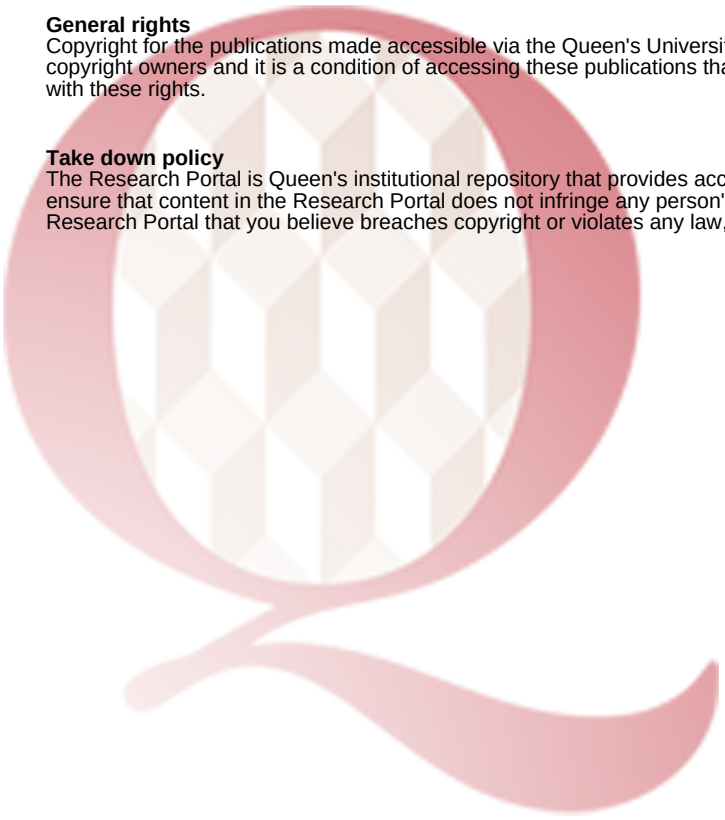
An amino acid substitution in Fasciola hepatica P-glycoprotein from triclabendazole-resistant and triclabendazole-susceptible populations

Wilkinson, R., Law, C. J., Hoey, E. M., Fairweather, I., Brennan, G. P., & Trudgett, A. (2012). An amino acid substitution in Fasciola hepatica P-glycoprotein from triclabendazole-resistant and triclabendazole-susceptible populations. Molecular and Biochemical Parasitology, 186(1), 69-72. DOI: 10.1016/j.molbiopara.2012.08.008

Published in:
Molecular and Biochemical Parasitology

Document Version:
Peer reviewed version



Title: An amino acid substitution in *Fasciola hepatica* P-glycoprotein from triclabendazole-resistant and triclabendazole-susceptible populations

Authors: Richard Wilkinson, Christopher J Law, Elizabeth M Hoey, Ian Fairweather, Gerard P Brennan, Alan Trudgett

PII: S0166-6851(12)00218-6
DOI: doi:10.1016/j.molbiopara.2012.08.008
Reference: MOLBIO 10702

To appear in: *Molecular & Biochemical Parasitology*

Received date: 9-3-2012
Revised date: 31-8-2012
Accepted date: 31-8-2012

Please cite this article as: Wilkinson R, Law CJ, Hoey EM, Fairweather I, Brennan GP, Trudgett A, An amino acid substitution in *Fasciola hepatica* P-glycoprotein from triclabendazole-resistant and triclabendazole-susceptible populations, *Molecular & Biochemical Parasitology* (2010), doi:10.1016/j.molbiopara.2012.08.008

An amino acid substitution in *Fasciola hepatica* P-glycoprotein from triclabendazole-resistant and triclabendazole-susceptible populations.

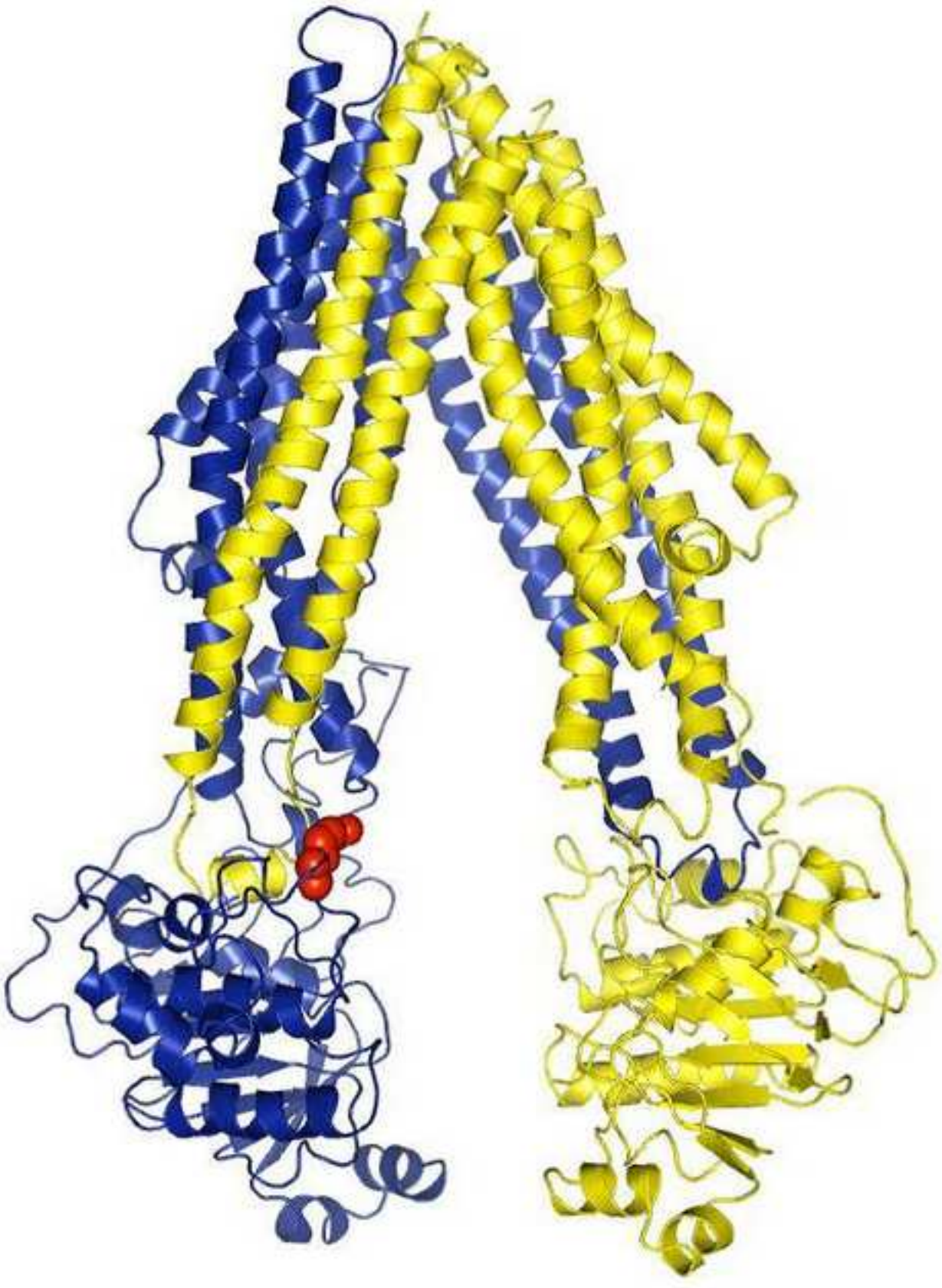
Richard Wilkinson[a], Christopher J Law[a], Elizabeth M Hoey[a], Ian Fairweather[a], Gerard P Brennan[a], Alan Trudgett[a]*

[a]School of Biological Sciences, The Queen's University of Belfast

*Corresponding author: Alan Trudgett, School of Biological Sciences, The Queen's University of Belfast, 97, Lisburn Road, Belfast BT9 7BL, N.Ireland, UK. Email a.trudgett@qub.ac.uk. Tel. +44 2890972125, Fax +44 2890975877.

Key Words: *Fasciola hepatica*; anthelmintic resistance; P-glycoprotein; SNP

Ribbon model of Liver Fluke P-glycoprotein with arginine substitution at residue 1144.



- Triclabendazole-resistant liver flukes differ in their P-glycoprotein genes.
- Triclabendazole-resistant flukes have greater allelic diversity.
- Triclabendazole resistance correlates with a change in the P-glycoprotein X-loop.
- This may provide the basis for a test for triclabendazole resistance.

Abstract

Control of fasciolosis is threatened by the development of anthelmintic resistance. Enhanced triclabendazole (TCBZ) efflux by ABC transporters such as P-glycoprotein (Pgp) has been implicated in this process. A putative full length cDNA coding for a Pgp expressed in adult *Fasciola hepatica* has been constructed and used to design a primer set capable of amplifying a region encoding part of the second nucleotide binding domain of Pgp when genomic DNA was used as a template. Application of this primer set to genomic DNA from TCBZ-resistant and -susceptible field populations has shown a significant difference in the alleles present. Analysis of an allele occurring at a three-fold higher frequency in the "resistant" population revealed that it was characterised by a serine to arginine substitution at residue 1144. Homology modelling studies have been used to locate this site in the Pgp structure and hence assess its potential to modify functional activity.

The common liver fluke (*Fasciola hepatica*) is a major cause of economic losses to agriculture in temperate regions, with cost estimated at US$2000 million per annum [1]. In the absence of an effective vaccine against the fluke control is achieved by chemotherapy. However, the effectiveness of the drug of choice, triclabendazole (TCBZ) - a benzimidazole derivative – is threatened by the development of resistant fluke populations; these have been reported from Australia, Europe and South America [reviewed in 2]. Work with TCBZ-resistant isolates [3] has indicated that, in contrast to resistance to benzimidazoles in nematodes [4], TCBZ resistance is not associated with changes in the likely target molecule -tubulin [5]. TCBZ-resistant isolates have been shown to process TCBZ more rapidly and their resistant phenotype can be reversed, under in vitro conditions, by the co-administration of inhibitors of P-glycoprotein (Pgp) drug efflux pumps or drug detoxification pathways [2].

F. hepatica is known to express a Pgp-like ABC transporter [6], although this protein has been reported to be predominately expressed in juvenile flukes, and, as such, generally will not be involved in TCBZ efflux from adult flukes established in the bile ducts of infected animals. Recent publication of EST libraries for *F. hepatica* using mature flukes [7] (specifically contigs 1473, 14497, 10329 and 15817) in conjunction with work in our laboratory, has allowed the construction of a putative full length sequence encoding a *F. hepatica* Pgp expressed in adult flukes (supplementary material 1). A primer set (forward primer 5' ttggtgttgtatcgcaggaa 3', reverse primer 5' agccgaagtagcttcatcca 3'; enclosing residues 1090 to 1175) derived from the region coding for the second nucleotide binding domain of Pgp has been successfully applied to genomic DNA extracted from flukes from infra-populations in cattle raised on farms associated with either TCBZ susceptibility or resistance [8] and the amplicons sequenced. (It should be noted that these are functional designations for the populations present on the two farms – the "resistant" populations contains up to 36% susceptible flukes and conversely, the "susceptible" population may contain a minority (<10%) of resistant flukes) [9]. The primers used were separated by 255bp in the constructed coding sequence, but a product of approximately 830bp was routinely obtained. Sequencing of these amplicons revealed that the increased size was due to the presence of a 569bp intron. Characteristic 5' (ATxGTGAG) and 3' (CAGxGGC) splice sites were present at the beginning and end of the non-coding sequence (supplementary material 2). Translation of the partial exons flanking the intron from a putatively TCBZ-susceptible fluke gave a sequence that was identical to the translation of the relevant region of our full length *F. hepatica* sequence and had 77.4% identity with the second ABC transporter region of SMDR2, a *Schistosoma mansoni* putative Pgp (EMBL L26287), confirming that the amplicons produced using this primer set were derived from a gene coding for a liver fluke Pgp. In order to compare the fluke populations present on the

"susceptible" and "resistant" farms, genomic DNA from 10 flukes from each population was prepared and subjected to a PCR reaction using the second nucleotide binding site primers. After sequencing of the amplicons in both directions, an alignment was constructed and edited to give sequences of equal length. The Phase 2.1 algorithm [10] incorporated in DnaSP [11] was used to reconstruct haplotypes where consistent evidence of heterozygosity (double peaks on sequencing chromatograms) was seen in order to allow the derivation of parameters of genetic diversity. The susceptible population contained five haplotypes with a nucleotide diversity (π) of 0.00228. Four of the ten flukes were heterozygous. The resistant population contained seven haplotypes with a nucleotide diversity (π) of 0.00358 and three of these were heterozygous. The presence of heterozygotes indicated that sexual reproduction with cross-fertilisation (with the possibility of recombination) had occurred in previous generations in each population. A consideration of the mitochondrial haplotypes of the individual heterozygous flukes [8] revealed that there was no association between Pgp alleles and maternal lineage, again indicating cross-fertilisation. There was significantly greater diversity (0.768 versus 0.442, P = 0.0002 by a one tailed t-test) in Pgp haplotypes seen in the fluke population from the farm exhibiting TCBZ resistance. Ten unique haplotypes were observed when the resistant and susceptible populations were taken *in toto*. The distribution of these haplotypes over the two populations and the Single Nucleotide Polymorphisms (SNPs) that define them are shown in Table 1. Haplotype 1 was the most common, being present in all of the flukes from the susceptible population and eight of the flukes from the resistant population. The remainder of the haplotypes were segregated between the susceptible and resistant populations. As the two farms are approximately 100 km apart these differences could reflect geographical separation, however, with regard to TCBZ resistance, the mitochondrial haplotypes have been shown to be acting as neutral markers with the most common haplotype being present in approximately one third of the flukes from both farms

[8]. We consider that the differences in frequency of alleles observed in this study are more likely to be due to different selective pressures having operated, directly or indirectly, on the expressed products of the Pgp genes present in the ancestors of the flukes presently at the two locations.

There were 11 individual SNPs, of which 3 were transitions and the rest transversions. Three of the changes were within the partial exons and also resulted in an amino acid change. They were T28A, resulting in a valine to glutamic acid substitution at residue 1112; G54A, resulting in an alanine to threonine substitution at residue 1121; and T687G, resulting in a serine to arginine change at residue 1144. The first of these changes (T28A) was seen in 3 out of 10 TCBZ "susceptible" flukes and in 2 out of 10 TCBZ "resistant" flukes populations and the second (G54A) was exclusively (4 out of 10) in the TCBZ "susceptible" population. The third change (T687G giving S1144R) was seen more frequently in the TCBZ "resistant" population (3 out of 10 for the "resistant" flukes versus 1 out of 10 for the "susceptible" flukes) and was located four residues before the signature motif (LSGGQ) which interacts with adenosine triphosphate (ATP), suggesting that it may have functional significance.

To allow us to investigate the potential significance of the S144R substitution we constructed a homology model of *F. hepatica* Pgp. The amino acid sequence of the constructed *F. hepatica* Pgp sequence was used to search against the NCBI (National Center for Biotechnology Information) nonredundant protein sequence database with the programme PSI-BLAST [12] and identified mouse ABCB1, a multidrug transporter [13], as the closest relative of known structure. The backbone co-ordinates for the core of the homology model were built based on the crystal structure of ABCB1a (PDB ID: 3G5U), in the inward-facing, nucleotide-free conformation as the template. Alignment of the reconstructed *F. hepatica* and *M. musculus* primary sequences revealed a 41% identity between the two proteins. The N-terminus (residues 1-33) of the template crystal structure was not resolved and electron

density for the flexible linker region (residues 627-683) that connects the two halves of the transporter via nucleotide binding domain 1 (NBD1) and transmembrane domain 2 (TMD2) was missing [13]. Therefore, the N-terminus (residues 1-16) and linker region (residues 631 to 674) were not built into our model of the TCBZ-resistant *F. hepatica* transporter. The final model corresponded to residues V17-H630 and V675-E1278 and consisted of 12 α-helices arranged in two 6-helical bundle transmembrane domains (TMD1 and TMD2) and two cytoplasmic nucleotide-binding domains (NBD1 and NBD2). The protein takes on an overall V-shaped conformation in which NBD1 is associated with helices 1, 2, 3 and 6 of TMD1 and helices 10 and 11 of TMD2, and NBD2 is associated with helices 4 and 5 of TMD1 and helices 7, 8, 9 and 12 of TMD2 (Fig. 1). The TMDs form two arms that straddle a substrate binding and translocation pore that is closed towards the extracellular side. The NBDs both contain the Walker A and B motifs and the conserved LSGGQ ABC transporter signature sequence, or C-motif. As shown in Figs. 1 and 2, the arginine residue (R1144) is located at the interface of NBD2 and its associated transmembrane domain, where it protrudes and introduces a positive charge into the space between the two domains. It is also positioned to participate in formation of the NBD dimer interface that occurs upon ATP binding.

The stereochemical properties of the model as determined by PROCHECK revealed that 93% and 6.5% of residues were in the allowed and generously allowed regions, respectively, of the Ramachandran plot. This compares favourably with the Ramachandran plot of the coordinates of the template structure in which 92.5% of residues were in the allowed regions, and 7.5% in the generously allowed. The modelling of the *F. hepatica* Pgp demonstrated that residue 1144 is located in the region termed the X-loop, a part of the molecule which is thought to be involved in the cross-linking of the long intracellular loops (ICLs) in response to ATP binding and may transmit conformational changes to the ICLs [14]. Molecular dynamic simulations [15] have indicated that the regions of the molecule with which the X-

loop interacts potentially change during the cycle of nucleotide attachment and release, emphasising its probable importance in the functional activity of efflux pumps. The S1144R substitution thus has the potential to modulate these conformational changes. In a recent study of cattle ticks resistant to the macrocyclic lactone ivermectin, it was shown that there was an up-regulation of mRNA coding for an ABC transporter in resistant isolates [16]. This transcript also had an arginine in the position corresponding to residue 1144 in our assembled liver fluke sequence whereas two other cattle tick ABC transporters which had a glycine rather than arginine at this position were not up-regulated in the resistant tick isolates, supporting the hypothesis that this substitution may have functional consequences.

The observation that the TCBZ-resistant population does not contain flukes homozygous for this allele may be due to the sample size; alternatively it may indicate that the S1144R substitution, whilst advantageous in the presence of TCBZ, is less appropriate for the other functions of Pgp required by the fluke. It is intended to investigate this issue using *F. hepatica* Pgp expressed *in vitro*.

In conclusion, the data presented in this study, although preliminary, support the suggestion that the development of TCBZ resistance in liver flukes is a multifactorial process involving changes in drug uptake, efflux and metabolism [2] and indicate that selection for variant Pgps may be part of this process. The S1144R substitution is located in a region of the protein which is likely to be associated with its transporter functions. If an enhanced ability to transport TCBZ and its metabolites is confirmed by expression studies with a multidrug efflux protein bearing this substitution this would strengthen its association with the development of TCBZ resistance in the liver fluke. Analysis of isolates characterised with regard to TCBZ susceptibility [3] and field populations will be necessary to establish the importance of the S1144R substitution in the development of the TCBZ-resistant phenotype in locations other than the Netherlands. If changes in this region of Pgp are widely associated

with TCBZ resistance it may be possible to develop an allele-specific molecular test analogous to that used for nematodes [17] for potentially TCBZ-resistant populations based upon genomic DNA extracted from fluke eggs. These are a source of *F. hepatica* genomic DNA readily available for epidemiological surveys.

Acknowledgements

This project was supported in part by the DELIVER project (FOOD-CT-200X-023025), part of the Framework VI Programme of the European Union. This organisation had no part in the study design; the collection, analysis and interpretation of data; the writing of the report or the decision to submit this article for publication.

References

[1] Boray JC. 1994. Disease of Domestic Animals Caused by Flukes. Rome. Food and Agricultural Organization of the United Nations. P.49.

[2] Fairweather I. 2011. Reducing the future threat from (liver) fluke: realistic prospect or quixotic fantasy? Vet. Parasitol. 180, 133−143.

[3] Fairweather I. 2011. Liver fluke isolates: a question of provenance. Vet Parasitol. 176, 1-8.

[4] Wolstenholme AJ, Fairweather I, Prichard R, von Samson-Himmelstjerna G, Sangster NC. 2004. Drug resistance in veterinary helminths. Trends Parasitol. 20, 469-476.

[5] Ryan LA, Hoey E, Trudgett A, Fairweather I, Fuchs M, Robinson MW, Chambers E, Timson DJ, Ryan E, Feltwell T, Ivens A, Bentley G, Johnston D. 2008. *Fasciola hepatica* expresses multiple alpha- and beta-tubulin isotypes. Mol Biochem Parasitol. 159, 73-78.

[6] Reed MB, Panaccio M, Strugnell RA, Spithill TW. 1998. Developmental expression of a *Fasciola hepatica* sequence homologous to ABC transporters. Int J Parasitol. 28, 1375-1381.

[7] Young ND, Hall RS, Jex AR, Cantacessi C, Gasser RB. 2010. Elucidating the transcriptome of *Fasciola hepatica* - a key to fundamental and biotechnological discoveries for a neglected parasite. Biotechnol Adv. 28, 222-231.

[8] Walker SM, Johnston C, Hoey EM, Fairweather I, Borgsteede F, Gaasenbeek C, Prodöhl PA, Trudgett A. 2011. Population dynamics of the liver fluke, *Fasciola hepatica*: the effect of time and spatial separation on the genetic diversity of fluke populations in the Netherlands. Parasitology. 138, 215-223.

[9] Moll L, Gaasenbeek CP, Vellema P, Borgsteede FH. 2000. Resistance of *Fasciola hepatica* against triclabendazole in cattle and sheep in The Netherlands. Vet Parasitol. 91, 153-158.

[10] Stephens M, Donnelly P. 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet. 73, 1162-1169.

[11] Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics. 19, 2496-2497.

[12] Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389-3402.

193 [13] Aller SG, Yu J, Ward A, Weng Y, Chittaboina S, Zhuo R, Harrell PM, Trinh YT, Zhang
194 Q, Urbatsch IL, Chang G. 2009. Structure of P-glycoprotein reveals a molecular basis for
195 poly-specific drug binding. Science. 323, 1718-1722.

196 [14] Dawson RJ, Locher KP. 2006. Structure of a bacterial multidrug ABC transporter.
197 Nature. 443, 180-185.

198 [15] Becker JP, Van Bambeke F, Tulken PM, Prévost M. 2010. Dynamics and structural
199 changes induced by ATP binding in SAV1866, a bacterial ABC exporter. J Phys Chem B.
200 114, 15948-15957.

201 [16] Pohl PC, Klafke GM, Carvalho DD, Martins JR, Daffre S, da Silva Vaz I Jr, Masuda A.
202 2011. ABC transporter efflux pumps: A defense mechanism against ivermectin in
203 *Rhipicephalus (Boophilus) microplus*. Int J Parasitol. 41, 1323-1333.

204 [17] Alvarez-Sanchez MA, Perez-Garcia J, Cruz-Rojo MA, Rojo-Vazquez FA. 2005. Real
205 time PCR for the diagnosis of benzimidazole resistance in trichostrongylids of sheep. Vet.
206 Parasitol., 129, 291−298.

207 [18] Canutescu AA, Dunbrack RL Jr. 2005. MollDE: a homology modeling framework you
208 can click with. Bioinformatics. 21, 2914-2916.

209 [19] Krivov GG, Shapovalov MV, Dunbrack RL. Jr. 2009. Improved prediction of protein
210 side-chain conformations with SCWRL4. Proteins. 77, 778-795.

211 [20] Xiang Z, Soto CS, Honig B. 2002. Evaluating conformational free energies: the colony
212 energy and its application to the problem of protein loop prediction. Proc. Natl. Acad. Sci.
213 USA 99, 7432–7437.

214

Legends to figures

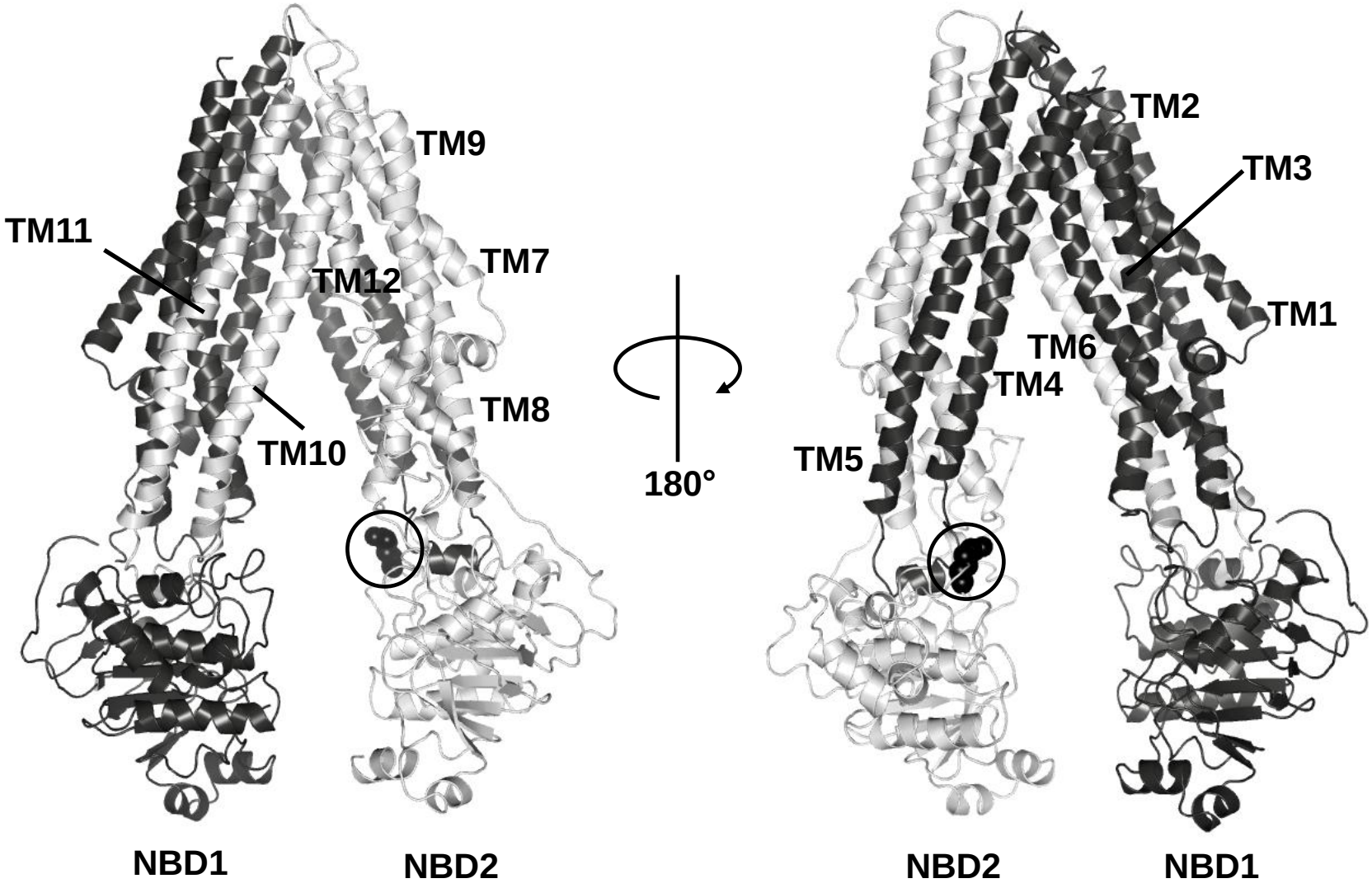
Fig 1. Homology model of the *F. hepatica* P-glycoprotein – front and rear views. The position of residue R1144 (represented as a red space-fill) is indicated by the circle and is located at the interface of NBD2 and its associated transmembrane domain, TM4. TM, transmembrane domain; NBD, nucleotide binding domain. Homology modelling was performed using MolIDE 1.7 [18]. PSIPRED was used to perform initial secondary structure prediction of the target, and the target-template sequence alignment was adjusted manually using the predicted secondary structure of the target and the experimental secondary structure of the template as a guide for gap placement. During building of the backbone core of the model, only those side-chains that are conserved between target and template were preserved. The remaining, non-conserved side chains were built onto the backbone with SCWRL4 [19], and loops modelled with Loopy [20]. The model was visualised using The PyMOL Molecular Graphics System, Version 1.3 (Schrödinger, LLC).

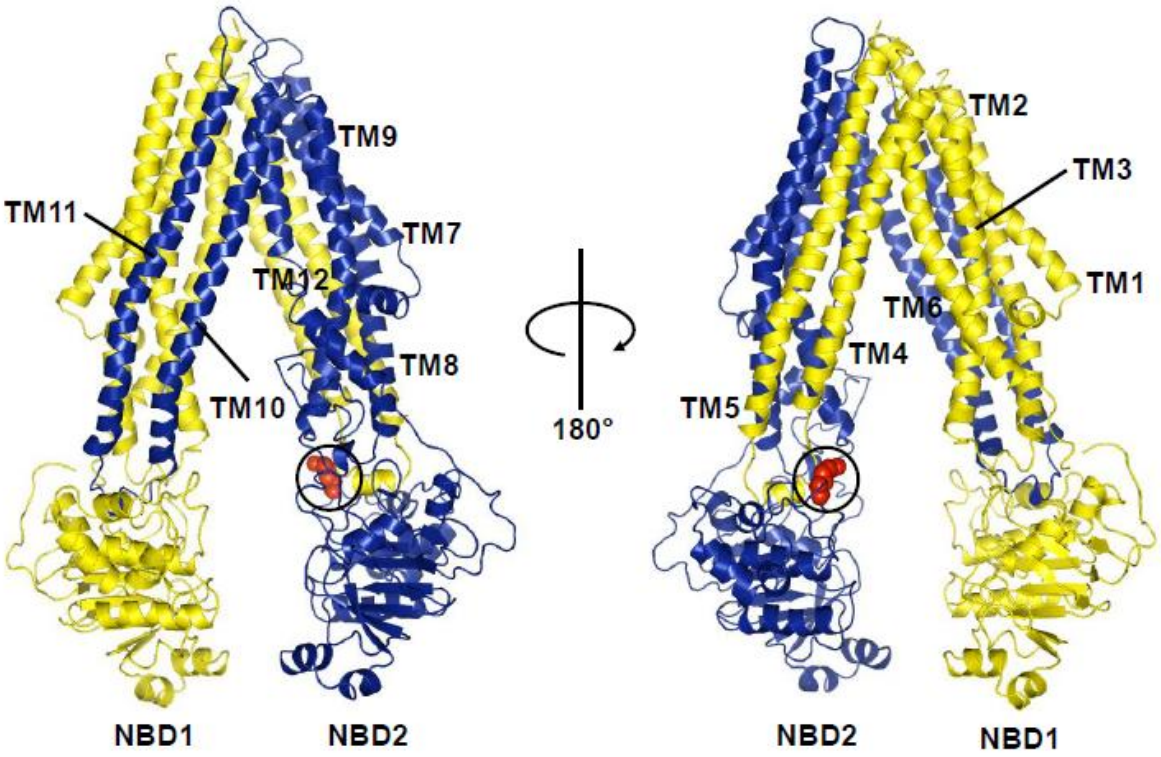
Fig 2. Close-up of the NBD2 region of the homology model showing residue 1144. A – modelled with arginine as residue 1144, B – modelled with serine as residue 1144.

Haplotype	Single Nucleotide Polymorphisms	Number (Resistant/ Susceptible)	Change in Protein
1	-	23 (Resistant and Susceptible)	-
2	A564G, T687G	1 (Susceptible)	S1144R
3	T28A, G54A, G181T, T267G, G323T, A564G,	1 (Susceptible)	V1112E, A1121T,
4	T28A, G54A, C60T,	2 (Susceptible)	V1112E, A1121T
5	G54A	1 (Susceptible)	A1121T
6	G181T, T267G, A617T	1 (Resistant)	
7	G181T, T267G, G323T, A617T, T687G	2 (Resistant)	S1144R
8	T267G, A617T,	6 (Resistant)	
9	T28A, G181T, T307A	1 (Resistant)	V1112E
10	G20C, T28A, T307A, T687G	1 (Resistant)	V1112E, S1144R

Table 2 Constructed haplotypes, their distribution in the two populations and non-synonymous changes in exons.

Figure(s)


TM9
TM11
TM7
TM12
TM8
TM10
NBD1
NBD2
180°
TM2
TM3
TM1
TM6
TM4
TM5
NBD2
NBD1



Figure(s)