Rev Biomed 2008; 19:53-60

Artículo Original

Topology predictions of membrane spanning region in *Plasmodium yoelii* ABC proteins

Iván Ferrer-Rodríguez

Department of Natural Sciences and Mathematics, Inter American University of Puerto Rico

SUMMARY

Introduction. Malaria drug resistance continues on the rise and constitutes a major health problem. Drug resistance in *Plasmodia* is a complex phenomenon often mediated by membrane proteins belonging to the ATP-Binding Cassette (ABC) superfamily of transporter, which are characterized by the presence of nucleotide-binding sites (NBS) and membrane spanning domains (MSD).
Objective. The main objective of this paper is to analyze the performance of a variety of tools to predict the transmembrane topology of ABC proteins in *P. yoelii*.
Material and methods. ABC proteins were identified in PlasmoDB 5.4 using the Search term tool querying with ABC as keyword. Protein sequences were analyzed for prediction of NBS and MSD using seven different bioinformatics tools. Each program was rated based on the number of correct and incorrect predictions.
Results. Seven of the 23 proteins identified contain the typical architecture structure of ABC proteins with transmembrane regions. The number of transmembrane domains in the proteins ranged from four to 11. TMHMM 1.0 provided the best comparison to the reference annotation in PlasmoDB (TMHMM 2.0) with 51 correct predictions, followed by Phobius, TMPRED and HMMTOP. MEMSAT and SPLIT have the lowest number of correct predictions.
Conclusions. We performed topology predictions of membrane spanning regions in *P. yoelii* ABC proteins. These analyses should provide further information about the structure of the ABC proteins and could guide researchers to understand better the role that these proteins can play in biological processes in the parasite.

Key words: Topology predictions, *Plasmodium yoelii*, ABC proteins

RESUMEN

Predicciones de topología de las regiones transmembranales de las proteínas ABC de *P. yoelii*

Introducción. El problema de resistencia a drogas en malaria continúa en aumento y representa un gran problema de salud. La resistencia a drogas en los *Plasmodia* es un fenómeno complejo frecuentemente mediado por proteínas de membrana de la familia ABC (ATP-Binding Cassette), las cuales se caracterizan por la presencia de lugares de enlaces a nucleótidos (NBS) y una región con dominios transmembranales (MSD).
Objetivo. El objetivo principal del trabajo es analizar el desempeño de diferentes herramientas cibernéticas para predecir la topología de las proteínas ABC en *P. yoelii*.
Material y Métodos. Se identificaron las proteínas ABC en PlasmoDB 5.4 utilizando la herramienta de búsqueda de términos, con ABC como la palabra clave. Las proteínas identificadas fueron analizadas utilizando siete herramientas cibernéticas que

Solicitud de sobretiros: Iván Ferrer-Rodríguez. Department of Natural Sciences and Mathematics, Inter American University of Puerto Rico, Bayamón Campus, 500 Dr. John Will Harris Road, Bayamón PR 00957 E-mail: iferrer@bc.inter.edu

Recibido: el 7 de abril de 2007. **Aceptado para publicación:** el 30 de abril de 2008.

Este artículo está disponible en http://www.revbiomed.uady.mx/pdf/rb081917.pdf

predicen NBS y MSD. Los programas fueron clasificados con base en el número de predicciones correctas e incorrectas.

Resultados. Siete de las 23 proteínas que fueron identificadas poseían la estructura de arquitectura típica de las proteínas ABC con regiones transmembranales. El número de dominios transmembranales en las proteínas varió entre 4-11. TMHMM 1.0 generó la mejor comparación en relación a la anotación de referencia en PlasmoDB (TMHMM 2.0) con 51 predicciones correctas, seguido por Phobius, TMPRED y HMMTOP. MEMSAT y SPLIT tuvieron el menor número de predicciones correctas.

Conclusiones. Se realizaron predicciones de topología de las regiones transmembranales de las proteínas ABC de *P. yoelii*. Estos análisis deben proveer información adicional sobre la estructura de las proteínas ABC y deben servir de guía a los investigadores para entender mejor el papel que estas proteínas podrían desempeñar en los procesos biológicos del parásito.

Palabras clave: Topología, predicciones, *Plasmodium yoelii*, proteínas ABC

INTRODUCTION

Human malaria remains one of the most important challenges to public health systems worldwide. According to the World Health Organization (WHO), malaria is the world's second biggest killer and is responsible for more than 300 million clinical cases globally, resulting in a call death between 1.5 and 2.7 million every year, with 80 percent of them in children from sub-Saharan Africa (1).

One of the major problems for malaria control is due to drug resistance, which is often mediated by membrane proteins belonging to the ATP-Binding Cassette (ABC) superfamily of transporter that hydrolyse ATP to energize the translocation of a wide variety of substrates across cell membranes. The ABC transporters are minimally constituted of a highly conserved ATP binding site and a less conserved transmembrane domain. These domains may be found in the same protein or on two different ones (2). The expression profile of the ABC transporters of *Plasmodium yoelii* and *P. berghei* was characterized by Szeto and coworkers (3), who demonstrated that 14 genes in *P. yoelii* and ten in *P. berghei* are transcribed in intraerythrocytic stages.

Even with the wealth of information available in genome sequencing projects on *P. berghei*, *P. chabaudi*, *P. falciparum*, *P. gallinaceum*, *P. knowlesi*, *P. reichenowi*, *P. vivax* and *P. yoelii* available in PlasmoDB (4), only a small number of genes have been shown to be involved in resistance to the quinoline-containing antimalarials (5-11).

Currently the post-genomic era, with its tools for accessing and analyzing data, provides an excellent opportunity for conducting bioinformatics approaches and comparative genomics to address important biological questions in the parasite, including drug resistance. A better understanding of the structure of the ABC proteins and their role in drug resistance may be crucial to develop novel strategies to treat the disease in the future. Since the topology of the majority of membrane proteins in *Plasmodium* remains biochemically uncertain, we analyzed the performance of different programs available in the Web for the prediction of transmembrane spanning regions in *P. yoelii* ABC proteins.

MATERIAL AND METHODS

Identification of *P. yoelii* ABC proteins. To identity all *P. yoelii* ABC proteins available in PlasmoDB 5.4 (www.plasmodb.org), the database was examined using the Search term tool querying with ABC as keyword. Other parameters used were a Max p-value exponent (BLAST Hits v. NRDB) of -30 and the following fields: Gene product, Gene notes, Protein domain names and descriptions, GO terms and definitions and Metabolic pathway names and descriptions. The genes encoding for proteins that were retrieved were used to consult the Orthologue Groups of Protein Sequences in

OrthoMCL DB and the list of domains in the group were identified using Pfam Domain Architecture. Protein sequences were downloaded in FASTA format and saved as a text document.

Plasmodium **ABC Sequence Analysis.** Protein sequences were analyzed for prediction of ABC and transmembrane regions using seven different bioinformatics tools. The programs used were TMHMM 1.0 (12), MEMSAT (13), HMMTOP (14, 15), TMPRED (16), SPLIT 4.0 SERVER (17) and Phobius (18). The topology predictions were compared to those available in PlasmoDB with TMHMM 2.0, which was previously categorized as the best performing transmembrane prediction program in its original version (19). Each one of the programs was rated by three values. First, the numbers of correct predictions, which mean the numbers of predictions that, were present in the reference annotation. Second, the negative incorrect predictions refer to MSD that, were not predicted by the program. Third, the positive incorrect predictions represent the MSD that were not present in the reference annotation.

RESULTS

The search in PlasmoDB 5.4 (www.plasmodb.org) using the Search term tool querying with ABC as keyword resulted in 23 genes (PY00207, PY00245, PY00547, PY00626, PY01655, PY01780, PY01826, PY02551, PY03961, PY04219, PY04270, PY04291, PY04349, PY05035, PY05246, PY06050, PY06054, PY06462, PY06546, PY06911, PY07750, PY07088, PY07089). These sequences ranged in sized form 498 bp (PY00547) to 5928 bp (PY05035).

The number of introns in the genes retrieved was as follows: 13 genes that contain no introns; six genes have one intron, two genes with two introns and a single gene containing three and four introns, respectively. A close examination of the Pfam Domain Architecture in PlasmoDB revealed that seven of the 23 *P. yoelii* ABC proteins contained the typical structure of ABC proteins with transmembrane regions: PY00207, PY00245, PY01826, PY05035, PY06054, PY06546 and PY07088 (**Table 1**).

Table 1

Descriptions of the *P. yoelii* ABC genes encoding for proteins with transmembrane regions

ABC Genes	Transcript Length (nucleotides)	Pfam Protein Domain Architecture	Product Description
PY00207	5514	* NBS-MSD	ABCG subfamily, breast cancer resistance protein gene
PY00245	4266	(MSD-NBS)2	ABCB subfamily, multidrug resistance protein 1
PY01826	2418	MSD-NDS	ABC transporter, TAP family
PY05035	5928	(MSD-NBS)2	Multidrug resistance ABC transporter, CT family MRP
PY06054	2667	MSD-NBS	Multidrug resistance protein 2
PY06546	2037	MSD-NDS	Transport protein, putative
PY07088	2409	MSD-NBS	ABC transporter, heavy metal transporter family

* NBS stands for nucleotide binding site and MSD for membrane spanning domain.

These sequences ranged in size from 2037 bp (PY06546) to 5928 bp (PY05035). The typical domain architecture of NBS and MSD was observed in all seven proteins (**Figure 1**).

The topology predictions provided by PlasmoDB with TMHMM 2.0 suggest that these seven proteins contain 51 MSD (**Table 2**). These predictions were used as the reference annotation to evaluate the prediction capability of various methods available in the Web. The number of transmembrane domains ranged from four (PY00207) to 11 (PY00245 and PY05035).

The performance of the different tools in terms of prediction of the individual MSD in the *P. yoelii* ABC proteins is presented in **Table 3**. The accuracy of prediction was measured considering both the number of transmembrane segments correctly predicted and the number of

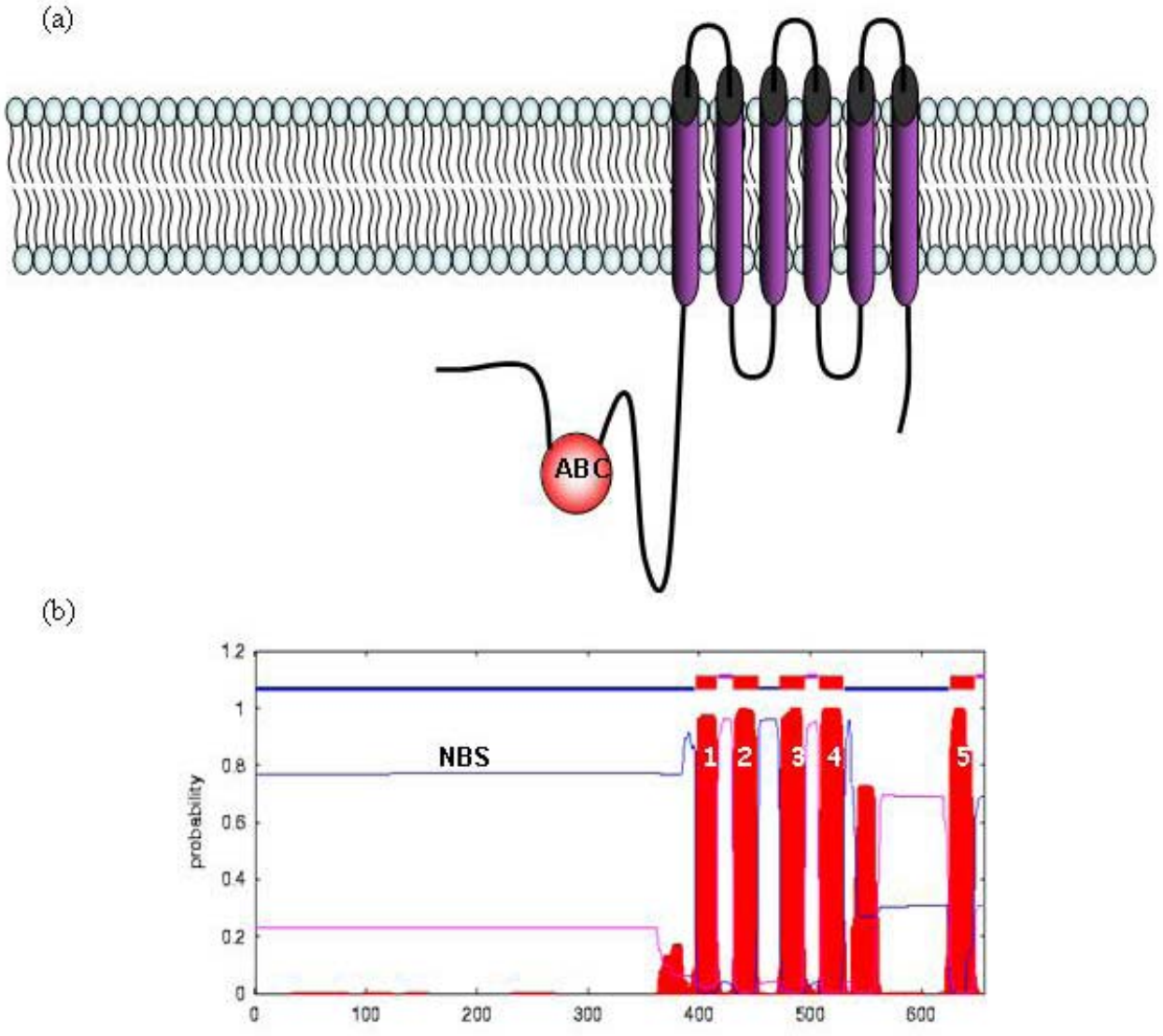

Figure 1. Typical membrane topology of an ABC protein with one nucleotide binding sites (NBS) and a membrane spanning domains (MSD). (a) A schematic representation of the topology of an ABC protein. (b) A representative prediction by TMHMM 1.0 showing the NBS followed by a MSD (1-5).

Table 2
Predictions of membrane spanning regions in *P. yoelii* ABC proteins as determined by PlasmoDB with TMHMM 2.0

ABC Proteins	Transmembrane helix (TM) predictions											
	TM-1	TM-2	TM-3	TM-4	TM-5	TM-6	TM-7	TM-8	TM-9	TM-10	TM-11	Total TM
PY00207	1610-32	1653-75	1688-1710	1806-28	-	-	-	-	-	-	-	4
PY00245	54-76	96-118	164-186	190-209	274-296	316-338	787-809	829-851	902-924	928-947	1023-45	11
PY01826	184-206	226-248	303-322	326-348	411-433	443-465	-	-	-	-	-	6
PY05035	129-151	171-193	333-355	365-387	447-469	1318-40	1377-99	1414-36	1449-68	1473-95	1554-76	11
PY06054	4-23	35-57	70-92	321-343	358-375	436-458	462-479	553-572	582-604	-	-	9
PY06546	101-123	138-158	213-235	245-262	321-340	-	-	-	-	-	-	5
PY07088	332-354	454-476	480-502	575-592	596-618	-	-	-	-	-	-	5

* Numbers in columns TM-1 through TM-6 represent amino acid numbers

Table 3
Performance of different bioinformatics tools predicting membrane spanning regions in *P. yoelii* ABC proteins

Methods	Links	* Correct predictions	Incorrect predictions		Total incorrect predictions
			† Negatives	‡ Positives	
TMHMM v1.0	http://www.cbs.dtu.dk/services/TMHMM/	51	0	3	3
Phobius	http://phobius.sbc.su.se/	49	2	5	7
TMPRED	http://www.ch.embnet.org/software/TMPRED_form.html	48	3	13	16
HMMTOP	http://www.enzim.hu/hmmtop/	45	6	11	17
MEMSAT 1.5	http://saier-144-37.ucsd.edu/memsat.html	39	1	4	5
SPLIT 4.0	http://split.pmfst.hr/split/4/	24	0	2	2
Total		256	12	38	50

* Correct predictions mean the numbers of predictions that were present in the reference annotation
† Negative incorrect predictions refer to membrane spanning domains (MSD) that were not predicted by the program
‡ Positive incorrect predictions represent the MSD that were not present in the reference annotation

incorrect predictions, as compared to the reference annotation. TMHMM 1.0 resulted in all 51 correct predictions and only three incorrect predictions. Phobius, TMPRED and HMMTOP resulted in 49, 48 and 45 correct predictions respectively, while MEMSAT and SPLIT have the lowest number of correct predictions with 39 and 24, respectively.

The performance of all the programs on each individual protein was evaluated. In three instances, PY00207, PY06546, PY07088, the consensus topology considering all the programs suggest that there are additional transmembrane regions than those suggested by TMHMM 2.0 (**Table 4**). These analyses suggest that there are two additional domains in PY00207 and one extra domain in PY06546 and PY07088.

DISCUSSION

In this report, we compared the performance of different bioinformatics tools available in the Web to predict transmembrane spanning regions in *P. yoelii* ABC proteins. Of the 23 genes identified in the search term tool querying with ABC as keyword, only seven proteins contained transmembrane regions.

These included three genes that have been related to drug resistance in a variety of organisms, including the breast cancer resistance protein gene (BCRP) (PY00207), the multidrug resistance protein 1 (MDR1) (PY00245) and the multidrug resistance associated protein gene (MRP) (PY05035) (5-9, 20-23). These genes have been shown to be expressed in intraerythrocytic stages of *P. yoelii* (3). The other genes identified were the multidrug resistance protein 2 (MDR2) (PY06054), a transporter of the TAP family (PY01826), a heavy metal transporter (PY07088) and a putative transporter (PY06546). Five of these ABC proteins are half transporters (PY00207, PY01826, PY06054, PY06546 and PY07088), which consist of a MSD and one NBS. PY00245 and PY05035 are full transporters containing two MSD and two NDS. TMHMM 2.0 in PlasmoDB predicted between four and nine transmembrane domains in the half transporters and 11 transmembrane domains in the full transporters.

Table 4
Consensus topology for *P. yoelii* ABC proteins by using different bioinformatics tools as compared with TMHMM 2.0

ABC Proteins	Method	Transmembrane helix (TM) predictions						
		TM-1	TM-2	TM-3	TM-4	TM-5	TM-6	Total TM
PY00207	TMHMM 2.0	-	1610-1632	1653-1675	1688-1710	-	1806-1828	4
PY06546	Consensus	1577-1596	1610-1632	1653-1675	1690-1708	1721-1739	1806-1828	6
	TMHMM 2.0	101-123	138-158	213-235	245-262	321-340	-	5
PY07088	Consensus	101-123	138-158	213-235	245-262	321-340	357-375	6
	TMHMM 2.0	332-354	-	454-476	480-502	575-592	596-618	5
	Consensus	332-354	373-394	454-476	480-502	575-592	596-618	6

* Numbers in columns TM-1 through TM-6 represent amino acid numbers

The evaluation of the different tools in terms of predictions of the individual MSD in the *P. yoelii* ABC proteins revealed that TMHMM 1.0 provided the best comparison to the reference annotation in PlasmoDB (TMHMM 2.0) with all 51 correct predictions, no false negatives and only three false positives. These results are in agreement with those of Möller *et al*. (19), who found that TMHMM was the best performing prediction program in a group of proteins with known biochemical characterization of membrane topology. The second best performance was provided by Phobius that showed 49 correct predictions with only two false negatives and five false positives. TMPRED, HMMTOP, MEMSAT and SPLIT followed in ranking. The low rankings displayed by MEMSAT and SPLIT could be partially due to their inability to predict the structures of all the proteins. No results were obtained for PY05035 with MEMSAT and for PY01826, PY05035, PY06546 and PY07088 with SPLIT.

The consensus topology considering all the programs suggests that in three proteins, PY00207, PY06546, PY07088, there were additional domains, as compared to TMHMM 2.0. Since the *P. yoelii* genome is not finished, the *Plasmodium* orthologs were consulted to validate our results. Two additional domains were identified in PY00207. These results are consistent with the orthologue gene in *P. falciparum*, PF14_0244, a half transporters with six transmembrane helices predicted. An additional domain was identified in the consensus topology in PY06546, PY07088, changing the TMHMM 2.0 prediction from five to six transmembrane helices.

In summary, we performed predictions of membrane spanning region in *Plasmodium yoelii* ABC proteins. Our results suggest that additional membrane spanning domains could be present in ABC transporters, as compared to the topology suggested by PlasmoDB. These analyses provide additional information regarding the structure and orientation of the ABC proteins in the parasite and should be considered an aid to the scientists to make an educated guess to understand the role that these proteins can play in drug resistance. Additional bioinformatics studies with multiple predictors should be conducted to further characterized the topology of *Plasmodium* ABC proteins.

Acknowledgment. The project described was supported by Grant P20 RR016470 from the National Center for Research Resources. The content is solely the responsibility of the authors and does not necessarily represent the official view of the National Center for Research Resources or the National Institutes of Health. Thanks to the Puerto Rico Louis Stokes Alliance for Minority Participation from the National Science Foundation (HRD-0601843). I am grateful to Dr. Adelfa E. Serrano for her collaboration and to Dr. Gary W. Gervais, Rafael Canales, Gloriene González and Dr. Oscar Ruiz for their comments and recommendations.

REFERENCES

1. **Malaria Fact sheet N°94.** May 2007. Available from: http://www.who.int/mediacentre/factsheets/fs094/en/.
2. **Saurin W, Hofnung M, Dassa E.** Getting in or out. Early segregation between importers and exporters in the evolution of ATP-binding cassette (ABC) transporters. J Mol Evol 1998; 48:22-41.
3. **Szeto AC, Perez-Rosado J, Ferrer-Rodriguez I, Vega J, Torruella-Thillet C, Serrano AE.** Identification and expression analysis of ABC genes in *Plasmodium yoelii* and *P. berghei*. Parasitol Res 2004; 92:1-11.
4. **Bahl A, Brunk B, Crabtree J, Fraunholz MJ, Gajria B, Grant GR,** *et al*. PlasmoDB: the *Plasmodium* genome resource. A database integrating experimental and computational data. Nucleic Acids Res 2003; 31(1):212-5.
5. **Ferrer-Rodríguez I, Pérez-Rosado J, Gervais GW, Peters W, Robinson BL, Serrano AE.** *Plasmodium yoelii*: Identification and partial characterization of an *mdr1* gene in an artemisinin resistant line. J Parasitol 2004; 90:152-60.
6. **Reed MB, Saliba KJ, Caruana SR, Kirk K, Cowman AF.** Pgh1 modulates resistance to multiple antimalarials in *Plasmodium falciparum*. Nature 2000; 403:906-9.
7. **Gervais G, Trujillo K, Robinson BL, Peters W, Serrano AE.** *P. berghei*: Identification of an *mdr*-like gene associated with drug resistance. Exp Parasitol 1999; 91:86-91.
8. **Foote SJ, Thompson JK, Cowman AF, Kemp DJ.** Amplification of the multidrug resistance gene in some chloroquine-resistant isolates of *P. falciparum*. Cell 1989; 57:921-30.

9. **Wilson CM, Serrano AE, Wasley A, Bogenshutz MP, Shankar AH, Wirth, DF.** Amplification of a gene related to mammalian *mdr* genes in drug-resistant *Plasmodium falciparum*. Science 1989; 244:1184-6.
10. **Hunt P, Cravo PVL, Donleavy P, Carlton J, Wallikker, D.** Chloroquine resistance in *Plasmodium chabaudi:* are chloroquine resistance transporter (*crt*) and multi-drug resistance (*mdr1*) orthologues involved? Mol Biochem Parsitol 2004; 133:27-35.
11. **Fidock AD, Nomura T, Talley KA, Cooper AR, Dzekunov MS, Ferdig TM,** *et al.* Mutations in the *P. falciparum* digestive vacuole transmembrane protein PfCRT and evidence for their role in chloroquine resistance. Mol Cell 2000; 6:861-71.
12. **Krogh A, Larsson B, Von Heijne G, Sonnhammer ELL.** Predicting transmembrane protein topology with a Hidden Markov Model: Application to complete genomes. J Mol Biol 2001; 305(3):557-80.
13. **Jones DT, Taylor WR, Thornton J M.** A model recognition approach to the prediction of all-helical membrane protein structure and topology. Biochemistry 1994; 33(10):3038-49.
14. **Tusnády GE, Simon I.** Principles governing amino acid composition of integral membrane proteins: Applications to topology prediction. J Mol Biol 1998; 283:489-506.
15. **Tusnády GE, Simon I.** The HMMTOP transmembrane topology prediction server. Bioinformatics 2001; 17:849-50.
16. **Hofmann K, Stoffel W.** TMBASE-A database of membrane spanning protein segments. Biol Chem Hoppe-Seyler 1993; 374:166.
17. **Juretić D, Zucić D, Lucić B, Trinajstić N.** Preference functions for prediction of membrane-buried helices in integral membrane proteins. Comput Chem 1998; 22(4):279-94.
18. **Käll L, Krogh A, Sonnhammer EL.** A Combined Transmembrane Topology and Signal Peptide Prediction Method. J Mol Biol 2004 338(5):1027-36.
19. **Möller S, Croning MD, Apweiler R.** Evaluation of methods for the prediction of membrane spanning regions. Bioinformatics 2001; (7):646-53. Erratum in: Bioinformatics 2002; 18(1):218.
20. **Di Pietro A, Conseil G, Perez-Victoria JM, Dayan G, Baubichon-Cortay H, Trompier D,** *et al.* Modulation by flavonoids of cell multidrug resistance mediated by P-glycoprotein and related ABC transporters. Cell Mol Life Sci 2002; 59(2):307-22.
21. **Janvilisri T, Venter H, Shahi S, Reuter G, Balakrishnan L, Van Veen HW.** Sterol transport by the human breast cancer resistance protein (ABCG2) expressed in *Lactococcus lactis*. J Biol Chem 2003; 278(23):20645-51.
22. **Legare D, Richard D, Mukhopadhyay R, Stierhof YD, Rosen BP, Haimeur A,** *et al*. The *Leishmania* ATP-binding cassette protein PGPA is an intracellular metal-thiol transporter ATPase. J Biol Chem 2001; 13(28):26301-7.
23. **Vezmar M, Georges E.** Direct binding of chloroquine to the multidrug resistance protein (MRP): possible role for MRP in chloroquine drug transport and resistance in tumor cells. Biochem Pharmacol 1998; 56(6):733-42.