Homology modeling and structural analysis of human P-glycoprotein

Hideaki Yamaguchi[1]*, Yumi Kidachi[2], Katsuyoshi Kamiie[2], Toshiro Noshita[3] & Hironori Umetsu[4]

[1]Department of Pharmacy, Faculty of Pharmacy, Meijo University, 150 Yagotoyama, Tenpaku, Nagoya 468-8503, Japan; [2]Department of Pharmacy, Faculty of Pharmaceutical Sciences, Aomori University, 2-3-1 Kobata, Aomori 030-0943, Japan; [3]Department of Life Sciences, Faculty of Life and Environmental Sciences, Prefectural University of Hiroshima, 562 Nanatsuka, Shobara 727-0023, Japan; [4]Laboratory of Food Chemistry, Department of Life Sciences, Junior College, Gifu Shotoku Gakuen University, 1-38 Nakauzura, Gifu 055-8288, Japan; Hideaki Yamaguchi – Email: hyamagu@meijo-u.ac.jp; Phone: +81 52 839 2721; *Corresponding author

Abstract:
Homology modeling and structural analysis of human P-glycoprotein (hP-gp) were performed with a software package the Molecular Operating Environment (MOE). A mouse P-gp (mP-gp; PDB code: 3G5U) was selected as a template for the 3D structure modeling of hP-gp. The modeled hP-gp showed significant 3D similarities at the drug biding site (DBS) to the mP-gp structure. The contact energy profiles of the hP-gp model were in good agreement with those of the mP-gp structure. Ramachandran plots revealed that only 3.5% of the amino acid residues were in the disfavored region for hP-gp. Further, docking simulations between 6-(methylsulfinyl)hexyl isothiocyanate (6-MITC) and the P-gp models revealed the similarity of the ligand-receptor bound location between the hP-gp and mP-gp models. These results indicate that the hP-gp model was successfully modeled and analyzed. To the best of our knowledge, this is the first report of a hP-gp model with a naturally occurring isothiocyanate, and our data verify that the model can be utilized for application to target hP-gp for the development of antitumor drugs.

Background:

ATP-binding cassette (ABC) transporters have been reported to act as efflux pumps at cellular membranes to transport molecules against concentration gradients, and consequently decrease intracellular concentrations of compounds that are structurally and/or functionally unrelated to the cells. They are endogenously present in cells, involved in absorption and elimination of compounds, and also affect the disposition of the compounds processed by cells. The transport of endogenous molecules is likely the primary function of ABC transporters, but they also transport drugs across cell membranes due to structural similarities between exogenous (foreign) and endogenous substances. Further, because of their wide presence in tissues such as the bile canaliculi of hepatocytes, the apical membranes of enterocytes and various blood-tissue barriers, ABC transporters play important roles in drug distribution as well. They also serve protective roles in many normal tissues of principal organs such as brain, liver, kidney and intestine, through the ATP-dependent efflux of xenobiotics and toxins. Taken together, ABC transporters play significant roles in absorption, elimination and disposition of compounds including drugs **[1]**.

BIOINFORMATION

ABC transporters are not only endogenously expressed in normal cells but also often overexpressed in tumor cells. As a consequence, they export antitumor drugs actively from the tumor cells, decrease these drugs below therapeutic thresholds and confer multidrug resistance (MDR) on the tumor cells. MDR is the phenomenon of resistance to many structurally unrelated cytotoxic agents, and is considered one of the major problems that need to be overcome in tumor chemotherapy [2]. At the surface of the cell membranes, MDR tumor cells overexpress efflux proteins (ABC transporters) such as P-glycoprotein (P-gp). The P-gp-mediated efflux of compounds is one of the well-studied mechanisms underlying MDR.

The size of P-gp has been reported to be 170 kDa, and it is encoded by the MDR1 gene in the ABC superfamily. P-gp is expressed endogenously in normal tissues of major organs such as brain, lung, kidney, liver and colon [3], which suggests that P-gp may have some specific physiological roles associated with specialized cell functions in different organs. P-gp is mainly distributed in the epithelia of excretory organs, and its ability to transport a wide range of lipophilic substrates indicates that P-gp functions in the detoxification process in the body. In tumor cells, P-gp is often overexpressed and acts as an ATP-dependent efflux pump to transfer a wide variety of compounds including xenobiotics, carcinogens and chemotherapeutic agents from cells. As a consequence, it decreases the intracellular concentrations of antitumor drugs, and the overexpression of P-gp has been positively correlated with poor prognosis in cancers. The antitumor drugs that have been identified as substrates of P-gp are anthracyclines, vinca alkaloids, epipodophyllotoxins and taxanes [4].

P-gp has also been investigated for their interactions with isothiocyanates (ITCs), and it was demonstrated that certain ITCs were inhibitors of P-gp [5]. We have been focusing our attention on ITCs as multifunctional agents for the prevention and treatment of cancer. Recently, we found that a naturally occurring ITC from wasabi (*Wasabia japonica*), a Japanese indigenous herb, and demonstrated significant growth inhibitory activity toward a macrophage-like tumor cell line [6]. These reports suggest that targeting P-gp with inhibitors from natural origin could be utilized for controlling the development and progression of cancer. P-gp consists of 12 transmembrane domains and 2 nucleotide binding folds with ATPase activity, and its biological functions are accompanied with its unique structural changes [7]. Thus, structural analysis of P-gp with its possible inhibitory ligands could be of importance for successful antitumor drug development. However, no human P-gp (hP-gp) structural model with naturally occurring inhibitors is available to the best of our knowledge, although a crystal structure of mouse P-pg (mP-gp) has been reported [7]. Molecular modeling has found widespread utility in the field of drug development [8, 9], and in the present study we will report the homology modeling and structural analysis of hP-gp by a highly sophisticated software package, the Molecular Operating Environment (MOE) 2010.10 (Chemical Computing Group Inc., Montreal, Canada).

Methodology:
Homology modeling of hP-gp
Homology modeling of hP-gp was executed as previously reported [10]. In brief, the hP-gp (NCBI reference sequence: NP_000918.2) sequences and the crystal structure coordinates of mP-gp [7] were loaded into the MOE. The primary structures of hP-gp and mP-gp were aligned, carefully checked to avoid deletions or insertions in conserved regions and corrected wherever necessary. A series of hP-gp models were independently constructed with the MOE using a Boltzmann-weighted randomized procedure [11] combined with specialized logic for the handling of sequence insertions and deletions [12]. The models with the best packing quality function were selected for full energy minimization and further inspection.

Assessment of the Modeled Structure
The qualities of the protein folds of the hP-gp homology model were evaluated with the MOE by calculating the effective atomic contact energies, comprising the desolvation free energies required to transfer atoms from water to the interior of the protein [13]. Briefly, the contact desolvation energies were calculated for 18 different atom types of the 20 common amino acids that were resolved based on the clustering pattern of their properties. The contact potentials for each atom type were measured within a contact range of 6 Å by explicitly accounting for neighboring interactions. The overall geometric and stereochemical qualities of the final modeled structure of hP-gp were examined using Ramachandran plots generated within the MOE [14].

Binding Site Selection and Exploration
The binding site selection and exploration for hP-gp was executed as previously reported [10]. In brief, the Site Finder module of the MOE was used to identify possible ligand-binding pockets within the newly generated 3D structures of hP-gp. Hydrophobic or hydrophilic alpha spheres served as probes denoting zones of tight atom packing. These alpha spheres were utilized to define potential ligand-binding sites (LBSs) and as centroids for the creation of dummy ligand atoms [15, 16]. The dummy atoms were matched to the corresponding alpha spheres during the characterization of the LBSs in hP-gp. This method generates bound conformations that approach crystallographic resolutions [17].

Alpha Sphere and Excluded Volume-Based Ligand-Protein Docking (ASE-Dock)
The docking and analysis of the ligand-protein interaction between 6-(methylsulfinyl)hexyl isothiocyanate (6-MITC) or cyclic-tris-(R)-valineselenazole (QZ59-RRR) and hP-gp were also performed with ASE-Dock in the MOE [18]. In the ASE-Dock module, ligand atoms have alpha spheres within 1 Å. Based on this property, concave models are created and ligand atoms from a large number of conformations generated by superimposition with these points can be evaluated and scored by maximum overlap with alpha spheres and minimum overlap with the receptor atoms. The scoring function used by ASE-Dock is based on ligand-protein interaction energies and the score is expressed as a U_{total} value. The ligand conformations were subjected to energy minimization using the MMF94S force field [19]. From the resulting 500,000 poses, the 200 poses with the lowest U_{total} values were selected for further optimization with the MMF94S force field. During the refinement step, the ligand was free to move within the binding pocket.
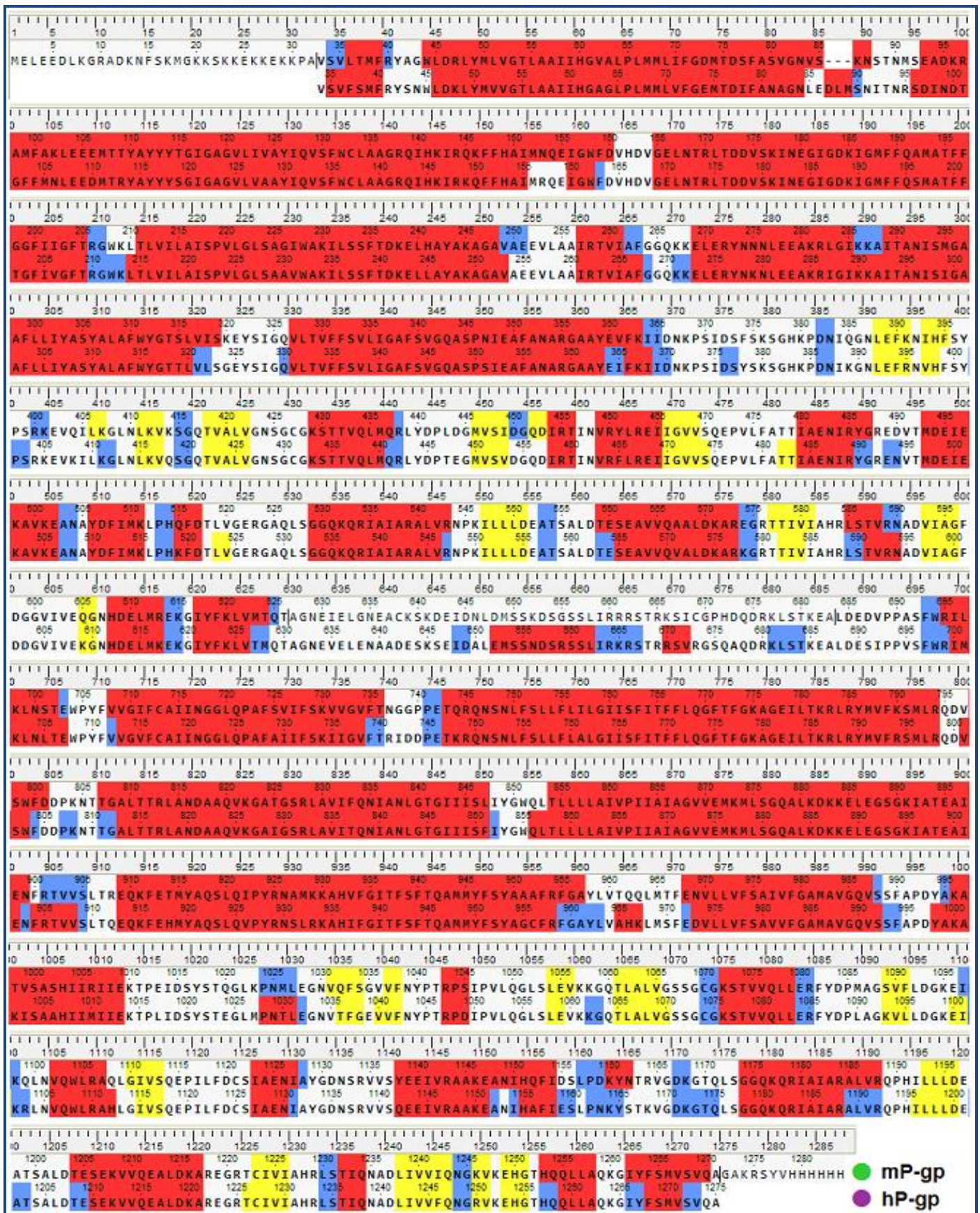
Figure 1: Evaluation of the 2D structures of the hP-gp model. Homology-aligned sequences of mP-gp (PDB code: 3G5U; green) and hP-gp (magenta). Red: α-helix. Blue: turn. Yellow: β-sheet.

Results:

Homology modeling of the hP-gp Structure

The sequence alignment of P-gp is shown in **(Figure 1)**. The alignment revealed that the critical active site residues Gln[721] (Gln[725] for hP-gp), Gln[986] (Gln[990] for hP-gp) and Ser[989] (Ser[993] for hP-gp) **[7]** at the drug biding site (DBS) were conserved in mP-gp and hP-gp. mP-gp (PDB code: 3G5U) was selected as a template **(Figure 2A)** for the present 3D structure modeling of hP-gp because of its sufficient crystal structure resolution (3.8 Å) and it was the only available structure of a mammalian P-gp with a relatively new information (from 2009) **[7]**. The % sequence identity between mP-gp and hP-gp was 88.2%. For the construction of the hP-gp model, 100 independent models of the target protein were built using a Boltzmann-weighted randomized modeling procedure in the MOE that was adapted from reports by Levitt **[11]** and Fechteler *et al.* **[12].** The intermediate models were evaluated by a residue packing quality function, which is sensitive to the degrees to which non-polar side-chain groups are buried and hydrogen bonding opportunities are satisfied. The hP-gp model with the best packing quality function and full energy minimization was utilized in the present study **(Figure 2B)**.

Analysis of the Contact Energies for the hP-gp Model

The effective atomic contact energies were calculated using the MOE for heavy atoms of standard amino acids within a contact range of 6 Å, assigning energy terms in kcal/mol for each contact pair **[13]**. These energies were summed for each residue, and in general, a large negative value indicated that the residue was predominantly in contact with hydrophobic atoms and therefore expected to be in a buried protein environment. Conversely, residues with positive energy terms indicated contacts with predominantly hydrophilic atoms, and were expected to be in more solvent-exposed regions of the proteins. The contact energy profiles of the homology-modeled hP-gp **Figure 3B (see supplementary material)** was compared with that of the X-ray structure of mP-gp **Figure 3A (see supplementary material)**. Although 1 out of 3 residues at the critical active site had differently charged contact energies, namely Gln[986] (mP-gp; -1.35 kcal/mol) and Gln[990] (hP-gp; 0.17 kcal/mol), the other residues had similarly charged contact energies, such as Gln[721] (mP-gp; -3.13 kcal/mol) and Gln[725] (hP-gp; -0.81 kcal/mol), and the trends of the variation in the contact energy in most parts of the hP-gp model were in good agreement with those of the X-ray structure of mP-gp.
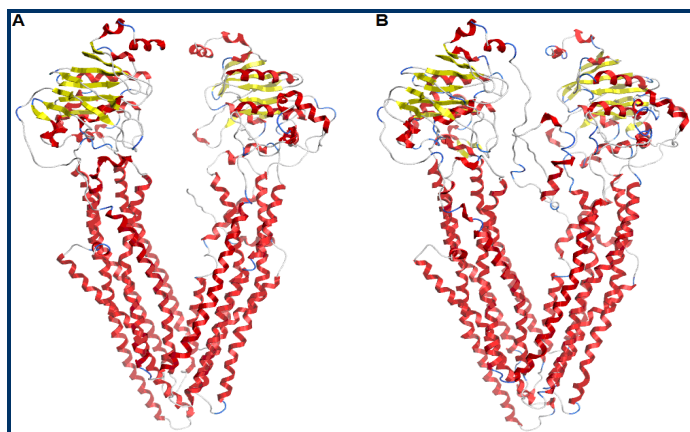
Evaluation of the Stereochemical Qualities of the hP-gp Model

The phi and psi backbone dihedral angles for hP-gp were scored using 2D probability distributions calculated on a high-resolution collection of X-ray structures containing approximately 100,000 residues from 500 protein structures **[20]**. Each probability distribution was estimated with 2-degree spacing for each of the phi and psi backbone dihedral angles with separate histograms for pre-proline, proline, glycine and general amino acids. The stereochemical qualities of the hP-gp model were assessed by Ramachandran plots **Figure 4 (see supplementary material)**. 83.0% of the residues were in the favored region, 13.5% were in the allowed region and only 3.5% were in the disfavored region. These results indicate that the phi and psi backbone dihedral angles in the hP-gp model are reasonably accurate.

Structural Comparisons of the P-gp Models

Root mean square deviation (RMSD) values between the main chain atoms of the mP-gp (3G5U) *vs* hP-gp after main chain fit were 1.3 Å **Figure 5 (see supplementary material)**. RMSD values for each residue were also analyzed. The RMSD values for the residues located in the DBS were about 2 Å or less. A superimposition of the template mP-gp (green) and hP-gp (magenta) models revealed that the P-gp models exhibited significant 3D similarities **Figure 6A (see supplementary material)**. They also presented similar structures at their DBSs **Figure 6B (see supplementary material)**.

Docking Simulations of an ITC to P-gp

It has been reported that P-gp interacts with ITCs and that certain ITCs are inhibitors of P-gp **[5]**. We recently found that 6-MITC, an ITC from a Japanese indigenous herb wasabi, demonstrated significant growth inhibitory activity toward a macrophage-like tumor cell line **[6]**, suggesting that 6-MITC can be utilized for the antitumor drug development. However, the precise mode of 6-MITC binding is unknown. The ASE-Dock **[18]** was performed to evaluate the present docking simulation and showed that 6-MITC bound at the DBS in the mP-gp **Figure 7A (see supplementary material)** and hP-gp **Figure 7B (see supplementary material)** models with the similar binding location. The similarity of the bound location at the DBS between the docked 6-MITC-mP-gp pose and the hP-gp model suggests that the present methods are capable of generating the 6-MITC-hP-gp model similar to the reported near-native P-gp complex. Further, a cyclic hexapeptide inhibitor, QZ59-RRR **[7]** was docked to the hP-gp model **Figure 7C (see supplementary material)**. The bound location of QZ59-RRR in hP-gp was similar to that of 6-MITC at the DBSs in the mP-gp and hP-gp models, which further indicates the accuracy of the present docking simulation.

Conclusion:

Examination of the hP-gp structures provides considerable insight into the drug transport mechanism of the protein and suggests approaches by which P-gp inhibitors with greater selectivity may be attainable. Further, detailed analysis of the ligand-protein interaction is of great significance in designing *in silico* hP-gp-inhibiotor models for successful development of antitumor drugs. The main objective in the present study was to create a hP-gp model. Analyses of the structural properties of the hP-gp and the docking simulations of the 6-MITC (or QZ59-RRR)-hP-gp pose suggest that the present methods are capable

Figure 2: The secondary structures of the hP-gp models **(A)** mP-gp (PDB code: 3G5U); **(B)** hP-gp.

of generating the hP-gp model similar to the near-native mP-gp. Consequently, it is proposed that the hP-gp in the present study will be suitable for further *in silico* structure-based *de novo* drug design. Furthermore, to the best of our knowledge, this is the first report of a hP-gp model with a naturally occurring isothiocyanate.

References:
[1] Schinkel AH & Jonker JW, *Adv Drug Deliv Rev.* 2003 **55:** 3 [PMID: 12535572]
[2] Pastan I & Gottesman MM, *Annu Rev Med.* 1991 **42:** 277 [PMID: 2035973]
[3] Gottesman MM & Pastan I, *Annu Rev Biochem.* 1993 **62:** 385 [PMID: 8102521]
[4] Bradshaw DM & Arceci RJ, *J Clin Oncol.* 1998 **16:** 3674 [PMID: 9817290]
[5] Hu K & Morris ME, *J Pharm Sci.* 2004 **93:** 1901 [PMID: 15176077]
[6] Noshita T *et al. Eur J Med Chem.* 2009 **44:** 4931 [PMID: 19716210]
[7] Aller SG *et al. Science.* 2009 **323:** 1718 [PMID: 19325113]
[8] Kurogi Y & Guner OF, *Curr Med Chem.* 2001 **8:** 1035 [PMID: 11472240]
[9] Jorgensen WL, *Science.* 2004 **303:** 1813 [PMID: 15031495]
[10] Yamaguchi H *et al. Eur J Med Chem.* 2011 **46:** 1325 [PMID: 21333409]
[11] Levitt M, *J Mol Biol.* 1992 **226:** 507 [PMID: 1640463]
[12] Fechteler T *et al. J Mol Biol.* 1995 **253:** 114 [PMID: 7473707]
[13] Zhang C *et al. J Mol Biol.* 1997 **267:** 707 [PMID: 9126848]
[14] Lüthy R *et al. Nature.* 1992 **356:** 83 [PMID: 1538787]
[15] Liang J *et al. Proteins.* 1998 **33:** 1 [PMID: 9741840]
[16] Liang J *et al. Proteins.* 1998 **33:** 18 [PMID: 9741841]
[17] Goto J *et al. J Med Chem.* 2004 **47:** 6804 [PMID: 15615529]
[18] Goto J *et al. J Chem Inf Model.* 2008 **48:** 583 [PMID: 18278891]
[19] Halgren TA, *J Comput Chem.* 1996 **17:** 490
[20] Lovell SC *et al. Proteins.* 2003 **50:** 437 [PMID: 12557186]

Edited by P Kangueane

Citation: Yamaguchi *et al. Bioinformation 8(22): 1066-1074 (2012)

Supplementary material:

Figure 3: Contact energy profiles of the P-gp models. **(A)** mP-gp (PDB code: 3G5U). The 3D structure of the mP-gp residues between Ala[630] and Ala[686] is somewhat unavailable in PDB. Therefore, the contact energies of these residues are unable to be calculated; **(B)** The constructed hP-gp model. The positions of the amino acid residues are shown on the x-axis, while the contact energies are shown on the y-axis.
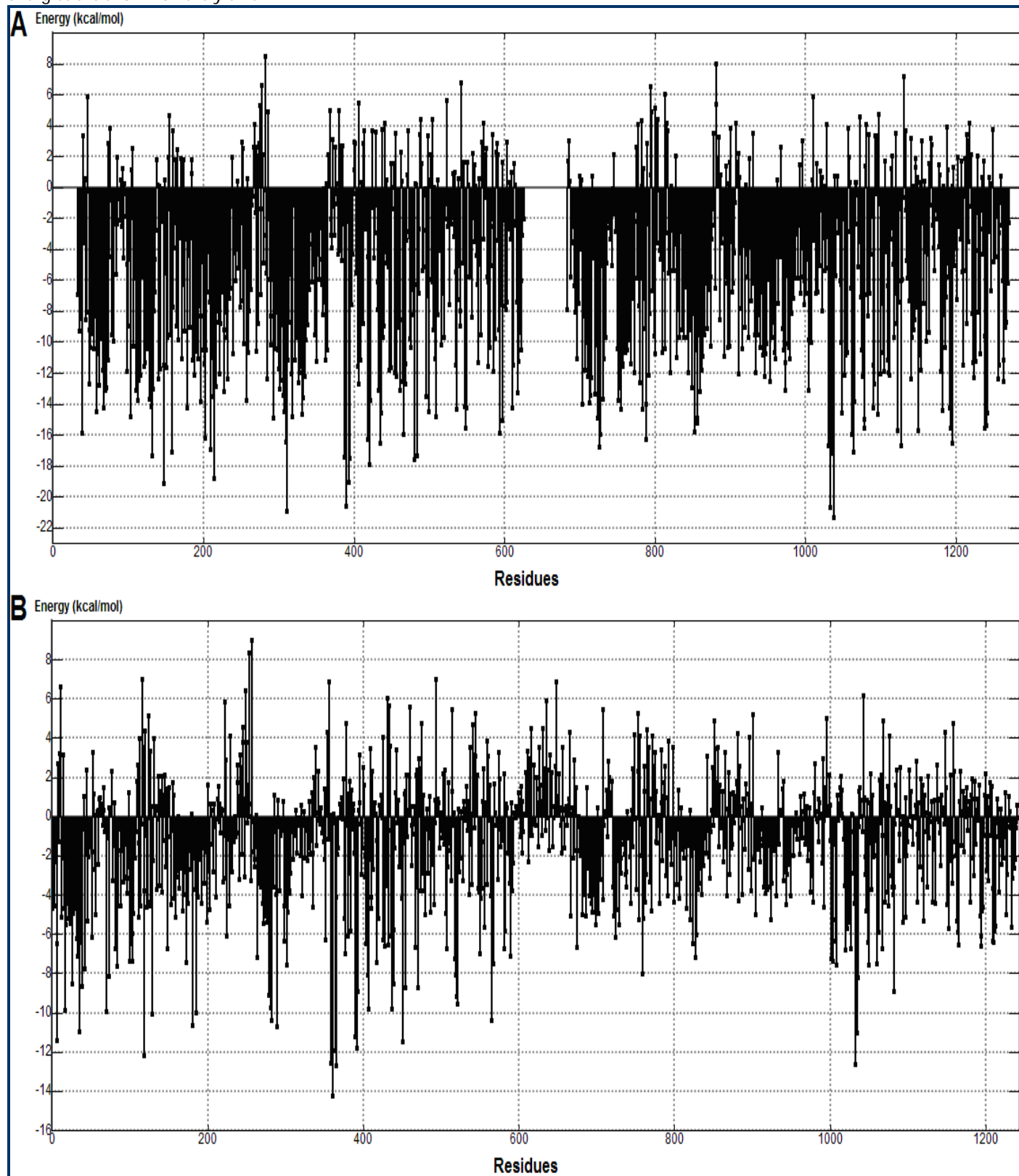
Figure 4: Ramachandran plots for the hP-gp model. Blue: favored region. Light-brown: allowed region.
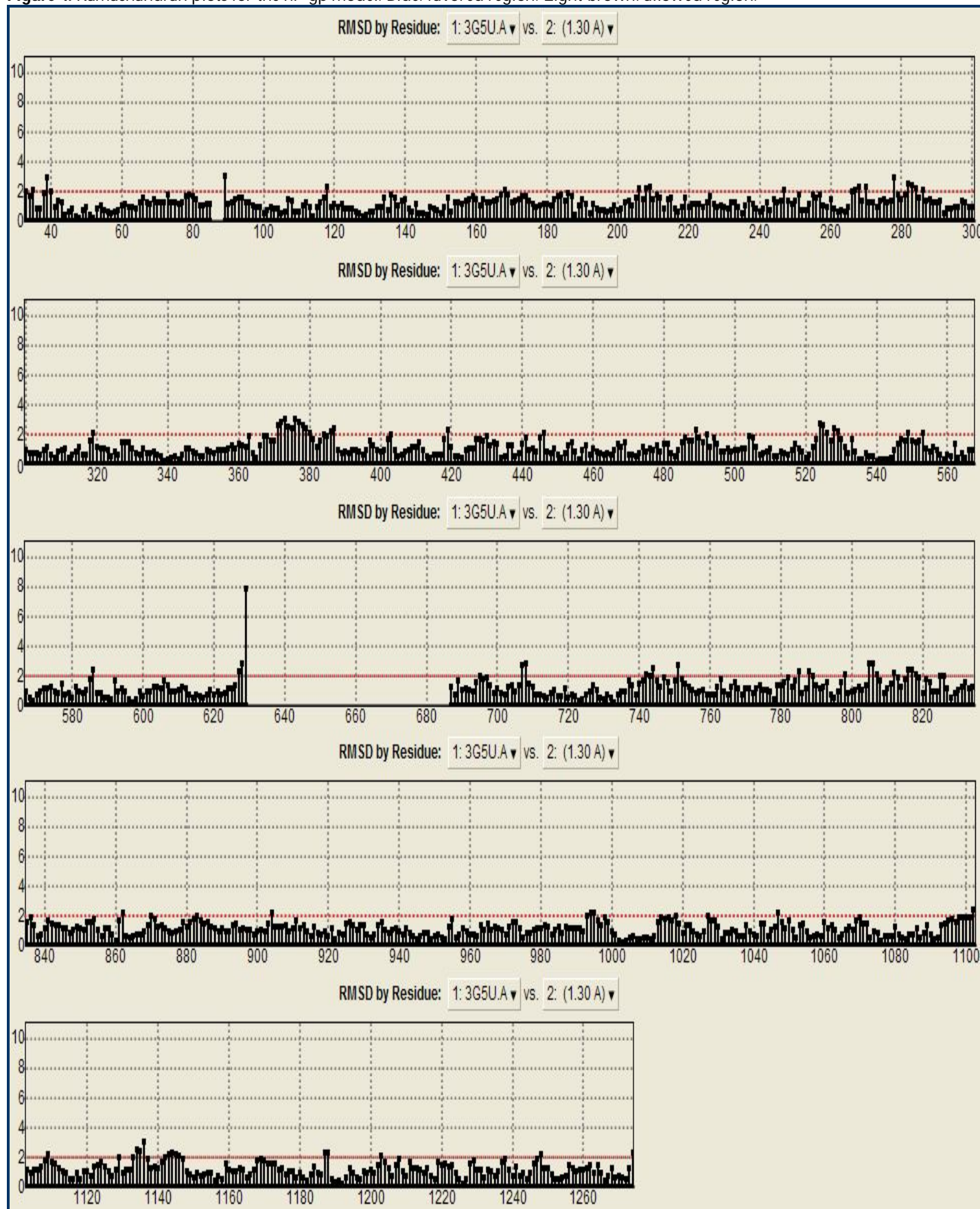
BIOINFORMATION

Figure 5: RMSD values between the main chain atoms of the mP-gp (PDB code: 3G5U) vs hP-gp after main chain fit. The 3D structure of the mP-gp residues between Ala630 and Ala686 is somewhat unavailable in PDB. Therefore, the RMSD values of these residues are unable to be calculated. The positions of the amino acid residues are shown on the x-axis, while the RMSD values are shown on the y-axis
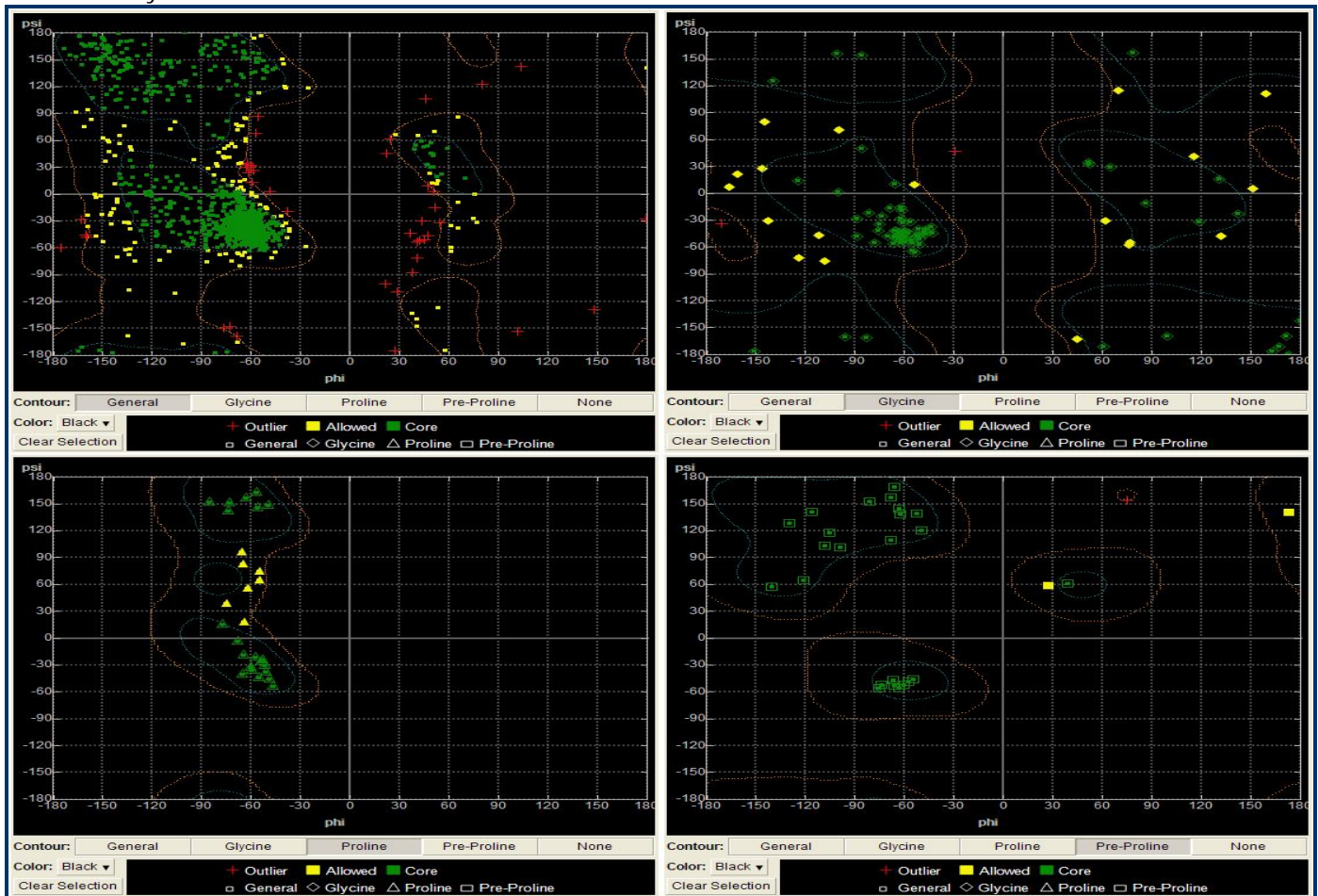

Figure 6: Structural comparison of the P-gp models: **(A)** A superimposition of the template mP-gp (PDB code: 3G5U; green) and hP-gp (magenta) models; **(B)** A superimposition of the mP-gp (green) and hP-gp (magenta) models at the DBSs. Blue: nitrogen. Red: oxygen. Yellow: sulfur.
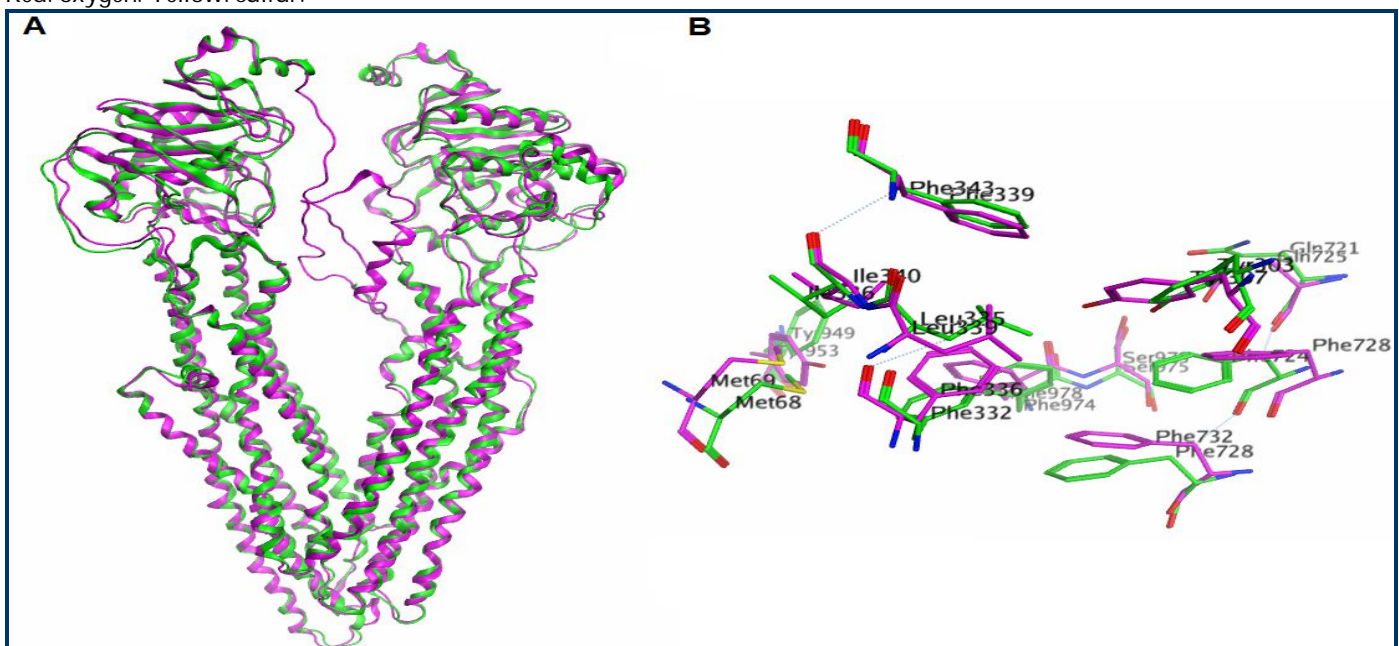
Figure 7: Docking simulations of 6-MITC and QZ59-RRR to the P-gp models. **(A)** ASE-Dock of 6-MITC for mP-gp (PDB code: 3G5U); **(B)** ASE-Dock of 6-MITC for hP-gp; **(C)** ASE-Dock of QZ59-RRR for hP-gp. Blue: nitrogen. Gray: carbon. Red: oxygen. Yellow: sulfur.