding for the chloramphenicol acetyl transferase (CAT), data not shown) (Fig. 3A and 3B). In the presence of VSP, the steady state accumulation level of RHO was increased (about 4 times), as compared to that in the absence of VSP, only in cells expressing Hco-Pgp-13 at passage 6, but not in untransfected (or CAT-expressing cells, data not shown). Interestingly, the accumulation kinetics in the presence of VSP is very similar to the one observed on the untransfected cells. Similar results were obtained with the fluorescent drug doxorubicin (DXR) (Fig. 3C and 3D). The steady state accumulation level of DXR when VSP was added to LLCPK1 cells expressing Hco-Pgp-13 was also increased relative to that in the absence of VSP (about 1.5 times), but less than what was observed for RHO. Altogether, these results suggest that RHO and DXR interact with and are transported by Hco-Pgp-13. In addition, VSP appears to interact with Hco-Pgp-13 and inhibits the transport of RHO and DXR mediated by this efflux protein.

Another strategy to test the function of Hco-Pgp-13, reflected by its capacity to efflux actinomycin D, is to determine if its expression can protect LLCPK1/HcoPgp13 transfected cells from ACD cytotoxic effect, as compared to the control cells transfected with the *cat* gene. Cell viability assays were thus also performed on LLCPK1 cells in the presence of actinomycin D at various concentrations. No increase of cell viability was observed in two different clones at passage 8 and 13 of cells transfected with *Hco-pgp-13* relative to *cat*-transfected cells (Fig. 4). This could be explained either by ACD not being a substrate of Hco-Pgp-13, or by a loss of the expression of Hco-Pgp-13 in LLCPK1 cells after several passages. Because the transport

assays of RHO also showed a decreasing effect of VSP on the transport of fluorescent dyes after several passages (data not shown), the second hypothesis was further investigated.

3. Passage-dependent expression and function of Hco-Pgp-13 in LLCPK1 cells

To better determine the effect of passages on Hco-Pgp-13 expression, a Western-blot was performed with anti-Hco-Pgp-13 antibodies following SDS-PAGE of membranes extracted from LLCPK1 cells transfected with *Hco-pgp-13* at passage 5 and passage 9. The result presented Fig. 5 showed a band of around 130 kDa in the membrane extract of passage 5, but not of passage 9, indicating that Hco-Pgp-13 is not expressed in the membrane of LLCPK1 at passage 9. This is consistent with the absence of observed function of Hco-Pgp-13 in LLCPK1 after several passages. This loss of expression of Hco-Pgp-13 at the membrane of LLCPK1 after passages can be explained by a decrease of the production of the protein or of its targeting to the cell membrane along passages.

Here we found that RHO could be transported by Hco-Pgp-13 and that VSP could inhibit this transport, as previously shown for Hco-Pgp-2, 9.1 and 16 (Godoy et al., 2015a, 2016; Godoy et al., 2015b). However, the LLCPK1 expression system proved to be limited for the study of the transport function of Hco-Pgp-13 and of the protection from the cytotoxic effect of drugs by the ABC transporter. Indeed, despite the codon optimization of the sequence of *Hco-pgp-13* for mammalian cells, a loss of its expression along passages was observed. Thus, we could not test the inhibition of the transport of fluorescent substrates by ML to investigate the interaction of Hco-Pgp-13 with these anthelmintic drugs.

Here, the pcDNA3.1 vector contains two independent promoters for the expression of *Hco-pgp-13* and of *neomycin*, so that this resistance gene can be expressed without the *Hco-pgp-13* gene simultaneously expressed. To overcome this limitation, a bicistronic pIRES vector could be used, which will lead to the transcription of *Hco-pgp-13* and of the resistance gene on the same mRNA, so that every resistant cell should express Hco-Pgp-13 protein. Also, a peptide signal could be added to ensure that the protein is expressed at the plasma membrane.

4. Functional characterization of Hco-Pgp-13 expressed in *Pichia pastoris*

Another possible strategy is a transfection in *P. pastoris* cells, which is known to yield high level expression of transgenes. The interaction between an ABC transporter and its putative substrates can then be studied by measuring the ATPase activity of the transporter in

presence of compounds of interest, as performed by Jin et al. (2012). This heterologous system was then chosen to express Hco-Pgp-13, as described in Part II, A.

Preliminary ATPase activity assays were performed with paclitaxel (PCT) on *P. pastoris* cells expressing Hco-Pgp-13. The results presented in Figure 6 indicated a dose-dependent stimulation of its basal ATPase activity by PCT. If reproducible, this would indicate that PCT can interact with Hco-Pgp-13.

Before more molecules were tested for their interaction with Hco-Pgp-13 on this experimental model, those that were the most potent to stimulate the ATPase activity of Cel-Pgp-1, which shows high homology with Hco-Pgp-13, were tested *in silico* for their affinity to Hco-Pgp-13. Other molecules of interests, such as valspodar and cholesterol, were also tested in order to screen the putative substrates of Hco-Pgp-13, before performing ATPase activity assays on *P. pastoris* cells expressing Hco-Pgp-13.

5. *In silico* docking calculations on Hco-Pgp-13 3D structural models

The *in silico* docking calculations of molecules of interest were run on the two 3D structural models of Hco-Pgp-13 constructed by homology with Cel-Pgp-1 4F4C. Molecules tested were 8 drugs known to be substrates of mammalian Pgp: paclitaxel (PCT), verapamil-S (VRP), doxorubicin (DXR) and valspodar (VSP), the fluorescent dye rhodamine 123 (RHO) (commonly used in mammalian Pgp transport assays), the endogenous substrate cholesterol (CLS), and finally the two ML mostly used in veterinary medicine, ivermectin (IVM) and moxidectin (MOX). Some of these tested molecules have been studied for *in vitro* experiments aimed at characterizing heterogenously expressed Hco-Pgp-13, among them ACD and PCT has been assayed for ATPase activity, and RHO, DXR and VSP for cell transport measurements.

a) Docking of substrates of mammalian ABCB1 transporter

For PCT, the clusters of the energy clustering histograms were either scattered, containing a few number of poses, between 1 and 4, for docking on Hco-Pgp-13_04, or poses were mainly grouped in two lowest energy clusters, for Hco-Pgp-13_52. On the model n°04, the lowest energy pose (PCT1) was found at -10.5 kcal/mol and the second cluster (PCT2, 2 poses) at -9.3 kcal/mol, whereas on the model n°52, the two lowest energy clusters were found at -11.50 kcal/mol (PCT1, 2 poses) and -11.48 kcal/mol (PCT2, 13 poses), (Figure 7A and 7B and Table 1). The binding location of PCT was similar for PCT1 on Hco-Pgp-13_04 and PCT2

on Hco-Pgp-13_52, in the deepest part of the inner pocket of the transporter (Figure 7C). PCT1 on Hco-Pgp-13_52 was found in the middle of the pocket, whereas PCT2 on Hco-Pgp-13_04 was found in its lower part, close to the cytoplasmic opening of the pocket, as most of other poses of this clustering. These binding poses interacting with 14 to 17 residues, including 11 to 14 hotspots and forms 2 H-bonds, except for PCT2 of model 04 that interacted with 4 hotspots and formed 2 H-bonds (Table 1 and Table 2). Consistent with their mostly overlapping binding site, 12 residues were found common between PCT1 of 04 and PCT2 of 52, including 10 hotspots, and the two lowest energy clusters formed a H-bond with Q108 (Table 2). Overall, PCT was found to be able to bind to various sites of Hco-Pgp-13, partly depending on the orientation of amino-acids lining the binding pocket. However, the 3 lowest energy binding were between -10.5 and -11.5 kcal/mol indicating a high affinity binding, significantly (>2 kcal/mol) higher to what was found for Cel-Pgp-1 (-8.2 kcal/mol for the lowest energy pose, manuscript under writing, Part I, A). Their positions were all found deep in the pocket of Hco-Pgp-13, in contrast to what was found on Cel-Pgp-1, where PCT hardly reached the deepest part of the pocket (PCT4, 1 pose at -6.5 kcal/mol, manuscript under writing, Part I, A).

For VRP, the energy clustering histograms presented scattered clusters of few poses, but for docking on both models, the localization of the first energy cluster (VRP1) matched that of other clusters found within 2 kcal/mol energy, with various orientations, so that only VRP1 was analyzed. VRP1 was found at -7.7 kcal/mol (2 poses) on the model n°04 and at -7.3 kcal/mol (10 poses) on the model n°52 (Figure 8A and 8B and Table 1). The binding location of VRP1 was similar on Hco-Pgp-13_04 and Hco-Pgp-13_52, burried within the TMDs of the transporter in a much folded conformation (Figure 8C). VRP thus bound to model 04 and 52 by interacting with 17 and 15 residues, including 13 and 10 hotspots and forming 2 and 1 H-bonds, respectively (Table 1 and Table 2). As its binding sites on both models mostly superimposed, 9 common residues were found, including 6 hotspots, although they formed H-bonds with different residues (Table 2). VRP was then found to bind to one major site of Hco-Pgp-13, regardless of the model considered, which increases the strength of this finding. This position is similar to what was found on Cel-Pgp-1, manuscript under writing, Part I, A). The binding energy found, around -7.5 kcal/mol, indicates a medium affinity of binding to Hco-Pgp-13, slightly higher to what was found for Cel-Pgp-1 (-6.5 kcal/mol energy bonding for VRP1, manuscript under writing, Part I, A).

For RHO, a major cluster was found in each clustering, either presenting the lowest binding energy (RHO1), with 76 poses at -3.9 kcal/mol on model n°04 of Hco-Pgp-13, or the

third lowest energy (RHO3), with 61 poses at -3.6 kcal/mol on model n°52 (Figure 9A and 9B and Table 1). The lowest energy cluster (RHO1) on this model was almost identical to RHO3 with 4 poses at -4.0 kcal/mol. These three clusters were found overlapping in the most inner part of the binding pocket of Hco-Pgp-13 (Figure 9C), interacting with 7 to 11 residues including 5 to 8 hotspots, and forming 1 or 2 H-bonds (Table 1 and Table 2). 10 residues were common to at least two of these poses, 5 being common to the three of them, 8 being hotspots and E21 of TMa-b forming a H-bond with the three poses (Table 2). This main binding position on Hco-Pgp-13 is similar to the 4th lowest energy cluster found on Cel-Pgp-1, although only one cluster was found at this location in Cel-Pgp-1 (manuscript under writing, Part I, A). On the two transporters, the almost identical binding energy of around -3.9 kcal/mol indicates a very low affinity of binding of RHO.

For DXR, the two clusterings presented one or several clusters containing the majority of poses. On the model n°04, the lowest energy cluster (DXR1, 3 poses) was found at -8.8 kcal/mol and the second cluster (DXR2, maximum of 25 poses) at -8.0 kcal/mol. On the model n°52, the lowest energy pose (DXR1) was found at -9.0 kcal/mol and DXR2 (2nd maximum of 17 poses) at -8.7 kcal/mol (Figure 10A and 10B and Table 1). The binding location of DXR on Hco-Pgp-13_04 was different for most of the clusters, from the lowest part of the pocket, next to its cytoplasmic opening, for DXR2, to its inner core with DXR3 (not shown, similar to DXR1 on model 52), and with most of clusters at various locations around the center of the pocket, such as DXR1 (Figure 10C). In contrast, all the clusters (in particular DXR3&4) found on model 52 were similar to either DXR1 or DXR2, at the inner core or in the middle of the protein binding pocket, respectively. Each of these poses interacted with 11 to 15 residues, including 3 to 9 hotspots and formed 2 to 5 H-bonds (Table 1 and Table 2). As these binding poses of DXR overlapped to various degrees, 6 interacting residues were found common between at least 3 of them, including 3 residues forming an H-bond with one or two of these poses, and E21 of TMa-b with the four of them (Table 2). Overall, DXR was thus able to bind to various sites of Hco-Pgp-13, with the highest affinity corresponding to the pose docked in the deepest part of the binding pocket. This was found for DXR1 on model n°52 at -9.0 kcal/mol, indicating a high affinity binding, similar to what was found for Cel-Pgp-1 (-9.3 kcal/mol for the lowest energy pose, also located in the deepest part of the pocket (manuscript under writing, Part I, A).

For VSP, the lowest energy cluster VSP1 (2 poses) found on Hco-Pgp-13_04 was at -10.5 kcal/mol and the two lowest energy clusters VSP1 (20 poses) and VSP2 (24 poses) were found at -8.2 kcal/mol and -6.2 kcal/mol, respectively, on Hco-Pgp-13_52 (Figure 11A and 11B

and Table 1). The binding location of these poses of VSP, were similar and representative of all other clusters on both models, in the lowest part, next to the cytoplasmic opening of the binding pocket (Figure 11C). They interacted with 13 to 22 residues, including 2 to 4 hotspots without forming any H-bond (Table 1 and Table 2). As these 3 binding poses were almost superimposed, only varying in their orientations, 13 interacting residues were shared by at least 2 of them, including 2 hotspots (Table 2). VSP then bound to a single location, with various possible orientations in the cytoplasmic opening of the binding pocket of Hco-Pgp-13, as on Cel-Pgp-1 (manuscript under writing, Part I, A). In contrast to other molecules, this binding site contained a few number of hotspots and VSP formed no H-bond. Its lowest binding energy of -10.5 kcal/mol however suggested a high affinity of binding, slightly higher to what was found for Cel-Pgp-1 (-9.5 kcal/mol for the lowest energy pose).

For CLS, a major cluster was found in each clustering, either presenting the lowest binding energy (CLS1, 58 poses) at -11.1 kcal/mol on Hco-Pgp-13_04, or the second lowest energy (CLS2, 80 poses) at -10.46 kcal/mol on Hco-Pgp-13_52 (Figure 12A and 12B and Table 1). The lowest energy cluster (CLS1, 4 poses) at -10.50 kcal/mol on Hco-Pgp-13_52 was similar to the two previously presented cluster as well as all other clusters found on the two models, only slightly varying in their orientations. They overlapped in the most buried part of the binding pocket of Hco-Pgp-13 (Figure 12C), interacting with 14 or 15 residues including 10 or 11 hotspots, and forming no H-bond, 1 or 2 H-bonds (Table 1 and Table 2). 13 interacting residues were common to at least two of these poses, including 10 hotspots, and H-bonds were formed with very close residues of TMa-b. This very specific binding position on Hco-Pgp-13 contrasts with the predicted binding of CLS on Cel-Pgp-1 on various possible sites. CLS lowest energy cluster was also found in the inner core of Cel-Pgp-1, but with a binding energy (-9.4 kcal/mol) indicating a lower affinity than the high one found for CLS on Hco-Pgp-13.

Overall, these 6 molecules delineated a binding pocket showing a different shape in the two models of Hco-Pgp-13 (Fig. 13). The binding pocket of the model n°52 was very similar to the pocket previously described in Cel-Pgp-1 as containing two sub-binding domains. In contrast, the binding pocket of Hco-Pgp-13_04 looked more continuous from its cytosolic opening to its inner core, with the part corresponding to the narrow site between the two sub-domains of Cel-Pgp-1 being only constrained on one side.

When comparing the binding sites of the molecules between the two models, it appeared that VRP, CLS, RHO, ACD and VSP binding were similar whereas PCT and DXR binding locations were more model-dependent. However, these two molecules also bound to various

sites within each of the two models, so that the two models looked rather complementary in the higher number of possibilities they offer for molecules to bind due to the various positions of lateral chains of amino-acids lining the inner chamber.

b) Docking of the ML ivermectin and moxidectin

The two other molecules tested for their binding to Hco-Pgp-13 were the ML ivermectin (IVM) and moxidectin (MOX), which are of particular interest in the perspective of potential anthelmintic drug resistance phenomena in parasite nematodes.

For IVM, the docking calculations, performed with two starting conformations on each model of Hco-Pgp-13, showed very close lowest energy clusters presenting different binding sites. They were thus all analyzed. On the model n°04, the lowest energy cluster (IVM1, 19 poses) of one starting conformation was found at -11.2 kcal/mol whereas the other lowest energy cluster (IVM1', 28 poses) obtained with a different starting conformation was at -11.3 kcal/mol. On the model n°52, the two lowest energy clusters IVM1 (17 poses) and IVM1' (8 poses) were both found at -12.8 kcal/mol (Figure 14A, B, C, D and Table 3). The binding location of IVM1 and IVM1' on Hco-Pgp-13_04 was similarly close to the cytoplasmic opening of the pocket, with different orientations, whereas the binding site of IVM1 and IVM1' on Hco-Pgp-13_52 was in the deepest part of the inner pocket of the transporter, also with two possible orientations (Figure 14E). These binding poses interacted with 14 to 17 residues on model n°04, and 21 residues on model n°52, including 4 or 5 hotspots and forming 2 to 3 H-bonds on Hco-Pgp-13_04, and 14 or 15 hotspots and forming no or 2 H-bonds on Hco-Pgp-13_52 (Table 3 and Table 4). As expected from their binding locations, 9 common interacting residues were found between IVM1 and IVM1' of model n°04, including 3 hotspots, while 16 interacting residues were shared by IVM1 and IVM1' of model n°52, including 11 hotspots (Table 4). The binding locations found on the two models were almost separated, with only 5 residues found common between the positions found in the different models. The binding site of IVM on Hco-Pgp-13 was thus mostly depending on the orientation of amino-acids lining the binding pocket. This led to a lowest binding energy varying from -12.8 to -11.2 kcal/mol between the two models, and indicating high affinity binding in both cases. Interestingly, these binding energies were close to those found for IVM1 and IVM2 on Cel-Pgp-1 (-12.2 kcal/mol and -11.1 kcal/mol, respectively, manuscript under writing, Part I, A). Plus, the docking position of IVM1 and IVM1' on Hco-Pgp-13_52, both showing the lowest energy binding found across the two models, were occupying the same space as IVM1 on Cel-Pgp-1. Their macrocycle ring on Hco-

Pgp-13_52 was overlapping with the dissacharide moiety of IVM1 on Cel-Pgp-1, and vice versa, so that the macrocycle of IVM was found even more buried within the TMDs of Hco-Pgp-13_52 than of Cel-Pgp-1. Plus, IVM1 on model n°52 formed a H-bond with S1028 of Hco-Pgp-13, aligned with the residue T1028 of Cel-Pgp-1 that formed a H-bond with all ML tested. The two poses observed on Hco-Pgp-13_04 were, in contrast, very different from any low energy binding site found on Cel-Pgp-1 (manuscript under writing, Part I, A).

For MOX, the result of one starting conformation was chosen for docking on Hco-Pgp-13_04, whereas two different starting conformations were analyzed after docking on Hco-Pgp-13_52. MOX1 (90 poses) was found at -10.8 kcal/mol on model n°04 and MOX3 (3 poses) was found at -10.0 kcal/mol binding energy on this model. In contrast, MOX1 (90 poses) was found at -11.7 kcal/mol on model n°52 and MOX1' (65 poses) at -11.6 kcal/mol (Figure 15A, B, C and Table 3). The binding location of MOX1 on Hco-Pgp-13_04 was close to the cytoplasmic entrance of the binding pocket, whereas those of MOX3 on Hco-Pgp-13_04 and MOX1' on Hco-Pgp-13_52 was in the middle of the pocket, and finally, MOX1 on Hco-Pgp-13_52 bound to the deepest part of the inner pocket (Figure 15D). They interacted with 11 to 15 residues, including 3 to 12 hotspots, and forming no H-bond, except for MOX1 forming one on model n°04 (Table 4). MOX1 and MOX1' on model n°52 shared 8 common interacting residues, whereas MOX3 on model n°04 shared 3 interacting residues with each of them, and 2 with MOX1 on model n°04. As for IVM, the binding position of MOX on Hco-Pgp-13 depended on the model considered and thus on the orientation of lateral chains of residues lining the inner pocket. MOX lowest binding energy of -11.7 or -10.8 kcal/mol, depending on the model considered, revealed a high affinity for Hco-Pgp-13 whereas MOX1 was found at -10.5 kcal/mol on Cel-Pgp-1, indicating a putative higher affinity of MOX for Hco-Pgp-13 than for Cel-Pgp-1 (manuscript under writing, Part I, A). MOX was found in the deepest core of the inner pocket of Cel-Pgp-1, similarly to what was found for MOX3 on the model n°04, with inverted locations of the macrocycle ring and spiroketal moiety. MOX1 and MOX1' on the model n°52 of Hco-Pgp-13 had a location of their macrocycle ring close to that of MOX3 on model n°4, but with their spiroketal moieties even more buried within the inner core of the inner pocket of Hco-Pgp-13.

It is interesting to note that the H-bonds found for all ML on T1028 of Cel-Pgp-1 have no counterpart on Hco-Pgp-13, on which most of ML poses formed no H-bond, except for IVM1 on model n°52 of Hco-Pgp-13 that formed a H-bond with S1028 aligned with T1028 of Cel-Pgp-1. On the other hand, E22 that formed a H-bond with IVA2 and MOX2 on Cel-pgp-1,

and also interacted with IVM1, is aligned in Hco-Pgp-13 to E21 that formed a H-bond with IVM1 on model n°04. Interestingly, other similarities appeared when comparing the docking of IVM and MOX on Hco-Pgp-13 and on Cel-Pgp-1. First, as for the clustering profiles obtained on Cel-Pgp-1, those of MOX presented less scattered clusters than those of IVM on Hco-Pgp-13, whatever model considered. Thus, the absence of disaccharide also appeared to lower the possibilities of binding to different sites in the binding pocket of Hco-Pgp-13. Plus, whatever model considered, the difference of binding energy between IVM and MOX was of about 1kcal/mol, as previously observed on Cel-Pgp-1.

When comparing the binding sites of IVM and MOX in each of the two models, a striking difference appeared due to the possible location of IVM and MOX in the middle of the TMDs in model n°04 of Hco-Pgp-13. This was not observed in model n°52 and in Cel-Pgp-1, on which these poses correspond to an intermediate location between BD1 and BD2 (Fig. 16A and B). However, as observed in Cel-Pgp-1, some poses of IVM and MOX showed good alignments of the macrocycle ring. This was found for MOX1 and IVM1 on model n°04, with moieties at different locations around the macrocycle (Fig. 16C, left panel), for MOX1' and IVM1' on model n°52 (Fig. 16C, right panel), with almost superimposed molecules, but also for MOX3 on model n°04 and IVM1 on model n°52 (Fig. 16C, middle panel), with opposite orientations of the superimposed benzofurane moieties, showing that the two models are not totally different.

c) Overall binding site of Hco-Pgp-13 according to the two models

To better understand the difference observed between the binding sites obtained with each of the two models of Hco-Pgp-13, all molecules tested were superimposed for model n°04 alone (Fig. 17A), model n°04 together with model n°52 (Fig. 17B), or model n°52 alone (Fig. 17C), and compared to the overall binding site previously found for Cel-Pgp-1 (Fig. 17D). As expected, the overall binding pocket of model n°52 alone of Hco-Pgp-13 resembles that of Cel-Pgp-1 more than the binding pocket of model n°04, due in particular to the similar narrow site between the two major binding domains. However, when looking at the superimposed binding pockets of Hco-Pgp-13_04 and Hco-Pgp-13_52, most of the binding locations of the various molecules docked were comprised in the same area. This indicated some complementarity between the two models, with different protrusions observed in each of them forming a global shape resembling that of Cel-Pgp-1.

The only binding sites found in the Hco-Pgp-13_04 model that appear to be not compatible with Hco-Pgp-13_52 are DXR2, MOX1 and IVM1 that occupy the narrow site of Hco-Pgp-13_52, corresponding to the part of Cel-Pgp-1 that separates the two sub-domains. To better compare this site between the three 3D structures, the binding poses of DXR2, MOX1 and IVM1 on model n°04, together with ACD1 on model n°04 and n°52, were compared to the location of ACD1 and ACD2 on Cel-Pgp-1 (Fig. 18). The site that joins the two areas corresponding to BD1 and BD2 in Cel-Pgp-1 is mainly occupied by two elongated structures formed of aligned molecules in Hco-Pgp-13_04 (Fig. 18A). Interestingly, ACD1 of model n°52 also aligned with one of these structures, showing that the two models are not contradictory. In contrast, none of them was aligned with the center of ACD molecules forming the corresponding site in Cel-Pgp-1 binding domain (Fig. 18B and C), showing a major difference between Hco-Pgp-13 and Cel-Pgp-1. Thus, Hco-Pgp-13 appears to be composed, like Cel-Pgp-1, of two major binding sub-domains BD1 and BD2, linked by a site oriented differently than in Cel-Pgp-1, and either very narrow as in the *C. elegans* transporter, or wider, depending on the orientation of the sides chains of residues lining the inner pocket. ML could then bind to this intermediate area when Hco-Pgp-13 amino-acids allow it, by "opening" the middle area of the binding pocket, during which BD2 might be narrow and might not be able to fit ML, as only one MOX position was found there on model n°04. Then, the conformation of residues that form a more narrow site between BD1 and BD2 could allow a wider binding site opening in BD2 and thus fit IVM and MOX within this site with a higher affinity, so that their major binding site could be expected to have the location found on Hco-Pgp-13_52.

Of course, it is necessary to keep in mind that these docking calculations were performed using two 3D structures that were constructed by homology with the crystal of a close Pgp homolog, Cel-Pgp-1 (PDB 4F4C), sharing 54% similarity for AAs present in the TMDs. The N-terminal helix was thus designed by homology to form a hairpin within the TMDs, with many interacting residues found on this helix. In particular, E21 was found to form a H-bond with most of substrates tested, showing once again the critical role of this hairpin when present in a Pgp, as it reaches the lower part of BD2. It could be wondered whether its absence would have changed much the binding profile of molecules tested, but the closer crystal structure of a Pgp not showing this helix is the mouse Pgp structure, which has been much controversed. We then chose to perform our homology modelling with a single, never questioned, crystal structure, rather than a bi-template homology modelling that can be more precise only if the two templates are themselves sure to be accurate. This docking is then less accurate than one performed

directly with a crystallized structure, so that the error bar of binding energies can be expected to be higher than the ones found for docking on Cel-Pgp-1 (about 0.25 kcal/mol), as well as those of the number of poses in a cluster (within ≈10% on Cel-Pgp-1, manuscript under writing, Part I, A). Thus, binding energies found for a same molecule on Hco-Pgp-13 and Cel-Pgp-1 must be compared with caution.

On the other hand, this docking strategy with two models allows a vision somehow flexible of the transporter, with two putative conformations of Hco-Pgp-13 compatible with each other in the dynamic process of ligand binding and translocation. This revealed more possible binding sites of IVM and MOX on Hco-Pgp-13 than what was found on Cel-Pgp-1, and we cannot rule out that a flexible strategy for docking on Cel-Pgp-1 would also have revealed supplementary binding sites. However, the lowest energy ones of IVM and MOX were found on the model n°52 of Hco-Pgp-13, with a general binding pocket more resembling that of Cel-Pgp-1 than the pocket of Hco-Pgp-04, so that the binding poses found for IVM and MOX on Cel-Pgp-1 still appear as the most accurate possible after this study. On the other hand, some other molecules bound with a higher affinity to Hco-Pgp-13_04 than 52, as its BD1 appeared more "closed" and interacted thus more closely with smaller compounds. This shows that the multispecific binding of Pgps may be possibly also partly explained by their flexibility, which was not taken into account in the strategy used on Cel-Pgp-1. Although the highest difference of energy binding between the two models was of 2.2 kcal/mol for VSP, it was found to bind in the same area on the other model. The lowest energy binding of other molecules only varied in a range of 0.2 to 1 kcal/mol between the two models, always with close to superimposed locations of binding within the pocket of Hco-Pgp-13. This shows that a flexible docking strategy would be more precise in terms of location of binding sites, but would not change much the binding energy or the location of binding of molecules in the pocket of a transporter. Plus, flexible docking is limited to some residues in the protein that must be chosen as the most relevant for interaction with ligands. This can more easily be done once preliminary docking calculations have been performed with a rigid transporter. It would thus be interesting to perform more docking calculations with an Hco-Pgp-13 protein flexible on residues such as E21 or other residues found to interact with most of molecules. However, this would take a huge amount of time to gain a few precision in the result, whereas *in vitro* experiments with molecules found to interact with high affinity to Hco-Pgp-13 would expectedly bring much more functional information than previously gained with this *in silico* docking.

d) Confrontation with *in vitro* data

These *in silico* calculations indicated a very high affinity of binding for PCT, VSP and CLS, as well as for ACD (manuscript under writing, Part I, A) on Hco-Pgp-13 with binding energies below -10 kcal/mol. It should be noticed that, among them, CLS is the only "small ligand", which likely gives it a still more marked significance (see below). This is consistent with our *in vitro* experiments showing a stimulation of ATPase activity of Hco-Pgp-13 by ACD (manuscript under writing, Part I, A) and PCT. In addition, these experiments indicated an apparent affinity in the µM range for both of them, in consistence with their respective calculated Eb (more negative for ACD) and membrane partition (higher for PCT). VRP and DXR bound with energies below -7 kcal/mol, compatible with a rather high affinity of these substrates on Hco-Pgp-13. This is in line with the transport of DXR observed on LLCPK1/*Hco-pgp-13* cells. The *in silico* data are further consistent with the finding that VSP (20 µM) inhibited DXR transport (tested at 10 µM) in these *in vitro* experiments, presumably by a competitive mechanism since they share partially overlapping binding sites in the inner chamber of Hco-Pgp-13. Obviously, for confirmation, specific interaction of DXR and VSP with Hco-Pgp-13 should be further tested *in vitro* by assaying their effects on its ATPase activity.

In contrast, RHO binding energy on Hco-Pgp-13 was calculated above -4 kcal/mol, which suggests a very low affinity of binding on the transporter, or could even be considered as non-specific binding. However, functional assays of RHO accumulation in LLCPK1/*Hco-pgp-13* cells indicated a RHO transport by Hco-Pgp-13, revealed by a high increase of RHO accumulation in presence of VSP (even higher than for DXR). It could nevertheless be possible to reconcile this apparent discrepancy, by taking into account the possibility that Hco-Pgp-13 could be a membrane lipid translocase, especially handling cholesterol or other sterols. Indeed, ATPase activity measurements have shown that Hco-Pgp-13 displays a "basal" activity in the absence of any exogenous added compound. This could mean that it can actively handle some endogenous membrane components. The very negative Eb value, especially taking into account its small size, calculated for CLS docking suggests that such endogenous substrate could be cholesterol, or at least any structurally-related sterol present in the *P. pastoris* cell membrane. Of course, this hypothesis should be easily tested by adding methyl-beta-cyclodextrin to the Hco-Pgp-13-containing membranes in order to specifically manipulate their sterol content and follow the repercussion on its ATPase activity.

Within this hypothesis of Hco-Pgp-13 being an active cholesterol translocase, the composition and/or the distribution of cholesterol in the plasma membrane of LLCPK1 cells should be altered in the presence of Hco-Pgp-13. It could then also modify the membrane partition and passive diffusion of the amphiphilic dye RHO in LLCPK1 cells, as shown for other types of membrane perturbation (Schulz et al., 2013). As transport accumulation assays are the result of kinetics depending on passive transport and active transport by various proteins, it does not directly accounts for the affinity of RHO towards Hco-Pgp-13 observed *in silico* in docking calculations. To go further, measurements of the ATPase activity stimulation by RHO would be required to better analyze its binding capacity on Hco-Pgp-13. The same type of *in vitro* studies performed with other ligands tested in docking experiments would represent a very informative combination of tools to better understand the molecular mechanism of ligand binding on, and putative transport by, Hco-Pgp-13.

As a final conclusion, based on the whole of our data combining various approaches, Hco-Pgp-13 appears as a multidrug efflux transporter (6 high affinity ligands, 2 medium affinity ligands and only 1 low affinity to non-ligand out of 9 *in silico* tested molecules), especially handling MLs, but it seems also likely involved in membrane lipids, at least sterols, translocation. This would make it a fair functional homolog of mammalian Pgp.

Figures and Tables

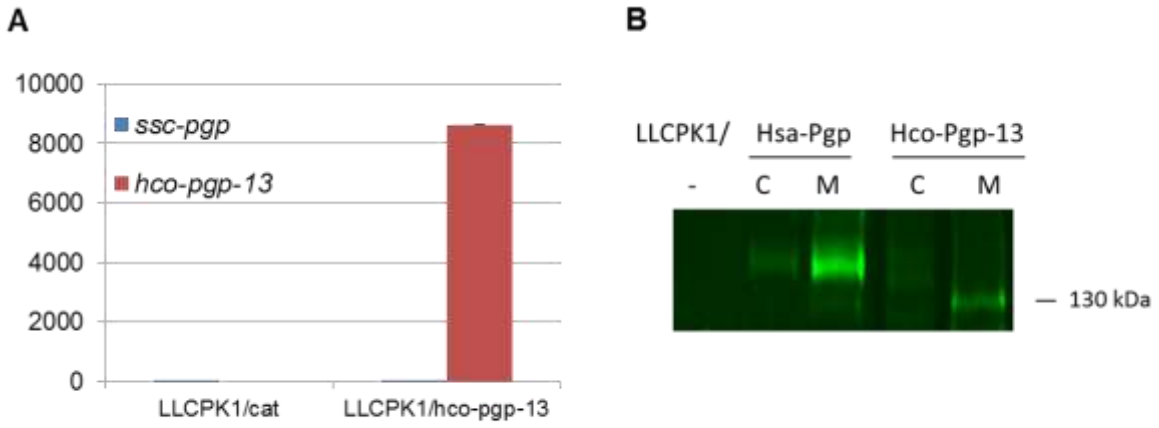


Figure 1. Expression level of Hco-pgp-13 mRNA and protein in LLCPK1 transfected cells. **A.** Quantitative Real-Time PCR on codon optimized *Haemonchus contortus* P-glycoprotein 13 transfected in pig kidney epithelial cells (LLCPK1/*Hco-pgp-13*), at passage 5 and in LLCPK1/*cat* cells transfected with the gene coding for the chloramphenicol acetyltransferase. Relative expression over the pig P-glycoprotein (*Ssc-Pgp*) endogenous gene was calculated by the $2^{-\Delta Ct}$ method. Mean of three replicate experiments ± S.D. are presented. **B.** Western-blot using antibody anti-His tag after SDS-PAGE with cytoplasmic (C) or membrane (M) extracts from LLCPK1 untransfected cells (lane 1), or transfected with *Hco-pgp-13* at passage 5 (lane 4 and 5), or C219 antibody recognizing mammalian Pgp on LLCPK1 cells transfected with human (Hsa-) Pgp (lane 2 and 3). Each lane was loaded with 10µg of protein.

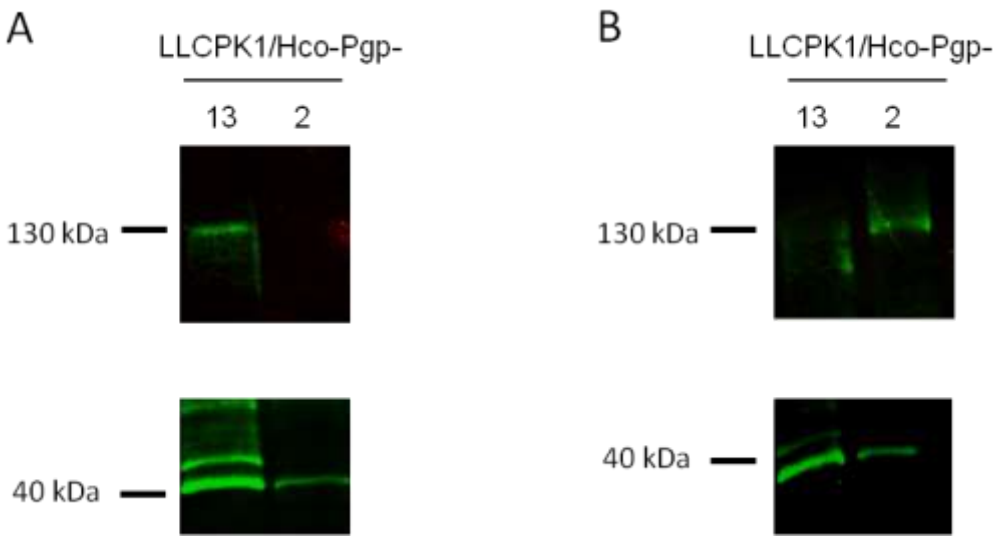


Figure 2. Specificity of Hco-Pgp-13 antibodies. A. Western-blot using antibody anti- epitope 1 and 2 of Hco-Pgp-13 after SDS-PAGE with membrane extracts from epithelial-like pig kidney LLCPK1 cells transfected with *Hco-pgp-13* (lane 1) or *Hco-pgp-2* (lane 2). **B.** Western-blot using antibody anti- Hco-Pgp-2 on membrane extracts from LLCPK1 cells transfected with *Hco-pgp-13* (lane 1) or *Hco-Pgp-2* (lane 2). Anti-human actin antibody was used as a loading control in lower panels. Each lane was loaded with 10µg of protein.

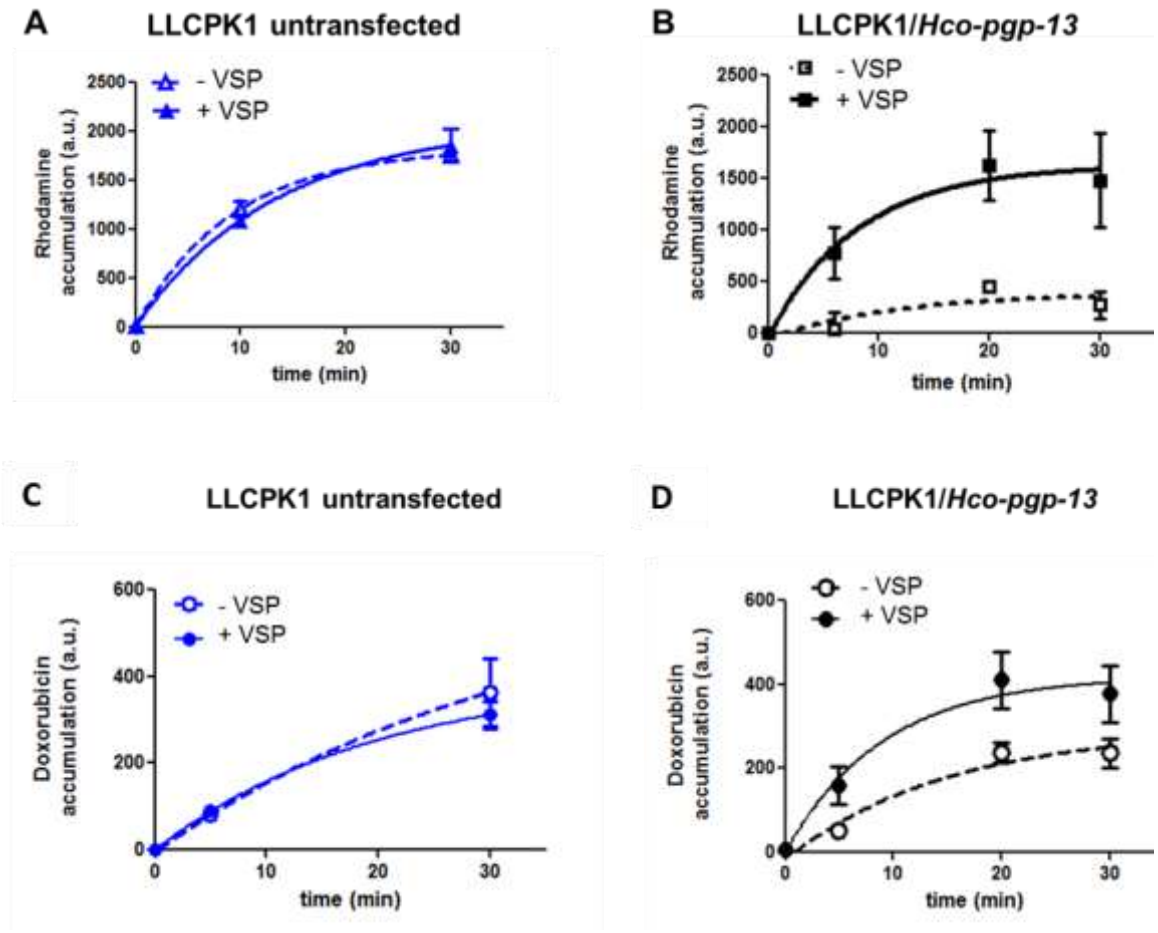


Figure 3. Functional transport assay of rhodamine 123 and doxorubicin fluorophore probes. Fluorescence accumulation is measured over time post-incubation of rhodamine 123 (**A, B**) or doxorubicin (**C, D**), in the absence or presence of valspodar (VSP). **A. C**. Untransfected epithelial-like pig kidney cells (LLCPK1). **B. D**. *Haemonchus contortus* P-glycoprotein 13 (*Hco-pgp-13*) transfectants at passage 6. Results are means of three replicate experiments ± S.E.M.

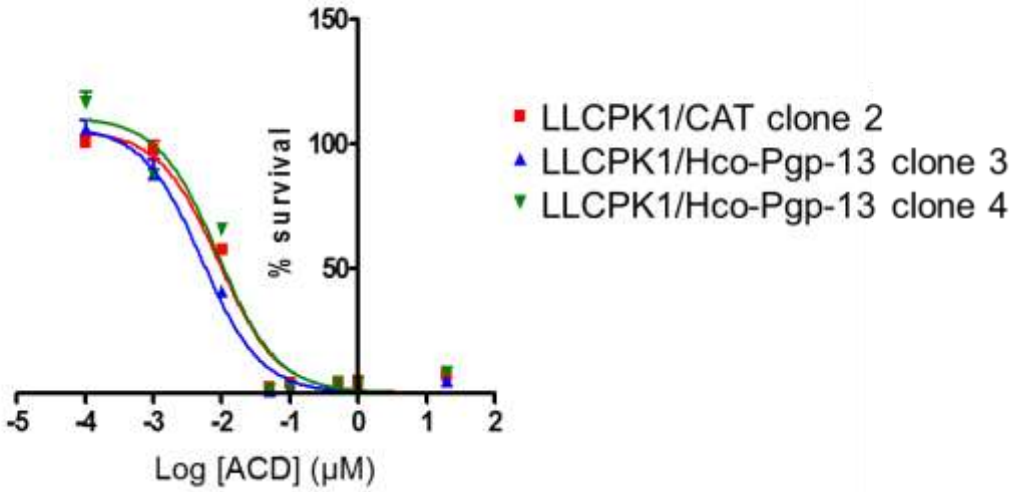


Figure 4. Cytotoxicity effect of actinomycin D (ACD) on pig kidney epithelial LLCPK1 cells transfected with *Hco-pgp-13* or *cat*. Cell survival percentage is expressed as compared to the 100% survival at the lowest concentration of ACD for each cell line. Two clones of cells transfected with *Hco-pgp-13* were tested, clone 3 at passage 8 and clone 4 at passage 13, and compared to cells transfected with the *cat* gene, coding for the chloramphenicol acetyl transferase.

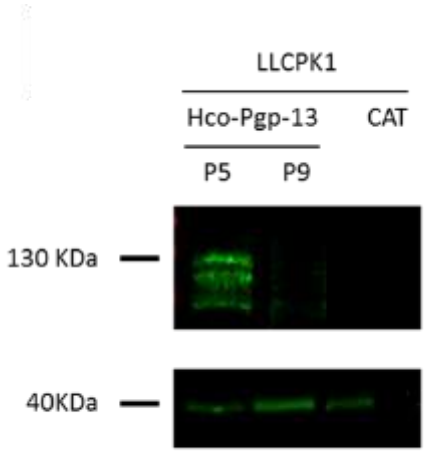


Figure 5. Effect of passages on the expression of Hco-Pgp-13 in pig kidney epithelial LLCPK1 cells. Western-blot using antibodies anti- epitope 1 and 2 of Hco-Pgp-13 after SDS-PAGE with membrane extracts from LLCPK1 cells transfected with *Hco-pgp-13* at passage 5 or passage 9, or on membrane extracts from LLCPK1 cells transfected with *cat* gene. Anti-human actin antibody was used as a loading control in lower panels. Each lane was loaded with 10µg of protein.

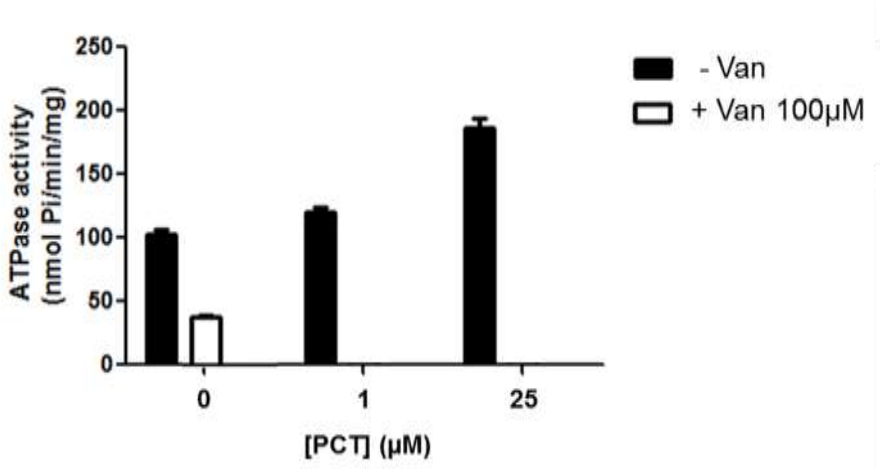


Figure 6. Stimulation of the ATPase activity of Hco-Pgp-13 expressed in *Pichia pastoris* membranes by paclitaxel. The ATPase activity is represented in nmol of Pi /min /mg of protein in *P. pastoris* membranes expressing Hco-Pgp-13, relative to the concentration of Paclitaxel in µM, in the absence (black bars), or in the presence of Vanadate (Van) 100 µM (white bars). Error bars have been calculated on triplicates within one experiment. No effect of PCT was observed on the basal ATPase activity of untransfected *P. pastoris* cells (data not shown).

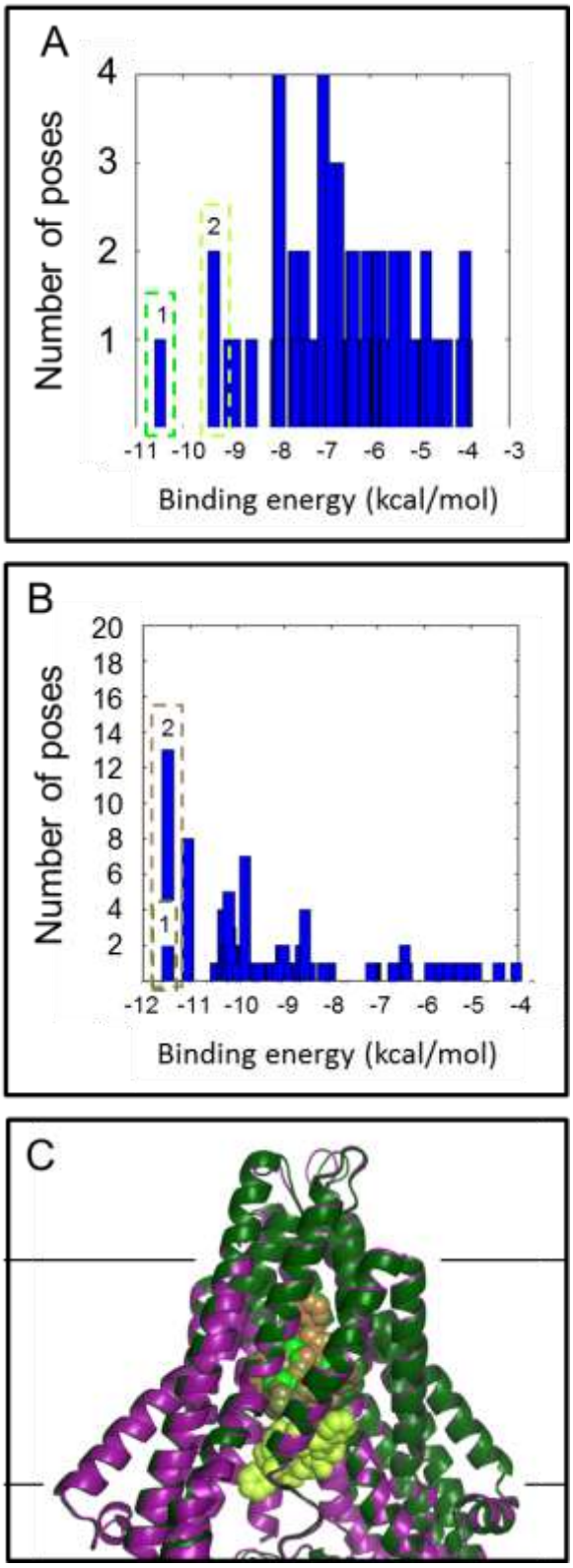


Figure 7. Paclitaxel binding to Hco-Pgp-13_04 (A, C) and Hco-Pgp-13_52 (B, C). A, B, energy clusterings of PCT docked to Hco-Pgp-13_04 (A) and Hco-Pgp-13_52 (B). C, binding sites of the 1st and 2nd lowest energy clusters of PCT on each protein. PCT1 and PCT2 are represented in bright green and light green spheres respectively for docking on Hco-Pgp-13_04 (purple ribbon) and in dark green and light brown spheres respectively for docking on Hco-Pgp-13_52 (green ribbon), using PyMol.

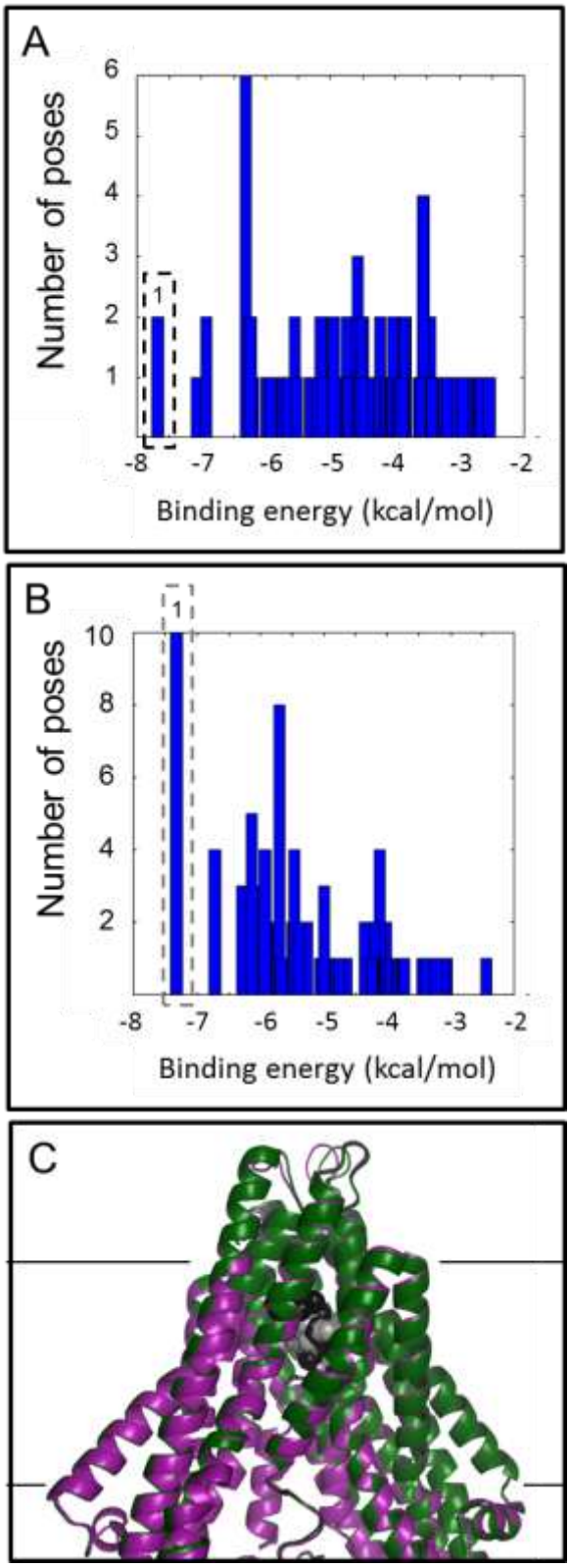


Figure 8. Verapamil binding to Hco-Pgp-13_04 (A, C) and Hco-Pgp-13_52 (B, C). A, B, energy clusterings of VRP docked to Hco-Pgp-13_04 (A) and Hco-Pgp-13_52 (B). C, binding sites of the lowest energy clusters of VRP on each protein. VRP1 is represented in black and gray spheres for docking on Hco-Pgp-13_04 (purple ribbon) and Hco-Pgp-13_52 (green ribbon), respectively, using PyMol.

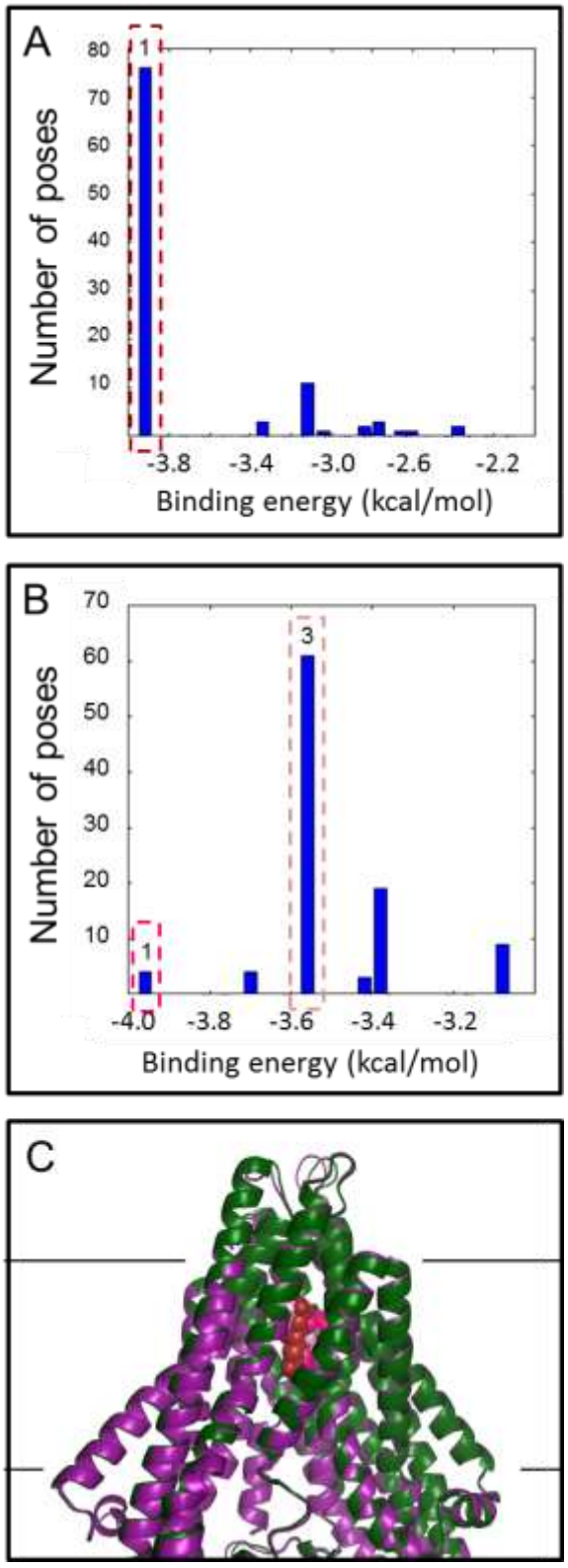


Figure 9. Rhodamine 123 binding to Hco-Pgp-13_04 (A, C) and Hco-Pgp-13_52 (B, C). A, B, energy clusterings of RHO docked to Hco-Pgp-13_04 (A) and Hco-Pgp-13_52 (B). C, binding sites of the 1st or 3rd lowest energy clusters of RHO on each protein. RHO1 is represented in dark red spheres for docking on Hco-Pgp-13_04 (purple ribbon), RHO1 and RHO3 are represented in dark pink and light pink spheres respectively for docking on Hco-Pgp-13_-52 (green ribbon), using PyMol.

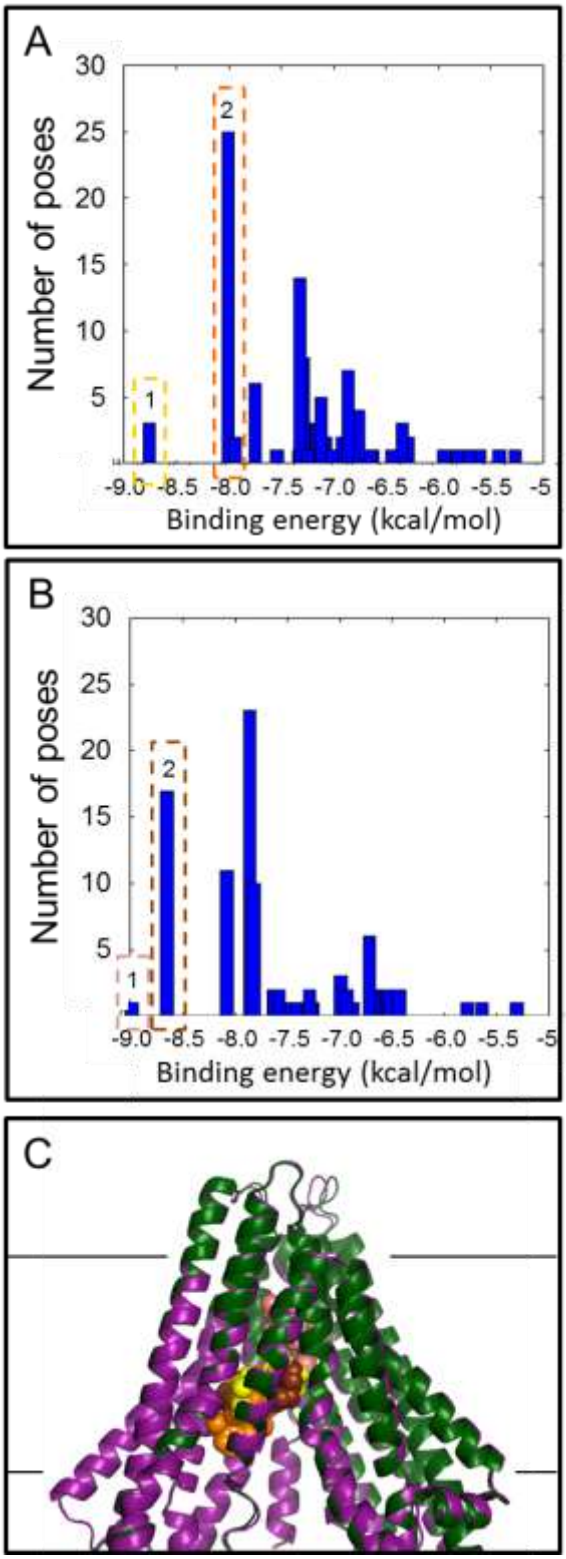


Figure 10. Doxorubicin binding to Hco-Pgp-13_04 (A, C) and Hco-Pgp-13_52 (B, C). A, B, energy clusterings of DXR docked to Hco-Pgp-13_04 (A) and Hco-Pgp-13_52 (B). C, binding sites of the 1st and 2nd lowest energy clusters of DXR on each protein. DXR1 and DXR2 are represented in yellow and brown spheres respectively for docking on Hco-Pgp-13_04 (purple ribbon) and in light pink and brown spheres respectively for docking on Hco-Pgp-13_52 (green ribbon), using PyMol.

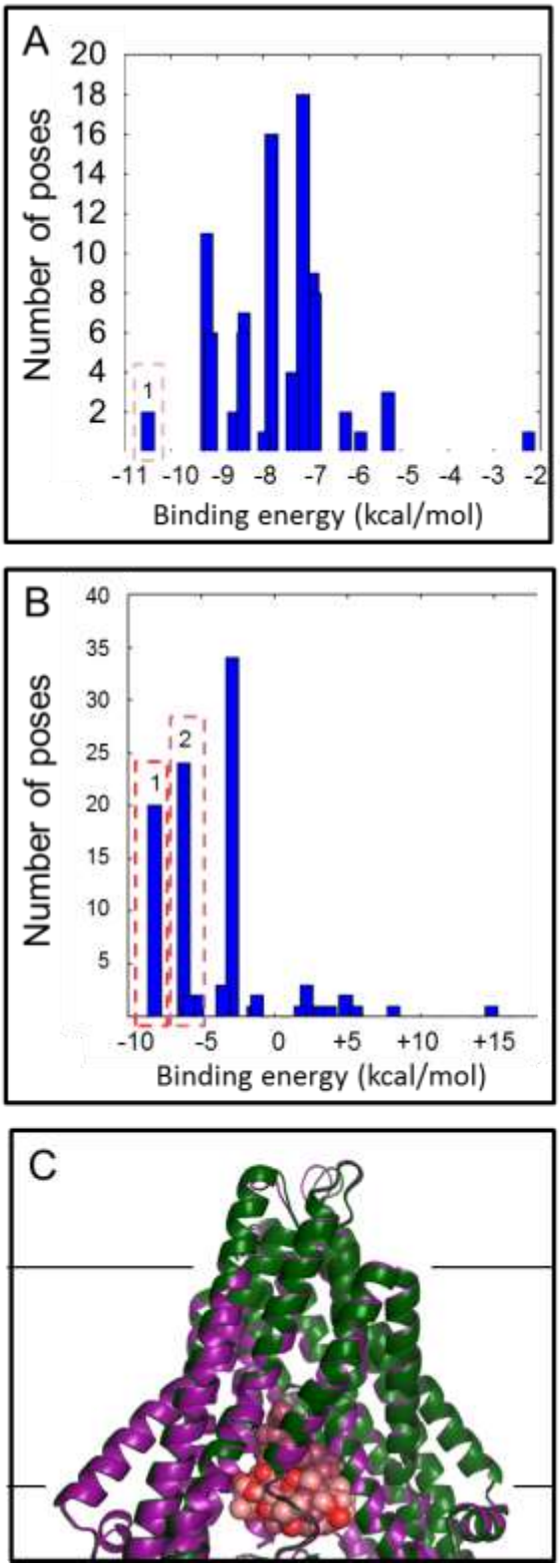


Figure 11. Valspodar binding to Hco-Pgp-13_04 (A, C) and _52 (B, C). A, B, energy clusterings of VSP docked to Hco-Pgp-13_04 (A) and Hco-Pgp-13_52 (B). C, binding sites of the 1st or 2nd lowest energy clusters of VSP on each protein. VSP1 is represented in light pink spheres for docking on Hco-Pgp-13_04 (purple ribbon), VSP1 and VSP2 are represented in red and dark pink spheres respectively for docking on Hco-Pgp-13_52 (green ribbon), using PyMol.

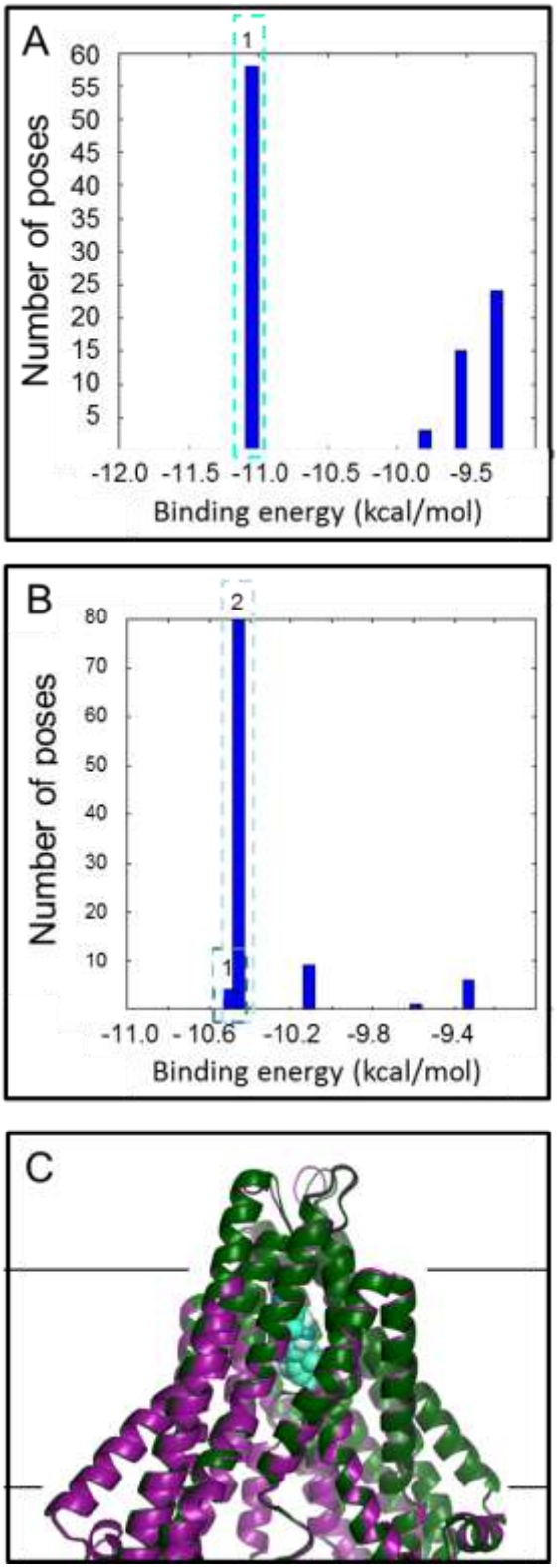


Figure 12. Cholesterol binding to Hco-Pgp-13_04 (A, C) and Hco-Pgp-13_52 (B, C). A, B, energy clusterings of RHO docked to Hco-Pgp-13_04 (A) and Hco-Pgp-13_52 (B). C, binding sites of the 1st or 3rd lowest energy clusters of CLS on each protein. CLS1 is represented in bright turquoise spheres for docking on Hco-Pgp-13_04 (purple ribbon), CLS1 and CLS2 are represented in dark turquoise and light blue spheres respectively for docking on Hco-Pgp-13_52 (green ribbon), using PyMol.

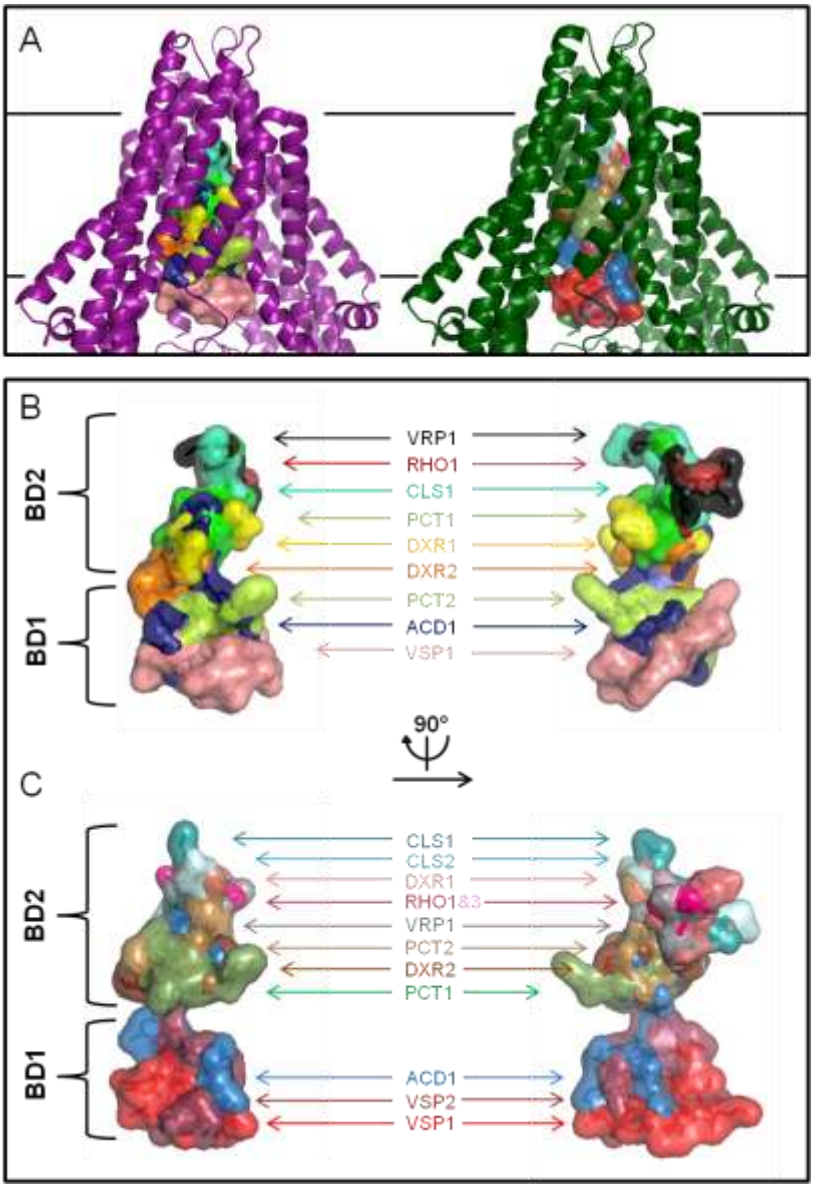


Figure 13. Binding domain characteristics of compounds docked on the Hco-Pgp-13 models 04 and 52. A. Front view of Hco-Pgp-13_04 and Hco-Pgp-13_52 represented in purple and green ribbon, respectively, with the binding sites of the representative lowest energy clusters of all the tested molecules represented in sticks and transparent surfaces, with similar color as in (B) and (C). **B.** Zoom, without Hco-Pgp-13_04, on the front and lateral views of the overlap between the binding sites of the first or second lowest energy clusters of paclitaxel (PCT1, bright green and PCT2, light green), verapamil (VRP1, black), rhodamine 123 (RHO1, dark red), doxorubicin (DXR1, yellow and DXR2, orange), valspodar (VSP1, light pink) and cholesterol (CLS1, bright turquoise). **C.** Zoom, without Hco-Pgp-13_52, on the front and lateral views of the overlap between the first, second or third lowest energy clusters of paclitaxel (PCT1, dark green and PCT2, light brown), verapamil (VRP1, gray), rhodamine 123 (RHO1, dark pink and RHO3, light pink), doxorubicin (DXR1, light pink and DXR2, brown), valspodar (VSP1, red and VSP2, dark pink) and cholesterol (CLS1, dark turquoise and CLS2, light blue). All molecules are represented in sticks and transparent surface.

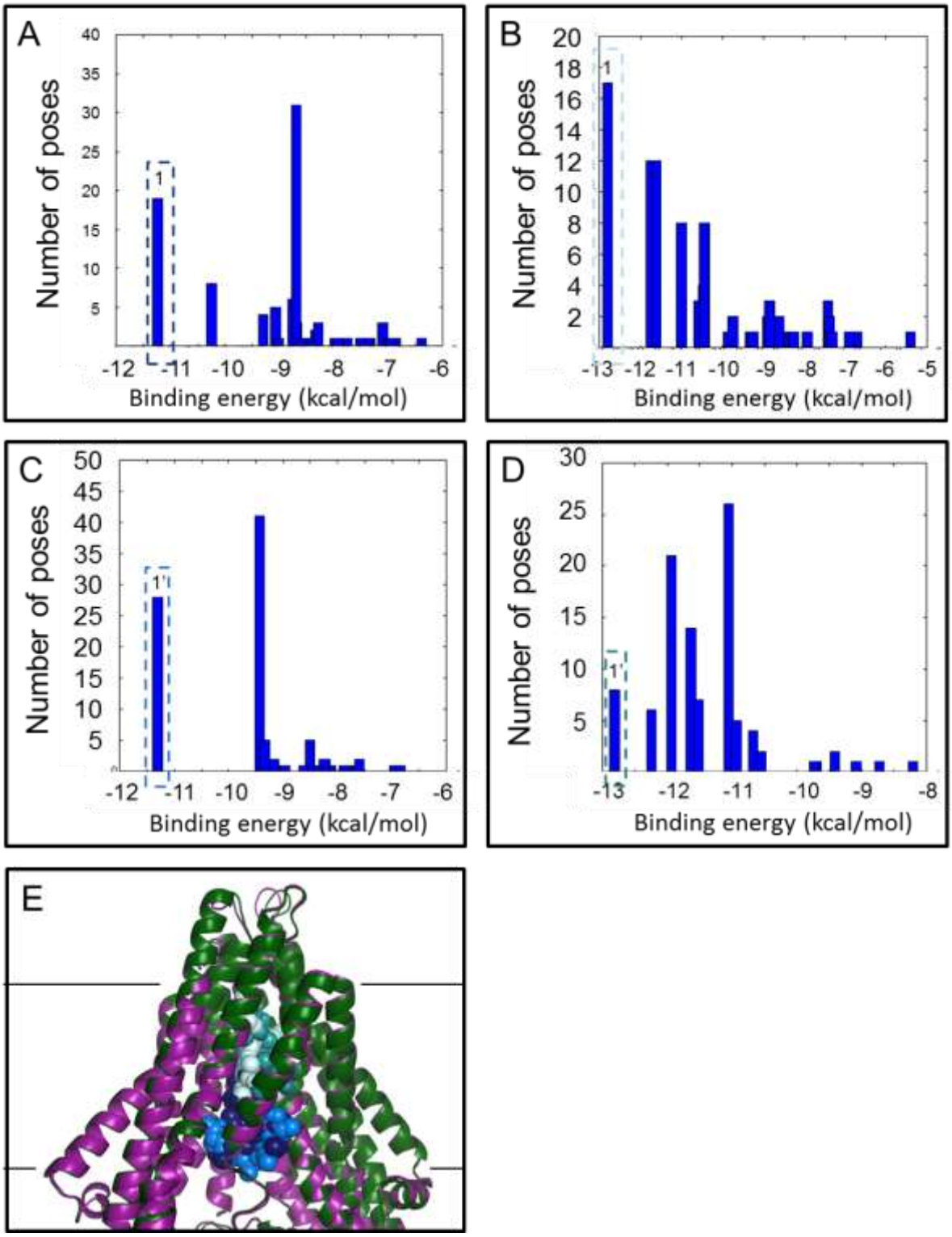


Figure 14. Ivermectin binding to Hco-Pgp-13_04 (A, C, E) and Hco-Pgp-13_52 (B, D, E). Energy clustering of two different starting conformers of IVM dokced to Hco-Pgp-13_04 (A, C) and Hco-Pgp-13_52 (B, D). E, binding sites of the 1st lowest energy cluster of each conformers. IVM1 and IVM1' are represented in dark blue and brigth blue spheres respectively for docking on Hco-Pgp-13_04 (purple ribbon), IVM1 and IVM1' are represented in light blue and dark turquoise spheres respectively for docking on Hco-Pgp-13_52 (green ribbon), using PyMol.

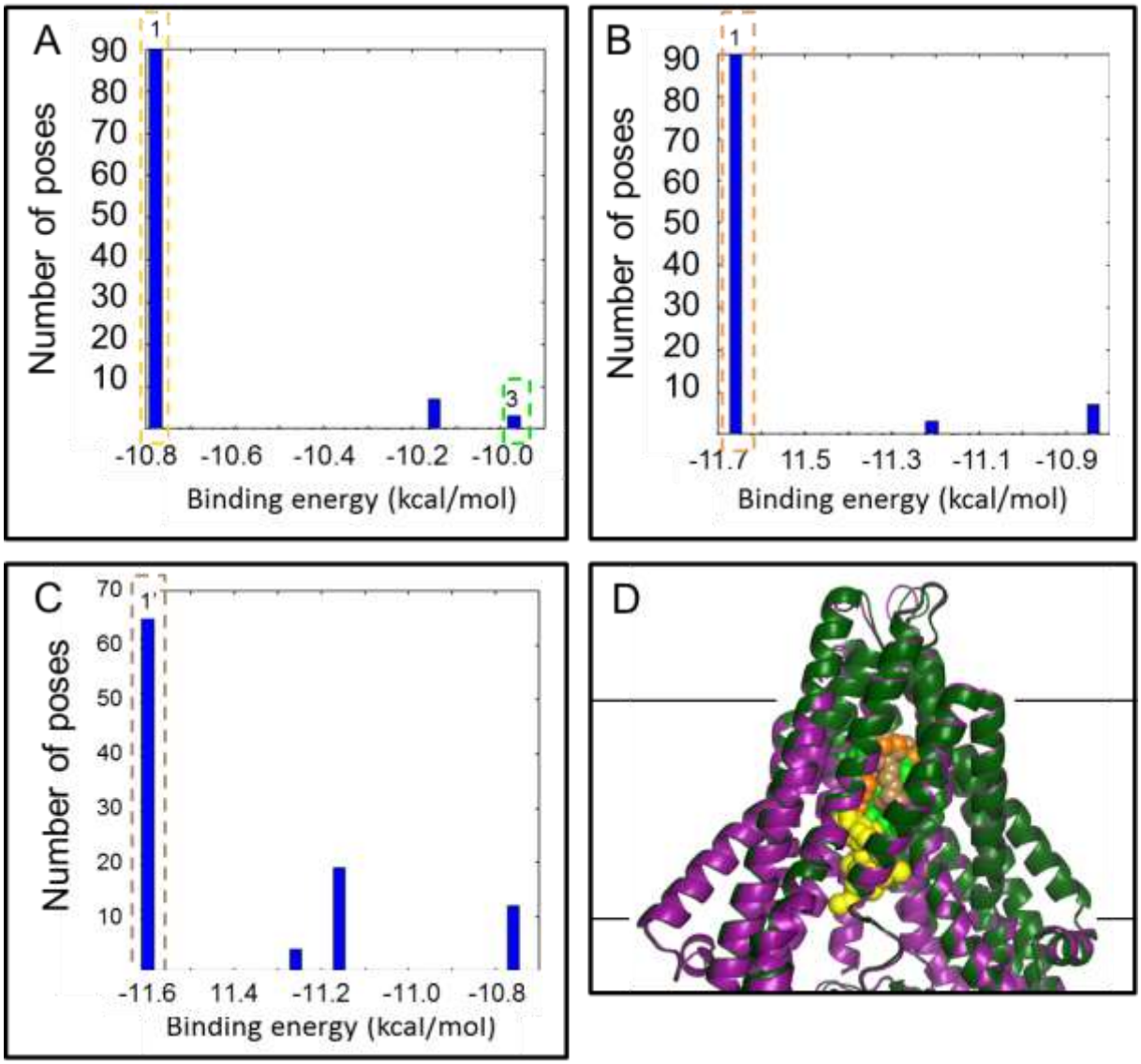


Figure 15. Moxidectin binding to Hco-Pgp-13_04 (A, D) and Hco-Pgp-13_52 (B, C, D). Energy clustering of the most representative starting conformer of MOX dokced to Hco-Pgp-13_04 (A) and of two most representative starting conformers of MOX Hco-Pgp-13_52 (B, C). D, binding sites of the 1st or 3rd lowest energy cluster of each conformers. MOX1 and MOX3 are represented in yellow and green spheres respectively for docking on Hco-Pgp-13_04 (purple ribbon), MOX1 and MOX1' are represented in orange and brown spheres respectively for docking on Hco-Pgp-13_52 (green ribbon), using PyMol.

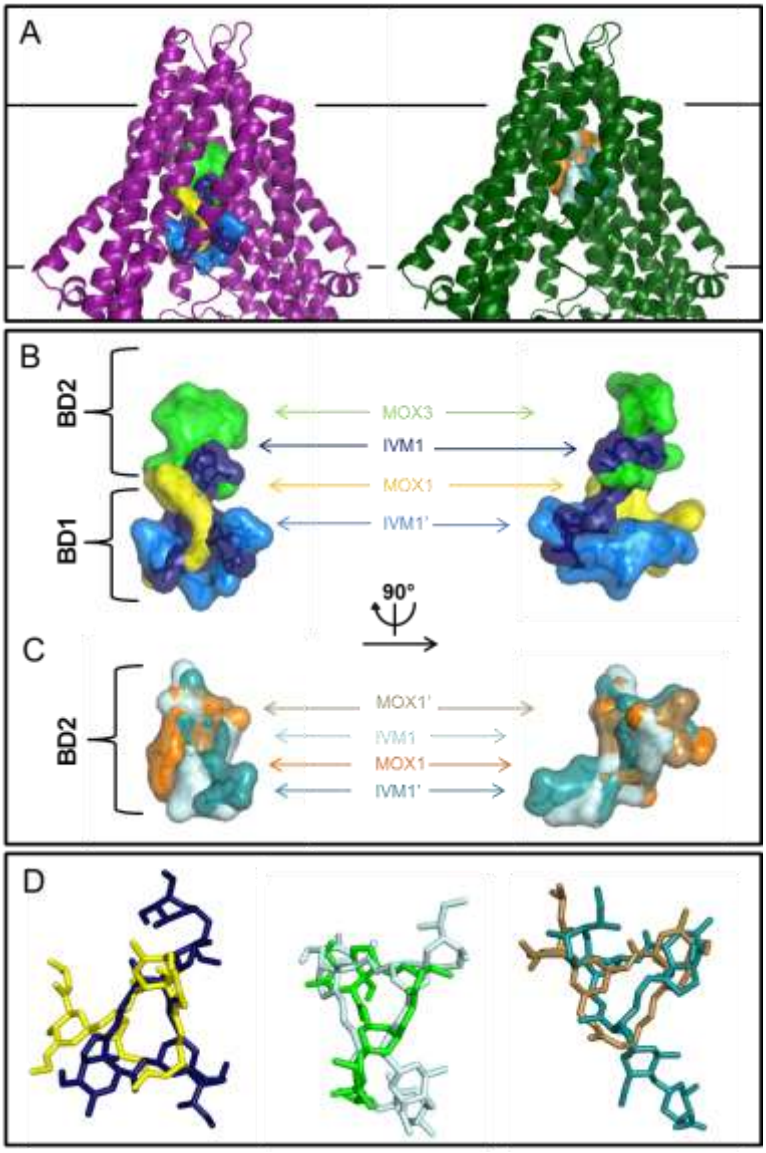


Figure 16. Binding domain characteristics of the macrocyclic lactones ivermectin (IVM) and moxidectin (MOX) on the Hco-Pgp-13 models 04 and 52. A. Front view of Hco-Pgp-13_04 and Hco-Pgp-13_52 represented in purple and green ribbon, respectively, with the binding sites of the representative lowest energy clusters of ivermectin and moxidectin all represented in sticks and transparent surfaces, with similar color as in (B) and (C). **B.** Zoom, without Hco-Pgp-13_04, on the front and lateral views of the overlap between the binding sites of representative lowest energy clusters of ivermectin (IVM1, dark blue and IVM1', bright blue) and moxidectin (MOX1, yellow and MOX3, green). **C.** Zoom, without Hco-Pgp-13_52, on the front and lateral views of the overlap between the binding sites of the representative lowest energy clusters of ivermectin (IVM1, light blue and IVM1', dark turquoise) and moxidectin (MOX1, orange and MO1', brown). All molecules are represented in sticks and transparent surface. **D.** Zoom on the superposition of the binding sites of molecules for which macrocycle rings are found to overlap: IVM1 and MOX1 docked to Hco-Pgp-13_04 (left panel), IVM1 docked to Hco-Pgp-13_52 and MOX3 docked to Hco-Pgp-13_04 (middle panel), IVM1' and MOX1' docked to Hco-Pgp-13_52 . Images were generated with PyMol.

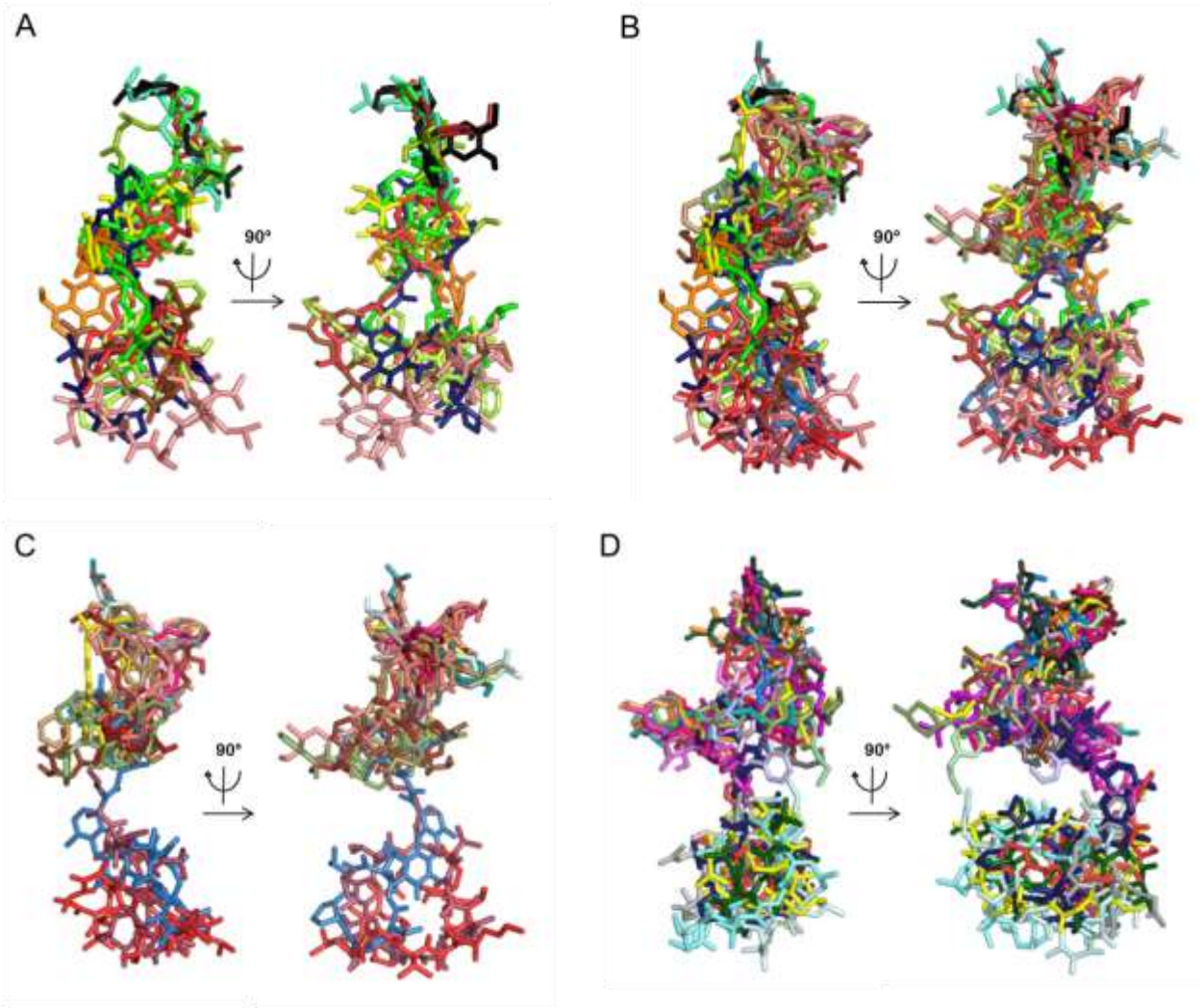


Figure 17. General binding pocket of Hco-Pgp-13 according to model 04 (A), models 04 and 52 superimposed (B), model 52 (C) and Cel-Pgp-1 according to PDB crystal structure 4F4C (D).

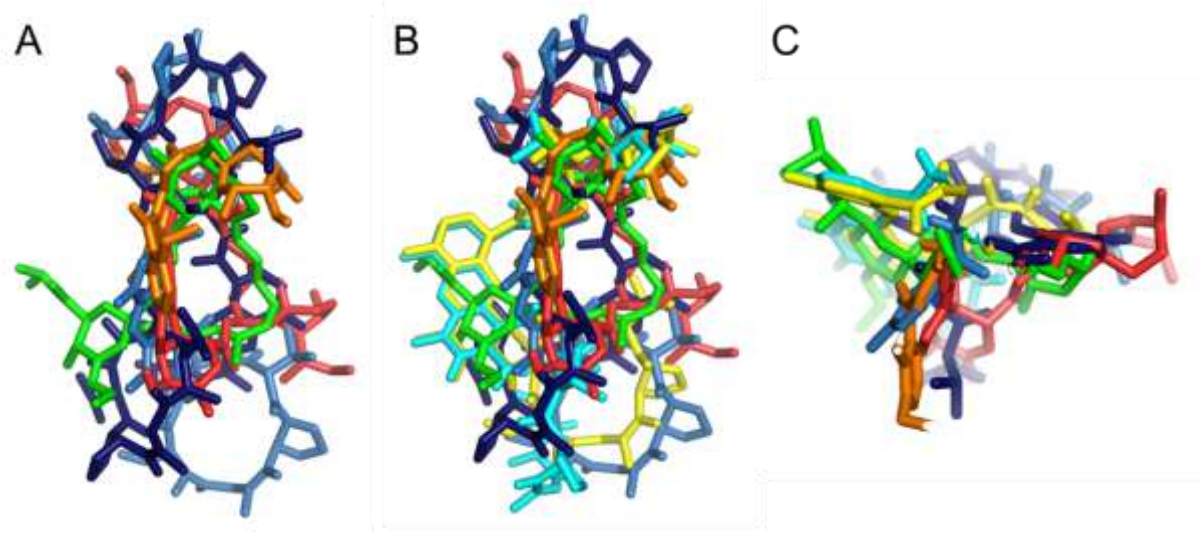


Figure 18. Comparison of the area separating BD1 and BD2 between the two models of Hco-Pgp-13 and Cel-Pgp-1 4F4C PDB structure.

Table 1. Physico-chemical, enzymological properties and docking characterization of the set of molecules tested for the validation step of the in silico procedure on HcoPgp13 models 4 and 52 : actinomycin D (ACD), doxorubicin (DXR), paclitaxel (PCT), rhodamine 123 (RHO), valspodar (VSP), verapamil (VRP) and cholesterol (CLS).

Molecule	DXR				PCT				RHO			VSP			VRP		CLS	
MW (Da)	504				854				381			1215			455		387	
logP [a]	1.9				3.3				3.3			4.7			5.0		7.1	
Hco-Pgp-13 Model	04		52		04		52		04		52	04		52	04	52	04	52
Cluster rank	1	2	1	2	1	2	1	2	1	1	3	1	1	2	1	1	1	1	2
Binding Energy (kcal/mol)	-8.8	-8.0	-9.0	-8.7	-10.5	-9.3	-11.5	-11.5	-3.9	-4.0	-3.6	-10.5	-8.2	-6.2	-7.7	-7.3	-11.1	-10.1	-10.5
Nb of poses	3	25	1	17	1	2	2	13	76	4	61	2	10	24	2	10	58	4	80
Nb of inter-acting residues	15	11	13	12	16	14	17	17	10	11	7	22	16	14	17	15	15	14	14
Nb of hotspot residues	9	3	8	6	13	4	11	14	8	8	5	4	4	2	13	10	11	10	11
Nb of H-bonds	3	5	2	3	1	2	1	1	2	1	1	0	1	0	2	1	2	1	0

(a) as calculated using Marvin Sketch with the consensus method (b) as reported in Jin et al, 2012. (*) cluster not positioned whithin the inner chamber.

Table 2. List of interacting residues of each transmembrane helix of HcoPgp13 models 4 and 52 with the indicated lowest energy clusters of actinomycin D (ACD), doxorubicin (DXR), paclitaxel (PCT), rhodamine 123 (RHO), valspodar (VSP) and verapamil (VRP).

Molecule	DXR				PCT				RHO			VSP			VRP			CLS	
Hco-Pgp-13 Model	4		52		4		52		4	52		4	52		4	52		4	52	
Cluster rank	1	2	1	2	1	2	1	2	1	1	3	1	1	2	1	1	1	1	2
TMa-b	S19 E21	S19 E21	S20 E21	S19 E21	E21	R11 D15 K25	E21	E21	R18 E21	E21	E21	R11 D15 K25 S26 V28 L29	Y8 R11 L29	R11 D15 K25 P32 A33 R45	R18 S20 E21	R18 S19 S20 E21	S20 E21	R18	R18
TM1			S115	Q108 M224	Q108 M224		Q108 M111	Q108							S115		S115	S115	
ECL1																		V118 T119	
TM3		L217 R220		M224	M224							K214		K214					
TM5			L335 M339						L335	L335					L335	L331 F334 L335		L335	F334 L335
TM6	L366 Y369	Y369	F358 V361	L365 L366 Y369	M362 L365 L366	Y369	L365 L366 Y369	F358 M362	V361 M362 L365	V361 M362 L365	F358 L365	Y369	Y369		F358 M362	V361 L364 L365	F358 M362 L365	F358 V361 M362 L365	F358 V361 M362 L364 L365

												G372	G372						
		L373				L373						L373	L373						
						P376						P376		P376					
						H377						**H377**		**H377**					
												V380							
TM7								**L772**	**L772**	**L772**					**L772**				
			F776		**F776**		**F776**	**F776**	**F776**	**F776**					**F776**	**F776**	**F776**	**F776**	
TM8			F803						F803										
							L909												
	L912						**L912**	**L912**											
	S915						S915												
TM10	M916	M916		M916	M916	M916	M916	M916					M916						
						I917													
	L919	L919				L919	L919					L919	L919	L919					
						A920						A920	A920	A920					
ICL4												L921							
												T922	T922						
	S988	**F985**					**S988**												
	Q989	**S988**			**Q989**			**Q989**											
	F991	Q989		Q989															
	T992				**T992**		**T992**	**T992**											
TM11								**Y993**											
								M995								**M995**			
															V996		**V996**		
																	C999		
															Y1000		**Y1000**	**Y1000**	**Y1000**
			F1018		**F1018**			**F1018**							**F1018**		**F1018**	**F1018**	**F1018**
															I1021		**I1021**		**I1021**
			I1022		**I1022**			**I1022**		**I1022**	**I1022**				**I1022**	**I1022**		**I1022**	**I1022**
	L1025		**L1025**		**L1025**			**L1025**	**L1025**						**L1025**	**L1025**	**L1025**	**L1025**	**L1025**
			L1026	**L1025**	**L1026**			**L1026**	**L1026**	**L1026**	**L1026**				**L1026**	**L1026**	**L1026**		
TM12	**S1028**				**S1028**		**S1028**												
	V1029			**S1028**	**V1029**		**V1029**	**V1029**								**V1029**	**V1029**		
				V1029			V1031												
						M1032	**M1032**	**M1032**					**M1032**						
						N1033						**N1033**	**N1033**						
						S1035						S1035	S1035						
												S1036	**S1036**	**S1036**					
Nter												P1039	P1039	P1039					
												E1040							

Bold: hotspot residues. Underscored: residues establishing a H-bond.

Table 3. Physico-chemical properties and docking characterization of the macrocyclic lactones ivermectin and moxidectin (MOX), to HcoPgp13 models 04 and 52.

Molecule : MW (Da) / logP [a]	**IVM:** 875 / 4.3				**MOX:** 640 / 5.4			
Hco-Pgp-13 Model	04		52		04		52	
Cluster rank	1	1’	1	1’	1	3	1	1’
Binding Energy (kcal/mol)	-11.2	-11.3	-12.8	-12.8	-10.8	-10.0	-11.7	-11.6
Nb of poses	19	28	17	8	90	3	90	65
Nb of inter-acting residues	14	17	21	21	11	13	15	15
Nb of hotspot residues	5	4	14	15	3	9	12	10
Nb of H-bonds	2	3	2	0	1	0	0	0

Table 4. List of interacting residues of each transmembrane helix of HcoPgp13 models 04 and 52 with the indicated lowest energy clusters of ivermectin (IVM) and moxidectin (MOX).

Molecule	IVM				MOX			
Hco-Pgp-13 Model	4		52		4		52	
Cluster rank	1	1'	1	1'	1	3	1	1'
TMa-b	S19 E21	R11 D15 K25	S20 E21	E21	D15 T22 K25		R18	S20 E21
TM1			**S115**	**S115**		**Q108** **M111**	**M111**	**S115**
ECL1			V118					
TM3		L217				M224		
TM5				**L335**			F334 **L335**	F334 **L335**
TM6	**L365** **Y369** L373	**Y369** G372 L373 P376 **H377**	F358 **M362** **L365**	F358 **V361** **M362** **L365**	**Y369** G372 L373 **H377**	**M362** **L366** **Y369**	F358 **V361** **M362** **L365** **L366**	F358 **V361** **L365**
TM7			**L772** **F776**	**F776**			**L772** **F776**	**F776**
TM8								F803
TM10	M916 I917	A913 M916 I917	**L912** A913 M916	L909 **L912** A913 M916	M916	M916		

	L919	L919			L919		
	A920	A920			A920		
IL10-11	L921						
				Q989			
TM11			**M995**				
			Y1000	**Y1000**		**Y1000**	**Y1000**
			F1018	**F1018**			**F1018**
				I1022		**I1022**	**I1022**
				M1024			
			L1025	**L1025**	**L1025**	**L1025**	**L1025**
			L1026	**L1026**	**L1026**	**L1026**	
TM12			**S1028**	**S1028**			
	V1029		**V1029**	**V1029**	**V1029**		**V1029**
			V1031	V1031			
	M1032	**M1032**	**M1032**		**M1032**		
		N1033					
	S1035	S1035					

Bold: hotspot residues. Underscored: residues establishing a H-bond.

Supplementary Table 1. Primers used for quantitative Real-Time-PCR on epithelial-like pig (Sus scrofa, Ssc) kidney cells (LLCPK1) transfected with *Haemonchus contortus* (Hco-) P-glycoprotein (Pgp) 13.

Primer	Nucleotide sequence (5' to 3')
Optimized *Hco-Pgp-13* Forward	ATCTCCGCCGGATACATTG
Optimized *Hco-Pgp-13* Reverse	GTGGTAATTTCCGGTAGCGTTT
Ssc-Pgp Forward	TGCCACCACGATAGCTGAGAACAT
Ssc-Pgp Reverse	ATGGCGATTCTCTGCTTCGTCCA
Ssc-gapdh Forward	AACTGCTTGGCACCCCTG
Ssc-gapdh Reverse	TTGGCAGCGCCGGTAGAA

gapdh = glyceraldehyde 3-phosphate dehydrogenase.

General Discussion

General Discussion and Prospects

In mammals, some ABC transporters, such as the human P-glycoprotein (ABCB1), transport numerous drugs and are strategically located on major protective barriers. For example, on the blood-brain barrier, they protect the brain from ivermectin entrance and neurotoxicity (Roulet et al., 2003; Schinkel, 1997; Schinkel et al., 1994). Some of these transporters are also involved in multidrug resistance in mammals and many other organisms (Eckford and Sharom, 2009; Jones and George, 2005; Koenderink et al., 2010; Lage, 2003; Leonard et al., 2003).

In nematodes, many Pgps orthologs have been identified. *H. contortus* is one of the most economically important parasites in small ruminants industry. It is one of the first nematodes, after *C. elegans*, to have its entire genome sequenced (Laing et al., 2011; Laing et al., 2013), and where 10 Pgps (i.e. full-size B sub-family members) were found (whereas 14 full Pgps and a pseudogene exist in *C. elegans)*. The individual role of each of these transporters remains to be determined. In particular, their involvement in multidrug transport, and eventually in resistance to anthelmintics, requires further investigation.

Some Pgps are located on the amphidial neurons, which appear as a possible route of entry of toxics, and which structure alteration has been linked with ML resistance in *C. elegans* and *H. contortus* (Dent et al., 2000; Freeman et al., 2003; Menez et al., 2016; Urdaneta-Marquez et al., 2014). As the AH targets are located in the nervous system of the worms, Pgps expressed in these organs might be critical in preventing these drugs from reaching their targets, thus avoiding the death of parasites by paralysis of somatic and pharyngeal muscles.

In *C. elegans*, two Pgps are found in these structures, Pgp-6 that has no ortholog in *H. contortus*, and Pgp-13. The latter has a close ortholog in *H. contortus*, thus called Hco-Pgp-13. We thus focused our interest on the characterization of the function Hco-Pgp-13. In this context, the release of a crystal structure for Cel-Pgp-1 (PDB: 4F4C) has been a major improvement in the deciphering of the function of individual nematode Pgps (Jin et al., 2012). It provides a first evidence of its multidrug transport function and a relevant template for modelling of close Pgp orthologs and drug docking.

1. Cel-Pgp-1 is a multidrug transporter with some homologies with mammalian Pgp

We further reinforced that Cel-Pgp-1 is a multidrug transporter by performing *in silico* docking calculations and we determine the binding modes of various molecules of pharmacological interest. This indicated a close homology of binding profile of Cel-Pgp-1 with that of Hsa-Pgp, as previously described by Jin et al. (2012). Such approach allowed us to identify residues of Cel-Pgp-1 implicated in the interaction with the various compounds tested. Many of them were aligned to hotspot residues, previously identified for being involved in ligand binding in several mammalian Pgps *in vitro* (Aller et al., 2009; Bessadok et al., 2011; Li et al., 2013; Loo et al., 2006a, b; Loo and Clarke, 2001, 2002; Shilling et al., 2006).

The shape of the full binding pocket of a Pgp was described for the first time as being composed of two interconnected major sub-domains, which resembles what was predicted earlier for a mammalian Pgp (Garrigues et al., 2002). This is also in line with previous *in vitro* (reviewed in (Ambudkar et al., 2006)), and *in silico* studies that had suggested the presence of at least two, possibly overlapping, binding sites for substrates on different Pgps (Chufan et al., 2013; Srinivasan et al., 2014).

2. Cel-Pgp-1 interacts with ML with high affinity

This modelling approach allowed us to demonstrate that some AH drugs currently used in veterinary and human medicine to treat helminth infections can bind to Cel-Pgp-1. The major finding arising from our approach concerns the high affinity binding of all MLs of current clinical importance, with a very specific binding site on Cel-Pgp-1. The drug docking provided here, combined with the ATPase activation by drugs with high affinity binding to Cel-Pgp-1 provided by Jin et al. (2012), allow us to confidently predict that all MLs tested here are certainly transported by Cel-Pgp-1. This is consistent with an increased sensitivity to IVM found *in vitro* for individual deletion strains of *C. elegans* (Janssen et al., 2013), although pgp-1 deletion strain was not the most affected.

Thus, this work opened the path to study homologies with other parasitic nematode Pgps. We then constructed two 3D models of Hco-Pgp-13 based on the 4F4C template, and they were found to have equivalent accuracies of 3D structure according to various evaluation tools.

3. Hco-Pgp-13 presents many homologies with Cel-Pgp-1, which makes it a putative ABC multidrug transporter

We performed *in silico* docking calculations on the two 3D models of Hco-Pgp-13, combined with *in vitro* transport and ATPase assays on two heterologous expression systems. One ligand was clearly identified by two approaches: actinomycin D, and three further pointed by preliminary experiments (paclitaxel, doxorubicin and valspodar). The high affinity binding predicted for various other, chemically-unrelated compounds on Hco-Pgp-13 in docking experiments indicates putative multispecific binding capacities for this transporter. In particular, due to their remarkably high affinity found *in silico*, the AH drugs IVM and MOX can be expected to be ligands, potentially transported, but this needs to be further tested by *in vitro* experiments.

Thus, the function of Hco-Pgp-13 looks somehow homologous to other *H. contortus* and *C. elegans* Pgps studied so far. Hco-Pgp-2, Hco-Pgp-9.1 and Hco-Pgp-16 were indeed shown to be able to transport rhodamine 123, in a way that can be inhibited by the ML IVM and MOX. This is consistent with a role for these proteins in the xenobiotics detoxication in *H. contortus* (Godoy et al., 2015a, 2016; Godoy et al., 2015b). In addition, in *C. elegans*, Cel-Pgp-2 has been suggested to be involved in pinocytosis, acidification of lysosomes and lipid storage in the intestine (Nunes et al., 2005; Schroeder et al., 2007), and this function could be conserved in Hco-Pgp-2 according to its localization partly in the intestine. Other functions described for individual Pgps in close nematodes are a possible role found for Cel-Pgp-3 in chloroquine and colchicine sensitivity (Broeks et al., 1995), and the involvement of the Pgp-9 of *C. elongatus* in sensitivity to KTZ, that can be modulated by ACD, VLN, DAU, and the three MLs: IVM, EPR and MOX (Kaschny et al., 2015). Plus, Pgps of *C. elegans* were all found individually involved in IVM sensitivity to various extent (Janssen et al., 2013).

Furthermore, we described a binding domain presenting a similar shape as that of Cel-Pgp-1, composed of two major sub-domains separated by a more or less narrow area of binding, depending on the model considered. We also identified interacting residues mostly conserved with Cel-Pgp-1 and mammalian Pgps in terms of alignment, although their nature was more variable. This transporter was then found to have a fair functional homology with Cel-Pgp-1 and Hsa-Pgp, consistent with a relatively conserved transport function of Pgps across species. This indicates that Hco-Pgp-13 might also be a multidrug transporter.

4. Hco-pgp-13 and Cel-Pgp-1 might also have homologies in their substrate recognition sites with other Pgps

One could then wonder whether most nematode Pgps will have more or less overlapping substrate recognition patterns, and hence biological function, and the possible selectivity of each isoform will remain to be determined. Indeed, this question of redundancy of the transport function should consider that the number of Pgps is larger in nematodes than in mammals, which express only 4 full transporters of the B sub-family (B1, B4, B5, B11). However, among these 4 transporters that display close sequence homologies, only B1 exhibits clear multidrug transport capacities.

Interestingly, most ABCB transporters across species are composed of 12 TM helices, but the crystal of Cel-Pgp-1 showed a supplementary N-term small transmembrane hairpin structure (TMa-b) never described before for any mammalian ABC protein. This could contribute the larger distance between NBDs observed in the open inward-facing conformation, as compared to the distance between NBDs of mouse Pgp that does not present TMa-b (Aller et al., 2009; Li et al., 2013; Szewczyk et al., 2015; Ward et al., 2013), due to sterical constraints. This structure might thus impact both the dimerization of NBDs and the substrate affinity at the level of TMDs, since these supplementary helices close them on one side within the membrane. In particular, it could both create a less easily reachable binding domain and change its residue nature, and thus its specificity for the potentially handled compounds.

Furthermore, 6/14 full Pgps (Pgp-1 and Pgp-10/11/12/13/14) have a long N-terminal sequence possibly forming a hairpin in *C. elegans*, whereas 2/10 Pgps (Pgp-10 and Pgp-13) in *H. contortus* are suspected to form such hairpin. Noticeably, Pgp-10 is the only ortholog in *H. contortus* of Pgp10/11 in *C. elegans*, and Pgp-13 is the only ortholog in *H.contortus* of Pgp12/13/14 in *C. elegans*, whereas Hco-Pgp-1, the closest ortholog of Cel-Pgp-1, presents a short N-terminal sequence. It thus will be desirable to evaluate whether these few TMa-b-containing isoforms display specific functional homologies. Of interest, the Pgps of the cluster Pgp-12/13/14/15 (Pgp-15 being a pseudogene) appeared to have different degrees of involvement in ML extruding in *C. elegans*, according to two studies showing an increase susceptibility of *C. elegans* to IVM individual Pgp deletion strains as compared to WT *C. elegans* (Ardelli and Prichard, 2013; Janssen et al., 2013). This could be due either to different affinities for each of them to these drugs or to their specific localization.

5. The expression of Hco-Pgp-13 in *Haemonchus contortus* is widely distributed, supporting an important function for this protein

Hco-Pgp-13 protein expression was located at the level of hypodermis and seam cell, both in the L3 larvae and adult. The L3 also showed staining of the membrane of epithelial cells of the gonad, and the adults appeared to express this transporter in pharyngeal glands or/and neurons, as well as in head neurons being possibly the amphids. Interestingly, this tissue expression pattern matched that of the whole cluster presenting the closest homologies of sequence in *C. elegans*, composed of Cel-Pgp-12/13/14/15.

Hco-Pgp-13 thus appeared to have a wider tissue distribution than other Pgps localized so far in this parasite or in the free-living worm *C. elegans*. In fact, all the nematode Pgps localized so far were found in a few tissues each, at variance with their closest human homolog ABCB1, nearly ubiquitious (although with various expression levels), but similarly to the very specific expression of the full transporters ABCB4 and B11 that play specific roles in the liver biliary canalicule.

Hco-Pgp-2 mRNA had been detected in the pharynx, mainly in pharyngeal glands, where Hco-Pgp-13 was also detected in our immunofluorescence assay, and in the intestine of *H. contortus* (Smith and Prichard, 2002), where we found no staining for Hco-Pgp-13. *Hco-pgp-2* mRNA was also detected in the vas deferens and lateral hypodermal chords anterior to the nerve ring, which correspond to the location of the amphidial nerves also stained in our study. However, the probe used for *in situ* hybridization had been designed on the partial sequence available at that time, not at the most specific region of Hco-Pgp-2, so that mRNA from other Hco-Pgps might have also been detected. Godoy et al. (2015a) precised the location of the protein of Hco-Pgp-2 by using specifically raised antibodies, and also found staining in the pharynx, in the deirid, a sensory cilia reaching the external environment on the side of the head, in head nerve cords and in the anterior intestine. Interestingly, Cel-Pgp-2, the closest unique ortholog of Hco-Pgp-2 in *C. elegans*, was also predicted to be expressed in the pharynx and intestine (Zhao et al., 2004). This localization was then confirmed, together with expression in the cell bodies of sensory neurons of the amphids (Nunes et al., 2005) (Schroeder et al., 2007). This showed a partly conserved localization of two orthologs across species.

On the opposite, HcoPgp-9.1 was found expressed at the level of the uterus of adult female *H. contortus* (Godoy et al., 2016), whereas Cel-Pgp-9 was predicted to be expressed in the pharynx and intestine (Zhao et al., 2004). However, in this case, two other orthologs of Cel-

Pgp-9 are thought to exist in *H. contortus*: Hco-Pgp-9.2 and Hco-Pgp-9.3, which could have different expression sites. In *C. elegans,* Cel-Pgp-1 and Pgp-3 were also suggested to be localized in the intestinal cells, as well as Pgp-3 in the excretory cell (Broeks et al., 1995) (Lincke et al., 1993). Cel-Pgp-3 had also been predicted to be expressed in the excretory cell and intestine by (Zhao et al., 2004), and other Cel-Pgps were also found each in some of the previously cited organs.

The tissue localization of Pgp-13 in the amphids of *C. elegans,* which were shown to be shortened in some IVM-resistant strains, and its detection at least in their close vicinity and in other relevant organs in *H. contortus*, together with its (multi)drug transport function suggested by our *in silico* analysis, makes it of high interest to study the possible involvement of this protein in ML resistance. The pharyngeal glands nerve cords containing glands and neurons were also found as a site of expression of Hco-Pgp-13. These structures could be very important to protect from ML, as the loss of pharyngeal pumping occurs in the response to these AH due to the flaccid paralysis of muscles by inhibition of electrogenic ion channels in the neurons innervating the pharynx, which is thought to cause the death of nematodes by starvation.

In addition, the expression of this Pgp in the hypodermis could also be critical, as these tissues are in close contact with muscles where neurons form junctions and where the paralysis occurs after IVM treatment. Since nematodes do not present blood circulatory system like mammals, the pathway used by exogenous compounds to reach their letal targets is not well described. Thus, an intestine location, as it is the case for Cel-Pgp-1, cannot be readily assessed for being important in protecting the nervous system of the treated nematode, in opposite to the situation in mammals, where the intestine is part of the primary route of entry of xenobiotics. As a whole, this shows that determination of the AH drug biodistribution inside a treated nematode is crucial for understanding its mode of action.

Finally, the hypothesis of the presence of only one ortholog in *H. contortus,* Hco-Pgp-13, as compared to the corresponding four genes in *C. elegans,* would make its effect comparable to the added effects of all these genes.

6. Perspectives for future investigation of Hco-Pgp-13 substrate profile

Based on previous *in vitro* experiments investigating the capacity of ACD and PCT to interact with Hco-Pgp-13, supplementary assays could be performed to test the interaction of various compounds, previously screened *in silico,* with Hco-Pgp-13. The *Pichia pastoris*

expression system allows the stable expression of many ABC transporters. It then permits a very direct measure of an interaction between a transporter and its putative substrates, by studying the stimulation by various compounds of the ATPase activity of the transporter present on isolated membranes. However, if this system allows to conclude confidently that drugs that activate the Pgp ATPase are definitively substrates, it cannot be used for some other compounds that do not stimulate ATPase activity. This has to be interpreted with caution as this does not obviously means no transport. Indeed, a basal activity exists, actually observed in our ATPase activity assays, which can be stimulated or unchanged in the presence of ligands, depending on the translocation rate of the studied compound as compared to the potential endogenous substrate. However, in contrast to other *in vitro* techniques, such as transport and cytotoxicity assays, the stimulation of ATPase activity permits the determination of the protein-ligand affinity, which can be comparable to the binding energy found *in silico* by taking into account the membrane partition of each molecule that depends on its hydrophobicity. Competition experiments, with one activator and a molecule that inhibits this activation, also allow gaining insight into the interaction of molecules which would otherwise show no modification of basal ATPase activity.

This heterologous expression system allowed in our project the successful expression of Hco-Pgp-13 protein at the membrane. We were able to start the investigation of Hco-Pgp-13 substrate profile, and found a stimulation of the ATPase activity of Hco-Pgp-13 by ACD, as well as PCT in preliminary experiments. This highly resembled the results found for Cel-Pgp-1 by Jin et al. (2012), and further experiments would help unravel the full recognition profile of Hco-Pgp-13, to determine if it also shows multispecific binding capacities. However, a limitation to the replication of this type of experiments has been the loss of the protein after freezing at -80°C and thawing of *P. pastoris* native membranes to perform functional assays, so that fresh cells were required to perform each experiment. The next experiments could then be performed on purified membranes reconstituted in proteoliposomes, presenting the advantage of a defined, controlled lipid composition. This would allow the direct study of ATPase activity stimulation by many compounds, without being hampered by the presence of other transporters in the system, so that the effect observed is the actual effect on the studied protein and does not require a negative control to validate the results. This would first allow determining the ability of Hco-Pgp-13 to transport ML, which is an important step in the elucidation of *H. contortus* resistance mechanisms to such drugs. More generally, this could help investigating *in vitro* the effect of molecules tested *in silico,* and thus increase the number

of mammalian Pgp substrates tested for their effect on Hco-Pgp-13, in order to gain insight into the multi-drug transport capacities of this transporter.

7. Could Hco-Pgp-13 also transport endogenous sterols?

The effect of the presence of sterols at various concentrations in the proteoliposome could help study the capacity of Hco-Pgp-13 to transport these putative endogenous substrates. Indeed, among the various mammalian ABC transporters having specific functions, some of them are transporters of lipids, with sometimes overlapping substrates (Le May et al., 2013; Lespine et al., 2009; Orlowski et al., 2007). In nematodes, membrane lipids are thus also relevant potential substrates for Pgps, as well as for the other ABC transporters. Nematodes require only a very low level of cholesterol, and they are auxotrophes for sterols, so that their membrane composition highly depends on the environment of each organism (Kurzchalia and Ward, 2003). For example, *C. elegans* feeds from ergosterol in the soil, whereas *H. contortus* is unlimited for cholesterol when it feeds from blood in its host. *P. pastoris* cells membranes are mainly composed of ergosterol, but their *in vitro* composition depends on the culture medium, whereas mammalian membranes are mostly made of cholesterol. It would thus be interesting to deplete the sterol level from proteoliposomes containing Hco-Pgp-13 and replace it or not with ergosterol or cholesterol. This would first permit to follow the basal activity variations, and then to compare the stimulation of ATPase activity by a same exogenous compound in presence of high concentrations of ergosterol, cholesterol, or few sterols. This way, we could identify the best initial composition of membranes that will yield a low basal activity of Hco-Pgp-13 ATPase, in order to obtain the highest stimulation effect possible by compounds investigated for their interaction with the transporter. This will then lead to a more precise comparison of the affinities of various ligands for Hco-Pgp-13.

However, we would need to keep in mind that a modification of ATPase activity stimulation level could be an indirect effect of the membrane composition changing the reaction environment. Indeed, parameters such as the thickness of the lipid bilayer, that is different in a membrane poor in sterol versus rich in sterol, can change the stability of a membrane protein. This possible importance of the composition of membranes can also be addressed by measuring its functionality when expressed in mammalian cells, allowing to wonder whether the higher level of cholesterol present in the mammalian membranes, as compared to the low level of sterols present in nematodes membranes, could lead to a higher basal transport, and eventual competition with a fluorescent substrate.

8. Further description of Hco-Pgp-13 molecular properties and implication for other Pgps

To further elucidate the molecular mechanisms of ligand binding on Hco-Pgp-13, mutagenesis studies could be performed on specific residues that seemed critical for ligand binding *in silico*. For example, the residue E21 was found to interact with all substrates tested except for VSP that bound in BD1 and CLS that bound very deep in BD2. Plus, it was found to form H-bonds with some poses of DXR, VRP and IVM. A change of the nature of this residue could be followed by expression in *P. pastoris* and ATPase activity assays with compounds that stimulated the ATPase of the WT protein, to see if the activation rate is changed, indication an alteration of the interaction. Such strategy could also help to determine the influence of the TMa-b hairpin by deleting its corresponding N-term sequence before transfection and functional assays. This was previously performed for Cel-Pgp-1 by Jin et al. (2012), who had not found much difference of function between the WT protein and the mutant Cel-Pgp-1 deleted for the first 56 amio-acids, in terms of protection form the cytotoxic effect of ACD and PCT, and stimulation of the ATPase activity by these two molecules. Thus, the absence of TMa-b in a protein that normally contains it might slightly modify the interaction site and affinity of substrates without dramatically changing the overall ligand profile of a Pgp.

This combination of *in silico* prediction of ligand binding and *in vitro* assay of their effect on the ATPase activity of a transporter is a powerful tool to better understand the molecular mechanism of ligand transport. If the first compounds predicted to be substrates of Hco-Pgp-13 *in silico* prove to have an effect of Hco-Pgp-13 ATPase activity *in vitro*, more docking calculations can be performed to screen interesting related compounds that can be further tested *in vitro*, which usually takes more time. By this way, the extensive multidrug capacity of a transporter can be more quickly studied in the wide variety of its substrates, with insights in the precise molecular mechanisms of binding that cannot be solved with biochemistry alone. Practically, both approaches are to be considered as valuably complementary, and should be combined as much as possible.

Moreover, the characterization for the first time of a protein of *H. contortus* at the molecular level will help shedding light on the transport mechanism of other Pgps of this organism, first considering Hco-Pgp-10 that also contains a long N-term sequence, and then Hco-Pgp-1 that is the closest ortholog in *H. contortus* to the Cel-Pgp-1 3D structure template. If Hco-Pgp-13 is proven to be a multidrug transporter, and to transport MLs, then its properties could be compared to those of the proteins Hco-Pgp-2/9/16 that have shown similar properties

in vitro. Then, the most direct Pgps to compare with will be other Pgps from pathogenic nematodes having the closest homology possible with Cel-Pgp-1 and Hco-Pgp-13, and those that have been proven as multidrug transporters. But only the full transporters of ABCB should be compared, as half-transporters should share more homology with the human ABCB10 protein, whose crystal structure has been solved (Shintre et al., 2013). This transporter, which showed less distance between its NBDs that Cel-Pgp-1, and whose conformation was almost independent of ATP presence, could also be used as template for homology modelling of close orthologs in species as close as possible.

9. Investigation of Hco-Pgp-13 importance in the living worm

To better understand the role of Hco-Pgp-13 *in vivo*, it will be interesting to study its involvement in ML resistance at the level of the living organism. This could be performed by injection of *H. contortus* with siRNA targeting *hco-pgp-13* and looking for phenotype changes, for example in the distribution of a fluorescent compound, as compared to the WT parasite.

The rescue of the function of deleted *pgp-13* in *C. elegans* by *Hco-pgp-13* could also be investigated. For this, a specific phenotype needs to be described for *C. elegans* deleted for *pgp-13* to evaluate drug toxicity in the absence and presence of various drugs, including ML. In particular, if Cel-Pgp-13 is critical for ML export and for protecting the worm from their toxicity, the motility, larval development or feeding capacity of *C. elegans* would be more altered in deleted strains than in WT in the presence of ML. Then, the sequence coding for *Hco-pgp-13* must be ligated to the sequence of the promoter of *Cel-pgp-13*, and fused in an expression vector for *C. elegans*. Gonad microinjection of the plasmid in *C. elegans* deletion strains finally allows obtaining *C. elegans* lineages expressing the transgene, and their phenotype can be compared with that of *C. elegans* deleted for Cel-Pgp-13. If a low sensitivity to ML is recovered in rescue strains as compared to deletion strains, Hco-Pgp-13 could be identified as important for ML transport and resistance *in vivo*.

However, as Hco-Pgp-13 appears to be the ortholog of Cel-Pgp-13, 12 and 14, double or triple deletion strains is required to see if the function of the whole group of Pgps could be rescued. Such *C. elegans* double KO line can be obtained by crossing two independant KO lines.

10. Perspective for fighting ML resistance in the field: design of inhibitors of Pgps which will have been identified as transporters of ML

Once the substrate binding profile of parasitic Pgps will have been characterized, and those involved in ML transport leading to resistance will have been defined in *H. contortus*, it will be interesting to screen for molecules that could be used as inhibitors of Pgp transport while MLs are administered. This could help getting back to their initial efficacy against parasitic nematode in the case of ML resistance, but an important point will be to make sure that these compounds block only the Pgps of the parasite and not the Pgp of the treated host. Otherwise, the ML treatment combined to suh inhibitor could be toxic for the host.

Given the substrate recognition pattern found for Cel-Pgp-1, that was relatively similar to mammalian Pgps (Jin et al., 2012), this task of looking for specificities between ABC transporters might be harder than expected. The Pgp of the sheep *Ovis aries* (Oar-Pgps), the main host of *H. contortus*, was found to show 58% mean similarity (according to BlastP) of residues within its two TMD domains with the corresponding domains of Cel-Pgp-1, which is equivalent to the similarity found between Hsa-Pgp and Cel-Pgp-1. As mentioned in Part II, A, of this manuscript, this similarity level is higher than that of all other Cel- or Hco-Pgps to Cel-Pgp-1, except for the Pgp-9 of the two species. It might thus be harder to find inhibitors of Cel/Hco-pgp-1, or Cel/Hco-Pgp-13 that do not inhibit Oar-Pgps than inhibitors of Cel/Hco-Pgp-10, showing the lowest homology of TMDs with Cel-Pgp-1 among all Pgps, that will not interact with mammalian Pgp. But this will require first to continue investigating the substrate profile of each Pgp one by one, in order to determine which one(s) has (/ve) the least similarity of binding profile, but are still involved in ML transport and thus are interesting to block in the parasite. As mammalian Pgps themselves can transport MLs, the balance between blocking *H. contortus* Pgps with inhibitors without being harmful for the host might be hard to find.

However, we have shown for Cel-Pgp-1 that the interacting residues, although presenting a high ratio of hotspot residues, have very few to no conservation of their nature, leading to the conclusion that the 3D recognition pattern clearly differed with that of the mammalian Pgp. Thus, the same drug binds to different places in the two multispecific chambers. The binding potencies of these multispecific domains are then qualitatively different, and it might be possible to rationally design an optimized ligand that will be able to distinguish them.

Finally, if the binding locations of compounds on various Pgps start to be described, the translocation mechanism of these transporters is not fully understood yet. In particular, the two

main binding sub-sites, when occupied, can be expected to have a different effect on the closing of NBDs required for ATP hydrolysis. Indeed, such transconformation appears harder when BD1 than when BD2 is occupied, and the ligands initially binding to BD1 might need one more step of movement within the protein before being expulsed in the external compartment. This would be coherent with a different effect on the stimulation of basal ATPase activity observed for substrates binding to BD2 and BD1, which translocation could take more or less time than that of endogenous substrates. Interestingly, the substrates observed *in vitro* to stimulate the most the basal ATPase activity of Cel-Pgp-1 were ACD, VNL and PCT, all binding to BD1. *In silico* molecular dynamics experiments could help elucidate this process involving a large amplitude transconformation of the protein, but only if performed with a sufficient time-scale to model the whole ligand translocation. Furthermore, as the lipid bilayer is hard to model around an ABC transporter, due to the variety of lipid molecules, the process described would be only approximate, as the lipids might play an important role during the dynamics of ligand translocation, probably more than during its first step of binding.

Thus, much still needs to be done to better understand the mechanistics of Pgps, and their possible involvement in the resistance to anthlemintics in parasitic nematodes. Plus, to overcome the problem of resistance in parasites of small ruminants, we need to keep in mind that the transport of drugs by Pgps might not be the only mechanism involved.

The approaches developed during this PhD have allowed gainining insights into responses important for the survival of parasitic nematodes when they are exposed to toxic drugs, and molecular targets to make drugs more efficient. This will hopefully contribute in optimizing the use of AH chemicals for the control of gastrointestinal parasitic nematodes of small ruminants.

Final Conclusion and Summary

We provided advancements in the domain of nematode Pgp molecular characterization and function, in particular in relation with anthelmintics (AH), including macrocyclic lactones (ML), transport. We reported the binding modes of various ligands on a Pgp of the nematode *C. elegans*, Cel-Pgp-1, using its recently published X-ray structural model, and in consistency with previous *in vitro* data (Jin et al., 2012). Overall, we also described for the first time the binding locations of a large set of AHs, and we showed that all ML bind to a very specific binding site with high affinity as compared to the validation set of molecules. In addition, we identified several hotspot residues for drug binding identified from alignment with mammalian Pgps as critical for ML binding on Cel-Pgp-1. Thus, we showed that a docking calculation strategy, usually performed to investigate the binding site of enzymes binding very specifically one substrate with a "key and lock" mechanism, could be a strong tool to gain insight into multiple binding sites of ligands on a multispecific ABC transporter harboring a wide binding domain.

Then, initially starting from its nucleotide sequence given by the large-scale sequencing of parasitic nematode *H. contortus* genome, we started shedding light on the substrates of the ABC transporter of Hco-Pgp-13 both *in silico*, and *in vitro*. Based on alignment with the crystal structure of Cel-Pgp-1 (PDB: 4F4C), we proposed 2 putative 3D structures for the Hco-Pgp-13 protein. The docking strategy previously validated was useful to screen for putative ligands of Hco-Pgp-13, which were for some of them tested *in vitro*. Various assays were performed in two heterologous Hco-Pgp-13 expression systems, the LLCPK1 mammalian cells and the yeast *Pichia pastoris*. ATPase activity assays performed on the latter showed that Hco-Pgp-13 can bind to and transport actinomycin D. More assays with other putative substrates will be required to discover the possibly wide variety of Hco-Pgp-13 substrates.

The localization of Hco-Pgp-13 was found to be at the level of hypodermis, seam cells and other epithelial cells of the L3, and at pharyngeal glands or/and neurons, as well as head neurons being possibly the amphids in adults. This reveals a wider expression of Hco-Pgp-13 than its close ortholog in *C. elegans*, Cel-Pgp-13, and it also corresponds to the locations of Cel-Pgp-12 and Cel-Pgp-14. More precise techniques, such as electron microscopy, will be necessary to determine if the neurons stained in the head with the two antibodies designed

against Hco-Pgp-13 are the amphids. Nonetheless, the various locations of expression of Hco-Pgp-13 suggest an important role for this protein in the parasite, either for transport of endogenous or exogenous compounds.

Then, the next question will be whether or not Hco-Pgp-13 can transport other drugs and among them, ML. If so, it will be important to study the consequences of this transport in ML resistance in *H. contortus*. Using the same method as the homology model built on Cel-Pgp-1, and docking calculations, the function of other Pgps of this parasite and others could be investigated, together with *in vitro* experiments. It will indeed be necessary to find which Hco-Pgp, among the last six not investigated individually so far, are able to bind to and transport ML. The docking strategy will then be very useful to search for compounds that could compete with ML for binding onto each Pgp and thus inhibit their transport, which could help fighting resistance in parasitic nematodes. Specificity towards the parasitic Pgps will be required to avoid toxicity for the host. This will demand a model of mammalian Pgp accurate enough to perform docking calculations on the ABC transporters of the host and avoid the use of inhibitors that would block it.

Finally, the global aim of this work will be to control the populations of *H. contortus* infecting goats and sheep, by targeting the worms that are already resistant to the main AHs used today in veterinary medicine. This would avoid the need to discover new anthelmintics, as the last ones released on the market were only efficient for a few years before leading to resistance. Plus, it could allow reducing the doses of AHs currently used by combination with inhibitors that can be selected for being less toxic.

References

Aceves, J., Erlij, D., and Martinez-Maranon, R. (1970). The mechanism of the paralysing action of tetramisole on Ascaris somatic muscle. British journal of pharmacology *38*, 602-607.

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

Alvarez, L.I., Solana, H.D., Mottier, M.L., Virkel, G.L., Fairweather, I., and Lanusse, C.E. (2005). Altered drug influx/efflux and enhanced metabolic activity in triclabendazole-resistant liver flukes. Parasitology *131*, 501-510.

Alvinerie, M., Dupuy, J., Kiki-Mvouaka, S., Sutra, J.F., and Lespine, A. (2008). Ketoconazole increases the plasma levels of ivermectin in sheep. Vet Parasitol *157*, 117-122.

Ambudkar, S.V., Dey, S., Hrycyna, C.A., Ramachandra, M., Pastan, I., and Gottesman, M.M. (1999). Biochemical, cellular, and pharmacological aspects of the multidrug transporter. Annu Rev Pharmacol Toxicol *39*, 361-398.

Ambudkar, S.V., Kim, I.W., and Sauna, Z.E. (2006). The power of the pump: mechanisms of action of P-glycoprotein (ABCB1). Eur J Pharm Sci *27*, 392-400.

Ambudkar, S.V., Kimchi-Sarfaty, C., Sauna, Z.E., and Gottesman, M.M. (2003). P-glycoprotein: from genomics to mechanism. Oncogene *22*, 7468-7485.

Ardelli, B.F., and Prichard, R.K. (2013). Inhibition of P-glycoprotein enhances sensitivity of Caenorhabditis elegans to ivermectin. Vet Parasitol *191*, 264-275.

Ashton, F.T., Li, J., and Schad, G.A. (1999). Chemo- and thermosensory neurons: structure and function in animal parasitic nematodes. Vet Parasitol *84*, 297-316.

Aubry, M.L., Cowell, P., Davey, M.J., and Shevde, S. (1970). Aspects of the pharmacology of a new anthelmintic: pyrantel. British journal of pharmacology *38*, 332-344.

Awadzi, K., Edwards, G., Opoku, N.O., Ardrey, A.E., Favager, S., Addy, E.T., Attah, S.K., Yamuah, L.K., and Quartey, B.T. (2004). The safety, tolerability and pharmacokinetics of levamisole alone, levamisole plus ivermectin, and levamisole plus albendazole, and their efficacy against Onchocerca volvulus. Ann Trop Med Parasitol *98*, 595-614.

Barrere, V., Beech, R.N., Charvet, C.L., and Prichard, R.K. (2014). Novel assay for the detection and monitoring of levamisole resistance in Haemonchus contortus. Int J Parasitol *44*, 235-241.

Bartley, D.J., McAllister, H., Bartley, Y., Dupuy, J., Menez, C., Alvinerie, M., Jackson, F., and Lespine, A. (2009). P-glycoprotein interfering agents potentiate ivermectin susceptibility in ivermectin sensitive and resistant isolates of Teladorsagia circumcincta and Haemonchus contortus. Parasitology *136*, 1081-1088.

Beech, R., Levitt, N., Cambos, M., Zhou, S., and Forrester, S.G. (2010). Association of ion-channel genotype and macrocyclic lactone sensitivity traits in Haemonchus contortus. Mol Biochem Parasitol *171*, 74-80.

Beech, R.N., Skuce, P., Bartley, D.J., Martin, R.J., Prichard, R.K., and Gilleard, J.S. (2011). Anthelmintic resistance: markers for resistance, or susceptibility? Parasitology *138*, 160-174.

Beis, K. (2015). Structural basis for the mechanism of ABC transporters. Biochem Soc Trans *43*, 889-893.

Bessadok, A., Garcia, E., Jacquet, H., Martin, S., Garrigues, A., Loiseau, N., Andre, F., Orlowski, S., and Vivaudou, M. (2011). Recognition of sulfonylurea receptor (ABCC8/9) ligands by the multidrug resistance transporter P-glycoprotein (ABCB1): functional similarities based on common structural features between two multispecific ABC proteins. J Biol Chem *286*, 3552-3569.

Blackhall, W.J., Liu, H.Y., Xu, M., Prichard, R.K., and Beech, R.N. (1998). Selection at a P-glycoprotein gene in ivermectin- and moxidectin-selected strains of Haemonchus contortus. Mol Biochem Parasitol *95*, 193-201.

Blaxter, M., and Denver, D.R. (2012). The worm in the world and the world in the worm. BMC Biol *10*, 57.

Blaxter, M., and Koutsovoulos, G. (2015). The evolution of parasitism in Nematoda. Parasitology *142 Suppl 1*, S26-39.

Blaxter, M.L., De Ley, P., Garey, J.R., Liu, L.X., Scheldeman, P., Vierstraete, A., Vanfleteren, J.R., Mackey, L.Y., Dorris, M., Frisse, L.M., *et al.* (1998). A molecular evolutionary framework for the phylum Nematoda. Nature *392*, 71-75.

Boray, J.C., Crowfoot, P.D., Strong, M.B., Allison, J.R., Schellenbaum, M., Von Orelli, M., and Sarasin, G. (1983). Treatment of immature and mature Fasciola hepatica infections in sheep with triclabendazole. The Veterinary record *113*, 315-317.

Borgers, M., and De Nollin, S. (1975). Ultrastructural changes in Ascaris suum intestine after mebendazole treatment in vivo. J Parasitol *61*, 110-122.

Borst, P., Evers, R., Kool, M., and Wijnholds, J. (1999). The multidrug resistance protein family. Biochim Biophys Acta *1461*, 347-357.

Boulin, T., Fauvin, A., Charvet, C.L., Cortet, J., Cabaret, J., Bessereau, J.L., and Neveu, C. (2011). Functional reconstitution of Haemonchus contortus acetylcholine receptors in Xenopus oocytes provides mechanistic insights into levamisole resistance. British journal of pharmacology *164*, 1421-1432.

Boulin, T., Gielen, M., Richmond, J.E., Williams, D.C., Paoletti, P., and Bessereau, J.L. (2008). Eight genes are required for functional reconstitution of the Caenorhabditis elegans levamisole-sensitive acetylcholine receptor. Proc Natl Acad Sci U S A *105*, 18590-18595.

Broeks, A., Gerrard, B., Allikmets, R., Dean, M., and Plasterk, R.H. (1996). Homologues of the human multidrug resistance genes MRP and MDR contribute to heavy metal resistance in the soil nematode Caenorhabditis elegans. The EMBO journal *15*, 6132-6143.

Broeks, A., Janssen, H.W., Calafat, J., and Plasterk, R.H. (1995). A P-glycoprotein protects Caenorhabditis elegans against natural toxins. The EMBO journal *14*, 1858-1866.

Bryan, J., Vila-Carriles, W.H., Zhao, G., Babenko, A.P., and Aguilar-Bryan, L. (2004). Toward linking structure with function in ATP-sensitive K+ channels. Diabetes *53 Suppl 3*, S104-112.

Bygarski, E.E., Prichard, R.K., and Ardelli, B.F. (2014). Resistance to the macrocyclic lactone moxidectin is mediated in part by membrane transporter P-glycoproteins: Implications for control of drug resistant parasitic nematodes. Int J Parasitol Drugs Drug Resist *4*, 143-151.

Chen, Z.S., Robey, R.W., Belinsky, M.G., Shchaveleva, I., Ren, X.Q., Sugimoto, Y., Ross, D.D., Bates, S.E., and Kruh, G.D. (2003). Transport of methotrexate, methotrexate polyglutamates, and 17beta-estradiol 17-(beta-D-glucuronide) by ABCG2: effects of acquired mutations at R482 on methotrexate transport. Cancer research *63*, 4048-4054.

Chermette (1982). L'haemonchose ovine et ses particularités : importance et situation actuelle en France. Le Point Vét *13*, 21-28.

Choudhury, H.G., Tong, Z., Mathavan, I., Li, Y., Iwata, S., Zirah, S., Rebuffat, S., van Veen, H.W., and Beis, K. (2014). Structure of an antibacterial peptide ATP-binding cassette transporter in a novel outward occluded state. Proc Natl Acad Sci U S A *111*, 9145-9150.

Chufan, E.E., Kapoor, K., Sim, H.M., Singh, S., Talele, T.T., Durell, S.R., and Ambudkar, S.V. (2013). Multiple transport-active binding sites are available for a single substrate on human P-glycoprotein (ABCB1). PLoS One *8*, e82463.

Cisternino, S., Mercier, C., Bourasset, F., Roux, F., and Scherrmann, J.M. (2004). Expression, up-regulation, and transport activity of the multidrug-resistance protein Abcg2 at the mouse blood-brain barrier. Cancer research *64*, 3296-3301.

Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res *16*, 10881-10890.

Coyne, M.J., Smith, G., and Johnstone, C. (1991). A study of the mortality and fecundity of Haemonchus contortus in sheep following experimental infections. Int J Parasitol *21*, 847-853.

Cromie, L., Ferry, M., Couper, A., Fields, C., and Taylor, S.M. (2006). Pharmacokinetics of a novel closantel/ivermectin injection in cattle. J Vet Pharmacol Ther *29*, 205-211.

Cully, D.F., Vassilatis, D.K., Liu, K.K., Paress, P.S., Van der Ploeg, L.H., Schaeffer, J.M., and Arena, J.P. (1994). Cloning of an avermectin-sensitive glutamate-gated chloride channel from Caenorhabditis elegans. Nature *371*, 707-711.

Cvilink, V., Lamka, J., and Skalova, L. (2009). Xenobiotic metabolizing enzymes and metabolism of anthelminthics in helminths. Drug metabolism reviews *41*, 8-26.

Dano, K. (1973). Active outward transport of daunomycin in resistant Ehrlich ascites tumor cells. Biochim Biophys Acta *323*, 466-483.

David, M.A., Orlowski, S., Prichard, R.K., Hashem, S., André, F., and Lespine, A. (2016). In silico analysis of the binding of anthelmintics to Caenorhabditis elegans P-glycoprotein 1. Int J Parasitol *In press*.

Davidson, A.L., Dassa, E., Orelle, C., and Chen, J. (2008). Structure, function, and evolution of bacterial ATP-binding cassette systems. Microbiol Mol Biol Rev *72*, 317-364, table of contents.

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

Dawson, R.J., and Locher, K.P. (2007). Structure of the multidrug ABC transporter Sav1866 from Staphylococcus aureus in complex with AMP-PNP. FEBS Lett *581*, 935-938.

Dean, M., Hamon, Y., and Chimini, G. (2001). The human ATP-binding cassette (ABC) transporter superfamily. J Lipid Res *42*, 1007-1017.

Deeley, R.G., Westlake, C., and Cole, S.P. (2006). Transmembrane transport of endo- and xenobiotics by mammalian ATP-binding cassette multidrug resistance proteins. Physiol Rev *86*, 849-899.

Dent, J.A., Smith, M.M., Vassilatis, D.K., and Avery, L. (2000). The genetics of ivermectin resistance in Caenorhabditis elegans. Proc Natl Acad Sci U S A *97*, 2674-2679.

Didier, A., and Loor, F. (1996). The abamectin derivative ivermectin is a potent P-glycoprotein inhibitor. Anticancer Drugs *7*, 745-751.

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

Doshi, R., and van Veen, H.W. (2013). Substrate binding stabilizes a pre-translocation intermediate in the ATP-binding cassette transport protein MsbA. J Biol Chem *288*, 21638-21647.

Doyle, L., and Ross, D.D. (2003). Multidrug resistance mediated by the breast cancer resistance protein BCRP (ABCG2). Oncogene *22*, 7340-7358.

Dupuy, J., Alvinerie, M., Menez, C., and Lespine, A. (2010). Interaction of anthelmintic drugs with P-glycoprotein in recombinant LLC-PK1-mdr1a cells. Chem Biol Interact *186*, 280-286.

Dupuy, J., Larrieu, G., Sutra, J.F., Eeckhoutte, C., and Alvinerie, M. (2001). Influence of verapamil on the efflux and metabolism of 14C moxidectin in cultured rat hepatocytes. J Vet Pharmacol Ther *24*, 171-177.

Dupuy, J., Larrieu, G., Sutra, J.F., Lespine, A., and Alvinerie, M. (2003). Enhancement of moxidectin bioavailability in lamb by a natural flavonoid: quercetin. Vet Parasitol *112*, 337-347.

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

Elmshauser, S., Straehle, L.C., Kranz, J., Krebber, R., and Geyer, J. (2014). Brain penetration of emodepside is increased in P-glycoprotein-deficient mice and leads to neurotoxicosis. J Vet Pharmacol Ther *38*, 74-79.

Eneroth, A., Astrom, E., Hoogstraate, J., Schrenk, D., Conrad, S., Kauffmann, H.M., and Gjellan, K. (2001). Evaluation of a vincristine resistant Caco-2 cell line for use in a calcein AM extrusion screening assay for P-glycoprotein interaction. Eur J Pharm Sci *12*, 205-214.

Fairweather, I., and Boray, J.C. (1999). Fasciolicides: efficacy, actions, resistance and its management. Vet J *158*, 81-112.

Fauvin, A., Charvet, C., Issouf, M., Cortet, J., Cabaret, J., and Neveu, C. (2010). cDNA-AFLP analysis in levamisole-resistant Haemonchus contortus reveals alternative splicing in a nicotinic acetylcholine receptor subunit. Mol Biochem Parasitol *170*, 105-107.

Feng, X.P., Hayashi, J., Beech, R.N., and Prichard, R.K. (2002). Study of the nematode putative GABA type-A receptor subunits: evidence for modulation by ivermectin. Journal of neurochemistry *83*, 870-878.

Forrester, S.G., Beech, R.N., and Prichard, R.K. (2004). Agonist enhacement of macrocyclic lactone activity at a glutamate-gated chloride channel subunit from Haemonchus contortus. Biochemical pharmacology *67*, 1019-1024.

Forrester, S.G., Prichard, R.K., and Beech, R.N. (2002). A glutamate-gated chloride channel subunit from Haemonchus contortus: expression in a mammalian cell line, ligand binding, and modulation of anthelmintic binding by glutamate. Biochemical pharmacology *63*, 1061-1068.

Forrester, S.G., Prichard, R.K., Dent, J.A., and Beech, R.N. (2003). Haemonchus contortus: HcGluCla expressed in Xenopus oocytes forms a glutamate-gated ion channel that is activated by ibotenate and the antiparasitic drug ivermectin. Mol Biochem Parasitol *129*, 115-121.

Freeman, A.S., Nghiem, C., Li, J., Ashton, F.T., Guerrero, J., Shoop, W.L., and Schad, G.A. (2003). Amphidial structure of ivermectin-resistant and susceptible laboratory and field strains of Haemonchus contortus. Vet Parasitol *110*, 217-226.

Gadsby, D.C., and Nairn, A.C. (1999). Control of CFTR channel gating by phosphorylation and nucleotide hydrolysis. Physiol Rev *79*, S77-S107.

Gallagher, S.R. (2006). One-dimensional SDS gel electrophoresis of proteins. Curr Protoc Immunol *Chapter 8*, Unit 8 4.

Gao, A., Wang, X., Xiang, W., Liang, H., Gao, J., and Yan, Y. (2010). Reversal of P-glycoprotein-mediated multidrug resistance in vitro by doramectin and nemadectin. J Pharm Pharmacol *62*, 393-399.

Garrigues, A., Loiseau, N., Delaforge, M., Ferte, J., Garrigos, M., Andre, F., and Orlowski, S. (2002). Characterization of two pharmacophores on the multidrug transporter P-glycoprotein. Mol Pharmacol *62*, 1288-1298.

Geary, T.G., Sims, S.M., Thomas, E.M., Vanover, L., Davis, J.P., Winterrowd, C.A., Klein, R.D., Ho, N.F., and Thompson, D.P. (1993). Haemonchus contortus: ivermectin-induced paralysis of the pharynx. Exp Parasitol *77*, 88-96.

Ghisi, M., Kaminsky, R., and Maser, P. (2007). Phenotyping and genotyping of Haemonchus contortus isolates reveals a new putative candidate mutation for benzimidazole resistance in nematodes. Vet Parasitol *144*, 313-320.

Ghosh, R., Andersen, E.C., Shapiro, J.A., Gerke, J.P., and Kruglyak, L. (2012). Natural variation in a chloride channel subunit confers avermectin resistance in C. elegans. Science *335*, 574-578.

Gilleard, J.S. (2006). Understanding anthelmintic resistance: the need for genomics and genetics. Int J Parasitol *36*, 1227-1239.

Glendinning, S.K., Buckingham, S.D., Sattelle, D.B., Wonnacott, S., and Wolstenholme, A.J. (2011). Glutamate-gated chloride channels of Haemonchus contortus restore drug sensitivity to ivermectin resistant Caenorhabditis elegans. PLoS One *6*, e22390.

Godoy, P., Che, H., Beech, R.N., and Prichard, R.K. (2015a). Characterization of Haemonchus contortus P-glycoprotein-16 and its interaction with the macrocyclic lactone anthelmintics. Mol Biochem Parasitol *204*, 11-15.

Godoy, P., Che, H., Beech, R.N., and Prichard, R.K. (2016). Characterisation of P-glycoprotein-9.1 in Haemonchus contortus. Parasit Vectors *9*, 52.

Godoy, P., Lian, J., Beech, R.N., and Prichard, R.K. (2015b). Haemonchus contortus P-glycoprotein-2: in situ localisation and characterisation of macrocyclic lactone transport. Int J Parasitol *45*, 85-93.

Gonzalez-Canga, A., Fernandez-Martinez, N., Sahagun-Prieto, A., Diez-Liebana, M.J., Sierra-Vega, M., and Garcia-Vieitez, J.J. (2009). A review of the pharmacological interactions of ivermectin in several animal species. Current drug metabolism *10*, 359-368.

Gordon, H.M. (1961). Thiabendazole: a highly effective anthelmintic for sheep. Nature *191*, 1409-1410.

Gottesman, M.M., and Pastan, I. (1993). Biochemistry of multidrug resistance mediated by the multidrug transporter. Annu Rev Biochem *62*, 385-427.

Gottesman, M.M., Pastan, I., and Ambudkar, S.V. (1996). P-glycoprotein and multidrug resistance. Curr Opin Genet Dev *6*, 610-617.

Griffin, J., Fletcher, N., Clemence, R., Blanchflower, S., and Brayden, D.J. (2005). Selamectin is a potent substrate and inhibitor of human and canine P-glycoprotein. J Vet Pharmacol Ther *28*, 257-265.

Haber, C.L., Heckaman, C.L., Li, G.P., Thompson, D.P., Whaley, H.A., and Wiley, V.H. (1991). Development of a mechanism of action-based screen for anthelmintic microbial metabolites with avermectinlike activity and isolation of milbemycin-producing Streptomyces strains. Antimicrob Agents Chemother *35*, 1811-1817.

Harder, A., Holden-Dye, L., Walker, R., and Wunderlich, F. (2005). Mechanisms of action of emodepside. Parasitology research *97 Suppl 1*, S1-10.

Heiman, M.G., and Shaham, S. (2009). DEX-1 and DYF-7 establish sensory dendrite length by anchoring dendritic tips during cell migration. Cell *137*, 344-355.

Hibbs, R.E., and Gouaux, E. (2011). Principles of activation and permeation in an anion-selective Cys-loop receptor. Nature *474*, 54-60.

Higgins, C.F. (1992). ABC transporters: from microorganisms to man. Annu Rev Cell Biol *8*, 67-113.

Higgins, C.F., and Linton, K.J. (2004). The ATP switch model for ABC transporters. Nat Struct Mol Biol *11*, 918-926.

Hohl, M., Briand, C., Grutter, M.G., and Seeger, M.A. (2012). Crystal structure of a heterodimeric ABC transporter in its inward-facing conformation. Nat Struct Mol Biol *19*, 395-402.

Holden-Dye, L., and Walker, R.J. (2007). Anthelmintic drugs. WormBook : the online review of C elegans biology, 1-13.

Hollenstein, K., Frei, D.C., and Locher, K.P. (2007). Structure of an ABC transporter in complex with its binding protein. Nature *446*, 213-216.

Hopfner, K.P., Karcher, A., Shin, D.S., Craig, L., Arthur, L.M., Carney, J.P., and Tainer, J.A. (2000). Structural biology of Rad50 ATPase: ATP-driven conformational control in DNA double-strand break repair and the ABC-ATPase superfamily. Cell *101*, 789-800.

Hugnet, C., Lespine, A., and Alvinerie, M. (2007). Multiple oral dosing of ketoconazole increases dog exposure to ivermectin. J Pharm Pharm Sci *10*, 311-318.

Hung, L.W., Wang, I.X., Nikaido, K., Liu, P.Q., Ames, G.F., and Kim, S.H. (1998). Crystal structure of the ATP-binding subunit of an ABC transporter. Nature *396*, 703-707.

Husada, F., Gouridis, G., Vietrov, R., Schuurman-Wolters, G.K., Ploetz, E., de Boer, M., Poolman, B., and Cordes, T. (2015). Watching conformational dynamics of ABC transporters with single-molecule tools. Biochem Soc Trans *43*, 1041-1047.

Hyde, S.C., Emsley, P., Hartshorn, M.J., Mimmack, M.M., Gileadi, U., Pearce, S.R., Gallagher, M.P., Gill, D.R., Hubbard, R.E., and Higgins, C.F. (1990). Structural model of ATP-binding proteins associated with cystic fibrosis, multidrug resistance and bacterial transport. Nature *346*, 362-365.

Jabbar, A., Iqbal, Z., Kerboeuf, D., Muhammad, G., Khan, M.N., and Afaq, M. (2006). Anthelmintic resistance: the state of play revisited. Life Sci *79*, 2413-2431.

Jacquiet, P., Cabaret, J., Cheikh, D., and Thiam, A. (1995). Experimental study of survival strategy of Haemonchus contortus in sheep during the dry season in desert areas of the Mauritania. J Parasitol *81*, 1013-1015.

James, C.E., and Davey, M.W. (2009). Increased expression of ABC transport proteins is associated with ivermectin resistance in the model nematode Caenorhabditis elegans. Int J Parasitol *39*, 213-220.

Jani, M., Makai, I., Kis, E., Szabo, P., Nagy, T., Krajcsi, P., and Lespine, A. (2010). Ivermectin interacts with human ABCG2. J Pharm Sci *100*, 94-97.

Janssen, I.J., Krucken, J., Demeler, J., and von Samson-Himmelstjerna, G. (2013). Caenorhabditis elegans: modest increase of susceptibility to ivermectin in individual P-glycoprotein loss-of-function strains. Exp Parasitol *134*, 171-177.

Jara, G.E., Vera, D.M., and Pierini, A.B. (2013). Binding of modulators to mouse and human multidrug resistance P-glycoprotein. A computational study. J Mol Graph Model *46*, 10-21.

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

Jones, P.M., and George, A.M. (2004). The ABC transporter structure and mechanism: perspectives on recent research. Cell Mol Life Sci *61*, 682-699.

Jones, P.M., and George, A.M. (2005). Multidrug resistance in parasites: ABC transporters, P-glycoproteins and molecular modelling. Int J Parasitol *35*, 555-566.

Jonker, J.W., Merino, G., Musters, S., van Herwaarden, A.E., Bolscher, E., Wagenaar, E., Mesman, E., Dale, T.C., and Schinkel, A.H. (2005). The breast cancer resistance protein BCRP (ABCG2) concentrates drugs and carcinogenic xenotoxins into milk. Nature medicine *11*, 127-129.

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

Kaminsky, R., Ducray, P., Jung, M., Clover, R., Rufener, L., Bouvier, J., Weber, S.S., Wenger, A., Wieland-Berghausen, S., Goebel, T., *et al.* (2008a). A new class of anthelmintics effective against drug-resistant nematodes. Nature *452*, 176-180.

Kaminsky, R., Gauvry, N., Schorderet Weber, S., Skripsky, T., Bouvier, J., Wenger, A., Schroeder, F., Desaules, Y., Hotz, R., Goebel, T., *et al.* (2008b). Identification of the amino-acetonitrile derivative monepantel (AAD 1566) as a new anthelmintic drug development candidate. Parasitology research *103*, 931-939.

Kaplan, R.M. (2004). Drug resistance in nematodes of veterinary importance: a status report. Trends Parasitol *20*, 477-481.

Karwatsky, J., Lincoln, M.C., and Georges, E. (2003). A mechanism for P-glycoprotein-mediated apoptosis as revealed by verapamil hypersensitivity. Biochemistry *42*, 12163-12173.

Kaschny, M., Demeler, J., Janssen, I.J., Kuzmina, T.A., Besognet, B., Kanellos, T., Kerboeuf, D., von Samson-Himmelstjerna, G., and Krucken, J. (2015). Macrocyclic lactones differ in interaction with recombinant P-glycoprotein 9 of the parasitic nematode Cylicocylus elongatus and ketoconazole in a yeast growth assay. PLoS Pathog *11*, e1004781.

Kawalek, J.C., Rew, R.S., and Heavner, J. (1984). Glutathione-S-transferase, a possible drug-metabolizing enzyme, in Haemonchus contortus: comparative activity of a cambendazole-resistant and a susceptible strain. Int J Parasitol *14*, 173-175.

Kerboeuf, D., and Aycardi, J. (1999). Unexpected increased thiabendazole tolerance in Haemonchus contortus resistant to anthelmintics by modulation of glutathione activity. Parasitology research *85*, 713-718.

Kerboeuf, D., and Guegnard, F. (2011). Anthelmintics are substrates and activators of nematode P glycoprotein. Antimicrob Agents Chemother *55*, 2224-2232.

Kerboeuf, D., Hubert, J., Cardinaud, B., and Blond-Riou, F. (1995). The persistence of the efficacy of injectable or oral moxidectin against Teladorsagia, Haemonchus and Trichostrongylus species in experimentally infected sheep. The Veterinary record *137*, 399-401.

Kiki-Mvouaka, S., Menez, C., Borin, C., Lyazrhi, F., Foucaud-Vignault, M., Dupuy, J., Collet, X., Alvinerie, M., and Lespine, A. (2010). Role of P-glycoprotein in the disposition of

macrocyclic lactones: A comparison between ivermectin, eprinomectin, and moxidectin in mice. Drug Metab Dispos *38*, 573-580.

Kodan, A., Yamaguchi, T., Nakatsu, T., Sakiyama, K., Hipolito, C.J., Fujioka, A., Hirokane, R., Ikeguchi, K., Watanabe, B., Hiratake, J., *et al.* (2014). Structural basis for gating mechanisms of a eukaryotic P-glycoprotein homolog. Proc Natl Acad Sci U S A *111*, 4049-4054.

Koenderink, J.B., Kavishe, R.A., Rijpma, S.R., and Russel, F.G. (2010). The ABCs of multidrug resistance in malaria. Trends Parasitol *26*, 440-446.

Kozak, M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res *15*, 8125-8148.

Kraushaar, A. (1954). [Chemotherapeutic activity of halogenated salicylanilides in relation to their constitution]. Arzneimittelforschung *4*, 548-551.

Kurzchalia, T.V., and Ward, S. (2003). Why do worms need cholesterol? Nat Cell Biol *5*, 684-688.

Kwa, M.S., Okoli, M.N., Schulz-Key, H., Okongkwo, P.O., and Roos, M.H. (1998). Use of P-glycoprotein gene probes to investigate anthelmintic resistance in Haemonchus contortus and comparison with Onchocerca volvulus. Int J Parasitol *28*, 1235-1240.

Kwa, M.S., Veenstra, J.G., Van Dijk, M., and Roos, M.H. (1995). Beta-tubulin genes from the parasitic nematode Haemonchus contortus modulate drug resistance in Caenorhabditis elegans. Journal of molecular biology *246*, 500-510.

Lacey, E. (1988). The role of the cytoskeletal protein, tubulin, in the mode of action and mechanism of drug resistance to benzimidazoles. Int J Parasitol *18*, 885-936.

Lacey, E. (1990). Mode of action of benzimidazoles. Parasitol Today *6*, 112-115.

Lage, H. (2003). ABC-transporters: implications on drug resistance from microorganisms to human cancers. Int J Antimicrob Agents *22*, 188-199.

Laing, R., Hunt, M., Protasio, A.V., Saunders, G., Mungall, K., Laing, S., Jackson, F., Quail, M., Beech, R., Berriman, M., *et al.* (2011). Annotation of two large contiguous regions from the Haemonchus contortus genome using RNA-seq and comparative analysis with Caenorhabditis elegans. PLoS One *6*, e23216.

Laing, R., Kikuchi, T., Martinelli, A., Tsai, I.J., Beech, R.N., Redman, E., Holroyd, N., Bartley, D.J., Beasley, H., Britton, C., *et al.* (2013). The genome and transcriptome of Haemonchus contortus, a key model parasite for drug and vaccine discovery. Genome Biol *14*, R88.

Lambshead, P.J.D. (1993). Recent developments in marine benthic biodiversity research. Oceanis *19*, 5–24.

Lartillot, N., Lepage, T., and Blanquart, S. (2009). PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics *25*, 2286-2288.

Le May, C., Berger, J.M., Lespine, A., Pillot, B., Prieur, X., Letessier, E., Hussain, M.M., Collet, X., Cariou, B., and Costet, P. (2013). Transintestinal cholesterol excretion is an active metabolic process modulated by PCSK9 and statin involving ABCB1. Arterioscler Thromb Vasc Biol *33*, 1484-1493.

Lee, J.Y., Kinch, L.N., Borek, D.M., Wang, J., Urbatsch, I.L., Xie, X.S., Grishin, N.V., Cohen, J.C., Otwinowski, Z., Hobbs, H.H., *et al.* (2016). Crystal structure of the human sterol transporter ABCG5/ABCG8. Nature *533*, 561-564.

Lee, J.Y., Yang, J.G., Zhitnitsky, D., Lewinson, O., and Rees, D.C. (2014). Structural basis for heavy metal detoxification by an Atm1-type ABC exporter. Science *343*, 1133-1136.

Leonard, G.D., Fojo, T., and Bates, S.E. (2003). The role of ABC transporters in clinical practice. Oncologist *8*, 411-424.

Lespine, A. (2013). Lipid-like properties and pharmacology of the anthelmintic macrocyclic lactones. Expert opinion on drug metabolism & toxicology *9*, 1581-1595.

Lespine, A., Alvinerie, M., Vercruysse, J., Prichard, R.K., and Geldhof, P. (2008). ABC transporter modulation: a strategy to enhance the activity of macrocyclic lactone anthelmintics. Trends Parasitol *24*, 293-298.

Lespine, A., Chanoit, G., Bousquet-Melou, A., Lallemand, E., Bassissi, F.M., Alvinerie, M., and Toutain, P.L. (2006a). Contribution of lymphatic transport to the systemic exposure of orally administered moxidectin in conscious lymph duct-cannulated dogs. Eur J Pharm Sci *27*, 37-43.

Lespine, A., Dupuy, J., Alvinerie, M., Comera, C., Nagy, T., Krajcsi, P., and Orlowski, S. (2009). Interaction of macrocyclic lactones with the multidrug transporters: the bases of the pharmacokinetics of lipid-like drugs. Current drug metabolism *10*, 272-288.

Lespine, A., Dupuy, J., Orlowski, S., Nagy, T., Glavinas, H., Krajcsi, P., and Alvinerie, M. (2006b). Interaction of ivermectin with multidrug resistance proteins (MRP1, 2 and 3). Chem Biol Interact *159*, 169-179.

Lespine, A., Martin, S., Dupuy, J., Roulet, A., Pineau, T., Orlowski, S., and Alvinerie, M. (2007). Interaction of macrocyclic lactones with P-glycoprotein: structure-affinity relationship. Eur J Pharm Sci *30*, 84-94.

Lespine, A., Menez, C., Bourguinat, C., and Prichard, R.K. (2012). P-glycoproteins and other multidrug resistance transporters in the pharmacology of anthelmintics: Prospects for reversing transport-dependent anthelmintic resistance. Int J Parasitol Drugs Drug Resist *2*, 58-75.

Lewis, H.A., Buchanan, S.G., Burley, S.K., Conners, K., Dickey, M., Dorwart, M., Fowler, R., Gao, X., Guggino, W.B., Hendrickson, W.A., *et al.* (2004). Structure of nucleotide-binding domain 1 of the cystic fibrosis transmembrane conductance regulator. The EMBO journal *23*, 282-293.

Li, J., Ashton, F.T., Gamble, H.R., and Schad, G.A. (2000a). Sensory neuroanatomy of a passively ingested nematode parasite, Haemonchus contortus: amphidial neurons of the first stage larva. The Journal of comparative neurology *417*, 299-314.

Li, J., Jaimes, K.F., and Aller, S.G. (2013). Refined structures of mouse P-glycoprotein. Protein Sci *23*, 34-46.

Li, J., Zhu, X., Ashton, F.T., Gamble, H.R., and Schad, G.A. (2001). Sensory neuroanatomy of a passively ingested nematode parasite, Haemonchus contortus: amphidial neurons of the third-stage larva. J Parasitol *87*, 65-72.

Li, J., Zhu, X., Boston, R., Ashton, F.T., Gamble, H.R., and Schad, G.A. (2000b). Thermotaxis and thermosensory neurons in infective larvae of Haemonchus contortus, a passively ingested nematode parasite. The Journal of comparative neurology *424*, 58-73.

Lifschitz, A., Entrocasso, C., Alvarez, L., Lloberas, M., Ballent, M., Manazza, G., Virkel, G., Borda, B., and Lanusse, C. (2010). Interference with P-glycoprotein improves ivermectin activity against adult resistant nematodes in sheep. Vet Parasitol *172*, 291-298.

Lifschitz, A., Virkel, G., Ballent, M., Sallovitz, J., and Lanusse, C. (2009). Combined use of ivermectin and triclabendazole in sheep: in vitro and in vivo characterisation of their pharmacological interaction. Vet J *182*, 261-268.

Lifschitz, A., Virkel, G., Ballent, M., Sallovitz, J., Pis, A., and Lanusse, C. (2005). Moxidectin and ivermectin metabolic stability in sheep ruminal and abomasal contents. J Vet Pharmacol Ther *28*, 411-418.

Lifschitz, A., Virkel, G., Sallovitz, J., Imperiale, F., Pis, A., and Lanusse, C. (2002). Loperamide-induced enhancement of moxidectin availability in cattle. J Vet Pharmacol Ther *25*, 111-120.

Lifschitz, A.L., Virkel, G.L., Sallovitz, J.M., Pis, A., Imperiale, F.A., and Lanusse, C.E. (2004). Loperamide modifies the tissue disposition kinetics of ivermectin in rats. J Pharm Pharmacol *56*, 61-67.

Lin, D.Y., Huang, S., and Chen, J. (2015). Crystal structures of a polypeptide processing and secretion transporter. Nature *523*, 425-430.

Lincke, C.R., Broeks, A., The, I., Plasterk, R.H., and Borst, P. (1993). The expression of two P-glycoprotein (pgp) genes in transgenic Caenorhabditis elegans is confined to intestinal cells. EMBO J *12*, 1615-1620.

Lincke, C.R., The, I., van Groenigen, M., and Borst, P. (1992). The P-glycoprotein gene family of Caenorhabditis elegans. Cloning and characterization of genomic and complementary DNA sequences. Journal of molecular biology *228*, 701-711.

Lloberas, M., Alvarez, L., Entrocasso, C., Virkel, G., Ballent, M., Mate, L., Lanusse, C., and Lifschitz, A. (2013). Comparative tissue pharmacokinetics and efficacy of moxidectin, abamectin and ivermectin in lambs infected with resistant nematodes: Impact of drug treatments on parasite P-glycoprotein expression. Int J Parasitol Drugs Drug Resist *3*, 20-27.

Locher, K.P., Lee, A.T., and Rees, D.C. (2002). The E. coli BtuCD structure: a framework for ABC transporter architecture and mechanism. Science *296*, 1091-1098.

Loo, T.W., Bartlett, M.C., and Clarke, D.M. (2003). Methanethiosulfonate derivatives of rhodamine and verapamil activate human P-glycoprotein at different sites. J Biol Chem *278*, 50136-50141.

Loo, T.W., Bartlett, M.C., and Clarke, D.M. (2006a). Transmembrane segment 1 of human P-glycoprotein contributes to the drug-binding pocket. Biochem J *396*, 537-545.

Loo, T.W., Bartlett, M.C., and Clarke, D.M. (2006b). Transmembrane segment 7 of human P-glycoprotein forms part of the drug-binding pocket. Biochem J *399*, 351-359.

Loo, T.W., and Clarke, D.M. (1997). Correction of defective protein kinesis of human P-glycoprotein mutants by substrates and modulators. J Biol Chem *272*, 709-712.

Loo, T.W., and Clarke, D.M. (2001). Defining the drug-binding site in the human multidrug resistance P-glycoprotein using a methanethiosulfonate analog of verapamil, MTS-verapamil. J Biol Chem *276*, 14972-14979.

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

Lubega, G.W., and Prichard, R.K. (1991). Interaction of benzimidazole anthelmintics with Haemonchus contortus tubulin: binding affinity and anthelmintic efficacy. Exp Parasitol *73*, 203-213.

Mani, T., Bourguinat, C., Keller, K., Ashraf, S., Blagburn, B., and Prichard, R.K. (2016). Interaction of macrocyclic lactones with a Dirofilaria immitis P-glycoprotein. . Int J Parasitol.

Martin, R.J., Robertson, A.P., Buxton, S.K., Beech, R.N., Charvet, C.L., and Neveu, C. (2012). Levamisole receptors: a second awakening. Trends Parasitol *28*, 289-296.

McCavera, S., Rogers, A.T., Yates, D.M., Woods, D.J., and Wolstenholme, A.J. (2009). An ivermectin-sensitive glutamate-gated chloride channel from the parasitic nematode Haemonchus contortus. Mol Pharmacol *75*, 1347-1355.

Meaney, M., Savage, J., Brennan, G.P., Hoey, E., Trudgett, A., and Fairweather, I. (2013). Increased susceptibility of a triclabendazole (TCBZ)-resistant isolate of Fasciola hepatica to TCBZ following co-incubation in vitro with the P-glycoprotein inhibitor, R(+)-verapamil. Parasitology *140*, 1287-1303.

Menez, C., Alberich, M., Kansoh, D., Blanchard, A., and Lespine, A. (2016). Acquired tolerance to ivermectin and moxidectin after drug selection pressure in the nematode Caenorhabditis elegans. Antimicrob Agents Chemother.

Miller, J.E., Bahirathan, M., Lemarie, S.L., Hembry, F.G., Kearney, M.T., and Barras, S.R. (1998). Epidemiology of gastrointestinal nematode parasitism in Suffolk and Gulf Coast Native sheep with special emphasis on relative susceptibility to Haemonchus contortus infection. Vet Parasitol *74*, 55-74.

Moeller, A., Lee, S.C., Tao, H., Speir, J.A., Chang, G., Urbatsch, I.L., Potter, C.S., Carragher, B., and Zhang, Q. (2015). Distinct conformational spectrum of homologous multidrug ABC transporters. Structure *23*, 450-460.

Molento, M.B., Lifschitz, A., Sallovitz, J., Lanusse, C., and Prichard, R. (2004). Influence of verapamil on the pharmacokinetics of the antiparasitic drugs ivermectin and moxidectin in sheep. Parasitology research *92*, 121-127.

Molento, M.B., and Prichard, R.K. (1999). Effects of the multidrug-resistance-reversing agents verapamil and CL 347,099 on the efficacy of ivermectin or moxidectin against unselected and drug-selected strains of Haemonchus contortus in jirds (Meriones unguiculatus). Parasitology research *85*, 1007-1011.

Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., and Olson, A.J. (1998). Automated docking using lamarckian genetic algorithm and an empirical binding free energy function. J Comput Chem *19*, 1639-1662.

Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., and Olson, A.J. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J Comput Chem *30*, 2785-2791.

Mottier, L., Alvarez, L., Fairweather, I., and Lanusse, C. (2006). Resistance-induced changes in triclabendazole transport in Fasciola hepatica: ivermectin reversal effect. J Parasitol *92*, 1355-1360.

Mottier, M.L., and Prichard, R.K. (2008). Genetic analysis of a relationship between macrocyclic lactone and benzimidazole anthelmintic selection on Haemonchus contortus. Pharmacogenetics and Genomics *18*, 129-140.

Nare, B., Liu, Z., Prichard, R.K., and Georges, E. (1994). Benzimidazoles, potent anti-mitotic drugs: substrates for the P-glycoprotein transporter in multidrug-resistant cells. Biochemical pharmacology *48*, 2215-2222.

Neveu-Lemaire, M. (1936). Traité d'Helminthologie Médicale et Vétérinaire. Vigot Frères, Paris.

Neveu, C., Charvet, C., Fauvin, A., Cortet, J., Castagnone-Sereno, P., and Cabaret, J. (2007). Identification of levamisole resistance markers in the parasitic nematode Haemonchus contortus using a cDNA-AFLP approach. Parasitology *134*, 1105-1110.

Njue, A.I., Hayashi, J., Kinne, L., Feng, X.P., and Prichard, R.K. (2004). Mutations in the extracellular domains of glutamate-gated chloride channel alpha3 and beta subunits from ivermectin-resistant Cooperia oncophora affect agonist sensitivity. Journal of neurochemistry *89*, 1137-1147.

Nobili, S., Landini, I., Giglioni, B., and Mini, E. (2006). Pharmacological strategies for overcoming multidrug resistance. Current drug targets *7*, 861-879.

Nunes, F., Wolf, M., Hartmann, J., and Paul, R.J. (2005). The ABC transporter PGP-2 from Caenorhabditis elegans is expressed in the sensory neuron pair AWA and contributes to lysosome formation and lipid storage within the intestine. Biochem Biophys Res Commun *338*, 862-871.

Orlowski, S., Comera, C., Terce, F., and Collet, X. (2007). Lipid rafts: dream or reality for cholesterol transporters? Eur Biophys J *36*, 869-885.

Perez, C., Gerber, S., Boilevin, J., Bucher, M., Darbre, T., Aebi, M., Reymond, J.L., and Locher, K.P. (2015). Structure and mechanism of an active lipid-linked oligosaccharide flippase. Nature *524*, 433-438.

Perez, M., Blazquez, A.G., Real, R., Mendoza, G., Prieto, J.G., Merino, G., and Alvarez, A.I. (2009). In vitro and in vivo interaction of moxidectin with BCRP/ABCG2. Chem Biol Interact *180*, 106-112.

Portillo, V., Jagannathan, S., and Wolstenholme, A.J. (2003). Distribution of glutamate-gated chloride channel subunits in the parasitic nematode Haemonchus contortus. The Journal of comparative neurology *462*, 213-222.

Pouliot, J.F., L'Heureux, F., Liu, Z., Prichard, R.K., and Georges, E. (1997). Reversal of P-glycoprotein-associated multidrug resistance by ivermectin. Biochemical pharmacology *53*, 17-25.

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

Prajapati, R., Singh, U., Patil, A., Khomane, K.S., Bagul, P., Bansal, A.K., and Sangamwar, A.T. (2013). In silico model for P-glycoprotein substrate prediction: insights from molecular dynamics and in vitro studies. J Comput Aided Mol Des *27*, 347-363.

Prichard, R.K., and Roulet, A. (2007). ABC transporters and beta-tubulin in macrocyclic lactone resistance: prospects for marker development. Parasitology *134*, 1123-1132.

Qian, H., Martin, R.J., and Robertson, A.P. (2006). Pharmacology of N-, L-, and B-subtypes of nematode nAChR resolved at the single-channel level in Ascaris suum. FASEB journal : official publication of the Federation of American Societies for Experimental Biology *20*, 2606-2608.

Raj, A., van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A., and Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. Nat Methods *5*, 877-879.

Real, R., Egido, E., Perez, M., Gonzalez-Lobato, L., Barrera, B., Prieto, J.G., Alvarez, A.I., and Merino, G. (2011). Involvement of breast cancer resistance protein (BCRP/ABCG2) in the secretion of danofloxacin into milk: interaction with ivermectin. J Vet Pharmacol Ther *34*, 313-321.

Rice, A.J., Park, A., and Pinkett, H.W. (2014). Diversity in ABC transporters: type I, II and III importers. Crit Rev Biochem Mol Biol *49*, 426-437.

Riou, M., Guegnard, F., Le Vern, Y., and Kerboeuf, D. (2003). Modulation of the multidrug resistance (MDR) system in the nematode Haemonchus contortus by changing cholesterol content: effects on resistance to anthelmintics. The Journal of antimicrobial chemotherapy *52*, 180-187.

Riou, M., Koch, C., Delaleu, B., Berthon, P., and Kerboeuf, D. (2005). Immunolocalisation of an ABC transporter, P-glycoprotein, in the eggshells and cuticles of free-living and parasitic stages of Haemonchus contortus. Parasitology research *96*, 142-148.

Robertson, B. (1989). Actions of anaesthetics and avermectin on GABAA chloride channels in mammalian dorsal root ganglion neurones. British journal of pharmacology *98*, 167-176.

Roos, M.H., Kwa, M.S., Veenstra, J.G., Kooyman, F.N., and Boersema, J.H. (1993). Molecular aspects of drug resistance in parasitic helminths. Pharmacology & therapeutics *60*, 331-336.

Rosenberg, M.F., Callaghan, R., Ford, R.C., and Higgins, C.F. (1997). Structure of the multidrug resistance P-glycoprotein to 2.5 nm resolution determined by electron microscopy and image analysis. J Biol Chem *272*, 10685-10694.

Rosenberg, M.F., Callaghan, R., Modok, S., Higgins, C.F., and Ford, R.C. (2005). Three-dimensional structure of P-glycoprotein: the transmembrane regions adopt an asymmetric configuration in the nucleotide-bound state. J Biol Chem *280*, 2857-2862.

Roulet, A., Puel, O., Gesta, S., Lepage, J.F., Drag, M., Soll, M., Alvinerie, M., and Pineau, T. (2003). MDR1-deficient genotype in Collie dogs hypersensitive to the P-glycoprotein substrate ivermectin. Eur J Pharmacol *460*, 85-91.

Rufener, L., Baur, R., Kaminsky, R., Maser, P., and Sigel, E. (2010). Monepantel allosterically activates DEG-3/DES-2 channels of the gastrointestinal nematode Haemonchus contortus. Mol Pharmacol *78*, 895-902.

Rufener, L., Maser, P., Roditi, I., and Kaminsky, R. (2009). Haemonchus contortus acetylcholine receptors of the DEG-3 subfamily and their role in sensitivity to monepantel. PLoS Pathog *5*, e1000380.

Saddiqi, H.A., Jabbar, A., Sarwar, M., Iqbal, Z., Muhammad, G., Nisa, M., and Shahzad, A. (2011). Small ruminant resistance against gastrointestinal nematodes: a case of Haemonchus contortus. Parasitology research *109*, 1483-1500.

Sangster, N.C., Bannan, S.C., Weiss, A.S., Nulf, S.C., Klein, R.D., and Geary, T.G. (1999). Haemonchus contortus: sequence heterogeneity of internucleotide binding domains from P-glycoproteins. Exp Parasitol *91*, 250-257.

Sarkadi, B., Homolya, L., Szakacs, G., and Varadi, A. (2006). Human multidrug resistance ABCB and ABCG transporters: participation in a chemoimmunity defense system. Physiol Rev *86*, 1179-1236.

Sarkadi, B., Price, E.M., Boucher, R.C., Germann, U.A., and Scarborough, G.A. (1992). Expression of the human multidrug resistance cDNA in insect cells generates a high activity drug-stimulated membrane ATPase. J Biol Chem *267*, 4854-4858.

Sasaki, T., Takagi, M., Yaguchi, T., Miyadoh, S., Okada, T., and Koyama, M. (1992). A new anthelmintic cyclodepsipeptide, PF1022A. J Antibiot (Tokyo) *45*, 692-697.

Sauna, Z.E., and Ambudkar, S.V. (2007). About a switch: how P-glycoprotein (ABCB1) harnesses the energy of ATP binding and hydrolysis to do mechanical work. Mol Cancer Ther *6*, 13-23.

Savage, J., Meaney, M., Brennan, G.P., Hoey, E., Trudgett, A., and Fairweather, I. (2013a). Effect of the P-glycoprotein inhibitor, R(+)-verapamil on the drug susceptibility of a triclabendazole-resistant isolate of Fasciola hepatica. Vet Parasitol *195*, 72-86.

Savage, J., Meaney, M., Brennan, G.P., Hoey, E., Trudgett, A., and Fairweather, I. (2013b). Increased action of triclabendazole (TCBZ) in vitro against a TCBZ-resistant isolate of Fasciola hepatica following its co-incubation with the P-glycoprotein inhibitor, R(+)-verapamil. Exp Parasitol *135*, 642-653.

Savage, J., Meaney, M., Brennan, G.P., Hoey, E., Trudgett, A., and Fairweather, I. (2014). Disruption of vitellogenesis and spermatogenesis by triclabendazole (TCBZ) in a TCBZ-resistant isolate of Fasciola hepatica following incubation in vitro with a P-glycoprotein inhibitor. Parasitology *141*, 1064-1079.

Schinkel, A.H. (1997). The physiological function of drug-transporting P-glycoproteins. Semin Cancer Biol *8*, 161-170.

Schinkel, A.H., Smit, J.J., van Tellingen, O., Beijnen, J.H., Wagenaar, E., van Deemter, L., Mol, C.A., van der Valk, M.A., Robanus-Maandag, E.C., te Riele, H.P., *et al.* (1994). Disruption of the mouse mdr1a P-glycoprotein gene leads to a deficiency in the blood-brain barrier and to increased sensitivity to drugs. Cell *77*, 491-502.

Schmitt, L., Benabdelhak, H., Blight, M.A., Holland, I.B., and Stubbs, M.T. (2003). Crystal structure of the nucleotide-binding domain of the ABC-transporter haemolysin B:

identification of a variable region within ABC helical domains. Journal of molecular biology *330*, 333-342.

Schroeder, L.K., Kremer, S., Kramer, M.J., Currie, E., Kwan, E., Watts, J.L., Lawrenson, A.L., and Hermann, G.J. (2007). Function of the Caenorhabditis elegans ABC transporter PGP-2 in the biogenesis of a lysosome-related fat storage organelle. Mol Biol Cell *18*, 995-1008.

Schulz, S., Schmitt, S., Wimmer, R., Aichler, M., Eisenhofer, S., Lichtmannegger, J., Eberhagen, C., Artmann, R., Tookos, F., Walch, A., *et al.* (2013). Progressive stages of mitochondrial destruction caused by cell toxic bile salts. Biochim Biophys Acta *1828*, 2121-2133.

Schwarz, E.M., Korhonen, P.K., Campbell, B.E., Young, N.D., Jex, A.R., Jabbar, A., Hall, R.S., Mondal, A., Howe, A.C., Pell, J., *et al.* (2013). The genome and developmental transcriptome of the strongylid nematode Haemonchus contortus. Genome Biol *14*, R89.

Seelig, A., and Gerebtzoff, G. (2006). Enhancement of drug absorption by noncharged detergents through membrane and P-glycoprotein binding. Expert opinion on drug metabolism & toxicology *2*, 733-752.

Sheps, J.A., Ralph, S., Zhao, Z., Baillie, D.L., and Ling, V. (2004). The ABC transporter gene family of Caenorhabditis elegans has implications for the evolutionary dynamics of multidrug resistance in eukaryotes. Genome Biol *5*, R15.

Shilling, R.A., Venter, H., Velamakanni, S., Bapna, A., Woebking, B., Shahi, S., and van Veen, H.W. (2006). New light on multidrug binding by an ATP-binding-cassette transporter. Trends Pharmacol Sci *27*, 195-203.

Shintre, C.A., Pike, A.C., Li, Q., Kim, J.I., Barr, A.J., Goubin, S., Shrestha, L., Yang, J., Berridge, G., Ross, J., *et al.* (2013). Structures of ABCB10, a human ATP-binding cassette transporter in apo- and nucleotide-bound states. Proc Natl Acad Sci U S A *110*, 9710-9715.

Silvestre, A., and Cabaret, J. (2002). Mutation in position 167 of isotype 1 beta-tubulin gene of Trichostrongylid nematodes: role in benzimidazole resistance? Mol Biochem Parasitol *120*, 297-300.

Smith, H., and Campbell, W.C. (1996). Effect of ivermectin on Caenorhabditis elegans larvae previously exposed to alcoholic immobilization. J Parasitol *82*, 187-188.

Smith, J.M., and Prichard, R.K. (2002). Localization of p-glycoprotein mRNA in the tissues of Haemonchus contortus adult worms and its relative abundance in drug-selected and susceptible strains. J Parasitol *88*, 612-620.

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

Speel, E.J., Hopman, A.H., and Komminoth, P. (2006). Tyramide signal amplification for DNA and mRNA in situ hybridization. Methods Mol Biol *326*, 33-60.

Srinivasan, V., Pierik, A.J., and Lill, R. (2014). Crystal structures of nucleotide-free and glutathione-bound mitochondrial ABC transporter Atm1. Science *343*, 1137-1140.

Starich, T.A., Herman, R.K., Kari, C.K., Yeh, W.H., Schackwitz, W.S., Schuyler, M.W., Collet, J., Thomas, J.H., and Riddle, D.L. (1995). Mutations affecting the chemosensory neurons of Caenorhabditis elegans. Genetics *139*, 171-188.

Staudinger, J.L., Madan, A., Carol, K.M., and Parkinson, A. (2003). Regulation of drug transporter gene expression by nuclear receptors. Drug Metab Dispos *31*, 523-527.

Swan, G.E., Koeleman, H.A., Steyn, H.S., and Mulders, M.S. (1999). Intravascular plasma disposition and salivary secretion of closantel and rafoxanide in sheep. J S Afr Vet Assoc *70*, 75-79.

Szabo, K., Welker, E., Bakos, Muller, M., Roninson, I., Varadi, A., and Sarkadi, B. (1998). Drug-stimulated nucleotide trapping in the human multidrug transporter MDR1. Cooperation of the nucleotide binding domains. J Biol Chem *273*, 10132-10138.

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

Szakacs, G., Varadi, A., Ozvegy-Laczka, C., and Sarkadi, B. (2008). The role of ABC transporters in drug absorption, distribution, metabolism, excretion and toxicity (ADME-Tox). Drug Discov Today *13*, 379-393.

Szewczyk, P., Tao, H., McGrath, A.P., Villaluz, M., Rees, S.D., Lee, S.C., Doshi, R., Urbatsch, I.L., Zhang, Q., and Chang, G. (2015). Snapshots of ligand entry, malleable binding and induced helical movement in P-glycoprotein. Acta Crystallogr D Biol Crystallogr *71*, 732-741.

Thiebaut, F., Tsuruo, T., Hamada, H., Gottesman, M.M., Pastan, I., and Willingham, M.C. (1987). Cellular localization of the multidrug-resistance gene product P-glycoprotein in normal human tissues. Proc Natl Acad Sci U S A *84*, 7735-7738.

Thomaz-Soccol, V., Souza, F.P., Sotomaior, C., Castro, E.A., Milczewski, V., Mocelin, G., and Silva, M.C.P. (2004). Resistance of gastrointestinal nematodes to anthelmintics in sheep (Ovis aries). Brazilian Archives of Biology and Technology *47*, 41-47.

Tombline, G., Bartholomew, L.A., Tyndall, G.A., Gimi, K., Urbatsch, I.L., and Senior, A.E. (2004a). Properties of P-glycoprotein with mutations in the "catalytic carboxylate" glutamate residues. J Biol Chem *279*, 46518-46526.

Tombline, G., Bartholomew, L.A., Urbatsch, I.L., and Senior, A.E. (2004b). Combined mutation of catalytic glutamate residues in the two nucleotide binding domains of P-glycoprotein generates a conformation that binds ATP and ADP tightly. J Biol Chem *279*, 31212-31220.

Tombline, G., Muharemagic, A., White, L.B., and Senior, A.E. (2005). Involvement of the "occluded nucleotide conformation" of P-glycoprotein in the catalytic pathway. Biochemistry *44*, 12879-12886.

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

Urdaneta-Marquez, L., Bae, S.H., Janukavicius, P., Beech, R., Dent, J., and Prichard, R. (2014). A dyf-7 haplotype causes sensory neuron defects and is associated with macrocyclic lactone resistance worldwide in the nematode parasite Haemonchus contortus. Int J Parasitol *44*, 1063-1071.

Van Den Bossche, H., Verhoeven, H., Vanparijs, O., Lauwers, H., and Thienpont, D. (1979). Closantel, a new antiparasitic hydrogen ionophore [proceedings]. Arch Int Physiol Biochim *87*, 851-853.

Van den Brom, R., Moll, L., Kappert, C., and Vellema, P. (2015). Haemonchus contortus resistance to monepantel in sheep. Vet Parasitol *209*, 278-280.

Vasiliou, V., Vasiliou, K., and Nebert, D.W. (2009). Human ATP-binding cassette (ABC) transporter family. Hum Genomics *3*, 281-290.

Verdon, G., Albers, S.V., Dijkstra, B.W., Driessen, A.J., and Thunnissen, A.M. (2003a). Crystal structures of the ATPase subunit of the glucose ABC transporter from Sulfolobus solfataricus: nucleotide-free and nucleotide-bound conformations. Journal of molecular biology *330*, 343-358.

Verdon, G., Albers, S.V., van Oosterwijk, N., Dijkstra, B.W., Driessen, A.J., and Thunnissen, A.M. (2003b). Formation of the productive ATP-Mg2+-bound dimer of GlcV, an ABC-ATPase from Sulfolobus solfataricus. Journal of molecular biology *334*, 255-267.

Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. The EMBO journal *1*, 945-951.

Wang, B., Dukarevich, M., Sun, E.I., Yen, M.R., and Saier, M.H., Jr. (2009). Membrane porters of ATP-binding cassette transport systems are polyphyletic. J Membr Biol *231*, 1-10.

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

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

Ward, S., Thomson, N., White, J.G., and Brenner, S. (1975). Electron microscopical reconstruction of the anterior sensory anatomy of the nematode Caenorhabditis elegans.?2UU. The Journal of comparative neurology *160*, 313-337.

White, J.G., Southgate, E., Thomson, J.N., and Brenner, S. (1986). The structure of the nervous system of the nematode Caenorhabditis elegans. Philos Trans R Soc Lond B Biol Sci *314*, 1-340.

Wilkinson, R., Law, C.J., Hoey, E.M., Fairweather, I., Brennan, G.P., and Trudgett, A. (2012). An amino acid substitution in Fasciola hepatica P-glycoprotein from triclabendazole-resistant and triclabendazole-susceptible populations. Mol Biochem Parasitol *186*, 69-72.

Williamson, S.M., Storey, B., Howell, S., Harper, K.M., Kaplan, R.M., and Wolstenholme, A.J. (2011). Candidate anthelmintic resistance-associated gene expression and sequence polymorphisms in a triple-resistant field isolate of Haemonchus contortus. Mol Biochem Parasitol *180*, 99-105.

Williamson, S.M., and Wolstenholme, A.J. (2012). P-glycoproteins of Haemonchus contortus: development of real-time PCR assays for gene expression studies. Journal of helminthology *86*, 202-208.

Wolstenholme, A.J., and Rogers, A.T. (2005). Glutamate-gated chloride channels and the mode of action of the avermectin/milbemycin anthelmintics. Parasitology *131 Suppl*, S85-95.

Xu, M., Molento, M., Blackhall, W., Ribeiro, P., Beech, R., and Prichard, R. (1998). Ivermectin resistance in nematodes may be caused by alteration of P-glycoprotein homolog. Mol Biochem Parasitol *91*, 327-335.

Yan, R., Urdaneta-Marquez, L., Keller, K., James, C.E., Davey, M.W., and Prichard, R.K. (2012). The role of several ABC transporter genes in ivermectin resistance in Caenorhabditis elegans. Vet Parasitol *190*, 519-529.

Yates, D.M., Portillo, V., and Wolstenholme, A.J. (2003). The avermectin receptors of Haemonchus contortus and Caenorhabditis elegans. Int J Parasitol *33*, 1183-1193.

Zaitseva, J., Jenewein, S., Jumpertz, T., Holland, I.B., and Schmitt, L. (2005). H662 is the linchpin of ATP hydrolysis in the nucleotide-binding domain of the ABC transporter HlyB. The EMBO journal *24*, 1901-1910.

Zeng, Z., Andrew, N.W., Arison, B.H., Luffer-Atlas, D., and Wang, R.W. (1998). Identification of cytochrome P4503A4 as the major enzyme responsible for the metabolism of ivermectin by human liver microsomes. Xenobiotica *28*, 313-321.

Zeng, Z., Andrew, N.W., Green-Erwin, M.L., and Halley, B.A. (1996). Fate of 4"-epiacetylamino-4"-deoxyavermectin B1 in rats. Drug Metab Dispos *24*, 572-578.

Zeng, Z., Andrew, N.W., and Halley, B.A. (1997). Identification of cytochrome P4503A as the major enzyme sub-family responsible for the metabolism of 22,23-dihydro-13-O-[(2-methoxyethoxy)methyl]-avermectin B1 aglycone by rat liver microsomes. Xenobiotica *27*, 985-994.

Zhao, Z., Sheps, J.A., Ling, V., Fang, L.L., and Baillie, D.L. (2004). Expression analysis of ABC transporters reveals differential functions of tandemly duplicated genes in Caenorhabditis elegans. Journal of molecular biology *344*, 409-417.

TITLE: Identification and functional characterization of an ABC transporter of *Haemonchus contortus*, the P-glycoprotein 13

SUMMARY:

Macrocyclic lactones (ML) are paralyzing anthelmintics used in animals and humans against parasite nematodes. However, their therapeutic success is compromised by the spread of ML resistance. This might be at least partly due to P-glycoproteins (Pgps) ABC transporters that are selected and overexpressed in ML-resistant nematodes. Deciphering the role of the 10 Pgps expressed in the parasite of small ruminants *Haemonchus contortus* is thus of major importance to guaranty anthelmintic (AH) efficacy of various drugs. Here we focused on Hco-Pgp-13 due to the expression in the amphids of its closest ortholog in the model nematode *C. elegans*. Indeed, the amphids represent a putative entry route of drugs to reach AH targets in the nervous system and have been linked to AH susceptibility in *C. elegans* and *H. contortus*.

In order to predict the capacity of nematode Pgps to transport drugs, including ML and otherAH, we have developed an *in silico* drug docking model. We have used *C. elegans* Pgp-1 (Cel-Pgp-1) crystal structure and have showed a high affinity binding of several ligands that have been shown to be activators of its ATPase function. ML were also found to bind with high affinity to Cel-Pgp-1, on a specific binding site. This approach provides a valuable tool to predict drug-drug interactions and to rationally design new competitive inhibitors of nematode Pgps, in order to improve anthelmintic therapeutics.

We then predicted a putative 3D structure of Hco-Pgp-13 based on the recently released crystal of Cel-Pgp-1, with which it presented a high homology. This allowed the study of the interaction of Hco-Pgp-13 with potential substrates, in particular ML. We found similar affinities for various drugs previously tested on Cel-Pgp-1, supporting the good homology of these two proteins. Together with *in vitro* ATPase assay experiments that confirmed the substrate status of actinomycin D, this indicates a possible multispecifc recognition capacity of this parasitic transporter.

The determination of Hco-Pgp-13 localization using immunohistochemistry showed a wide tissue expression consistent with a critical role for Hco-Pgp-13 in endogenous and/or exogenous substrate transport.

In conclusion, this work provides insights into the role of nematode Pgps in transporting AH drugs, both at the level of the model organism *C. elegans* and of the parasitic nematode *H. contortus*. This suggests a high homology of function conserved between ABC tranporters in these species. The localization of such protein on amphidial structures and its possible involvement in drug resistance and survival of *H. contortus* to exposure to AH compounds remain to be precised.

KEYWORDS: ABC transporter; P-glycoprotein, *Haemonchus contortus*, *Caenorhabditis elegans*; nematodes; anthelmintics; macrocyclic lactones; *in silico* docking; multispecific recognition; substrate transport; multidrug resistance, amphids.

AUTEUR : Marion DAVID

TITRE : Identification et caracterisation fonctionelle d'un transporteur ABC de *Haemonchus contortus*, la P-glycoprotein 13

DIRECTEURS DE THESE : Dr Anne LESPINE et Pr Roger K. PRICHARD

LIEU ET DATE DE SOUTENANCE : Toulouse, le 14 Octobre 2016

RESUME :

Les lactones macrocycliques (LM) sont des anthelminthiques (AH) à effet paralysant très utilisés chez les animaux et les humains contre les nématodes parasites. Cependant, leur succès thérapeutique est compromis par la résistance croissante aux LM, qui pourrait être en partie dû aux ABC transporteurs P-glycoprotéines (Pgps) sélectionnés et surexprimés chez les nématodes résistants aux LM. Dans ce travail, nous avons étudié plus précisément la P-glycoprotéine 13 du parasite de petits ruminants, *Haemonchus contortus*. Son orthologue chez le modèle nématode *C. elegans*, Cel-Pgp-13, est exprimé dans les amphides, structures qui ont été associées à la sensibilité aux AH chez *C. elegans* et *H. contortus*.

Pour prédire la capacité des Pgps de nematode à transporter des drogues, incluant des LM et autres AH, nous avons développé un modèle de docking *in silico*. Nous avons utilisé la structure cristallographique de *C. elegans* Pgp-1 (Cel-Pgp-1), et nous avons montré la liaison avec une forte affinité de plusieurs ligands décrits comme activateurs de sa fonction ATPasique. Nous avons aussi décrit une forte affinité des LM, et un site spécifique de liaison de ces composés à Cel-Pgp-1. Cette approche représente un outil important pour prédire les interactions entre AH, et pour concevoir rationnellement de nouveaux inhibiteurs compétitifs des Pgps de nématode, dans le but d'améliorer les stratégies thérapeutiques.

Sur la base de cette approche, nous avons prédit la structure 3D de Hco-Pgp-13 à partir du cristal de Cel-Pgp-1 afin d'étudier son intéraction avec des substrats potentiels, en particulier les LM. Nous avons trouvé des affinités similaires pour différents composés précédemment testés sur Cel-Pgp-1. *In vitro*, la mesure de l'activité ATPasique montre que l'actinomycine D est un substrat de Hco-Pgp-13. Nos données démontrent la présence possible d'un domaine de reconnaissance multispécifique sur ce transporteur de parasite.

La détermination par immunofluorescence de l'expression de Hco-Pgp-13 a montré une distribution tissulaire large indiquant que Hco-Pgp-13 pourrait jouer un role important dans le transport de substrats endogènes et/ou exogènes.

En conclusion, ce travail permet de mieux comprendre le rôle des Pgps de nématodes dans le transport de médicaments AH, tant au niveau de l'organisme modèle *C. elegans* que du nématode parasite *H. contortus*. Cette étude suggère la conservation de la fonction de tranporteur ABC multidrogue dans ces espèces. La localisation de Hco-Pgp-13 sur les structures amphidiales, et son éventuelle implication dans la résistance aux médicaments et à la survie de *H. contortus* à l'exposition à des composés AH, restent à préciser.

MOTS-CLES : transporteur ABC ; P-glycoprotéine, *Haemonchus contortus, Caenorhabditis elegans*; nématodes; anthelmintiques; lactones macrocycliques; docking *in silico*; reconnaissance multispecifique; transport de substrats; résistance multidrogue; amphides.

DISCIPLINE ADMINISTRATIVE : Pathologie, toxicologie, génétique et nutrition.

INTITULE ET ADRESSE DE L'U.F.R. OU DU LABORATOIRE : Equipe Transporteurs Membranaires et Résistance, UMR 1331, TOXALIM, 180 chemin de Tournefeuille, BP 93173 – 31027 Toulouse Cedex 3