Available online at www.sciencedirect.com
ScienceDirect
Journal of Structural Biology 162 (2008) 63–74
Journal of Structural Biology
www.elsevier.com/locate/yjsbi

Homology modeling of breast cancer resistance protein (ABCG2)

Eszter Hazai, Zsolt Bikádi *

Virtua Drug, Ltd., H-1015 Budapest, Csalogány Street 4, Hungary

Received 14 September 2007; received in revised form 16 November 2007; accepted 6 December 2007
Available online 15 December 2007

Abstract

BCRP (also known as ABCG2, MXR, and ABC-P) is a member of the ABC family that transports a wide variety of substrates. BCRP is known to play a key role as a xenobiotic transporter. Since discovering its role in multidrug resistance, considerable efforts have been made in order to gain deeper understanding of BCRP structure and function. The recent study was aimed at predicting BCRP structure by creating a homology model. Based on sequence similarity with known structures of full-length, NB and TM domain of ABC transporters, TM, NB, and linker regions of BCRP were defined. The NB domain of BCRP was modeled using MalK as a template. Based on secondary structure prediction of BCRP and comparison of the transmembrane connecting regions of known structures of ABC transporters, the TM domain arrangement of BCRP was established and was found to resemble to that of the recently published crystal structure of Sav1866. Thus, an initial alignment of TM domain of BCRP was established using Sav1866 as a template. This alignment was subsequently refined using constrains derived from secondary structure and TM predictions and the final model was built. Finally, the complete homodimer ABCG2 model was generated using Sav1866 as template. Furthermore, known ligands of BCRP were docked to our model in order to define possible binding sites. The results of molecular dockings of known BCRP substrates to the BCRP model were in agreement with recently published experimental data indicating multiple binding sites in BCRP.

Keywords: BCRP; ABC transporter; Homology modeling; Transmembrane prediction; Secondary structure prediction; Molecular docking

1. Introduction

The ABC transporter superfamily is among the largest and most broadly expressed protein families (Dean et al., 2001). Members of this family translocate their substrates across extra- and intracellular membranes. The proteins use ATP to actively efflux a wide array of substrates from the cell including exogenous compounds such as drugs, and drug metabolites, as well as endogenous substances such as bile acids, peptides, steroids, ions, and phospholipids. They possess a relatively conserved structure which contains a combination of conserved nucleotide binding (NB)[1] and transmembrane (TM) domains. BCRP also seem to play pivotal role in membrane transporter mediated resistance (Choudhuri and Klaassen, 2006), along with P-gp (MDR1, ABCB1) and multidrug resistance protein 1 (MRP1, ABCC1). Inhibitors of the major ABC transporters that contribute to drug resistance have been developed and extensive research is being performed in an attempt to block the development of drug resistance in chemotherapy.

The G subfamily of ABC transporters—to which BCRP belongs- consists of half-transporters with a unique domain arrangement, i.e. *the NB domain is located at the N-terminus of the polypeptide chain followed by a linker region and transmembrane domain.* Since BCRP possesses one NB and one TM domain, it requires dimerization for expressing activity. There is experimental evidence that BCRP can form functional homodimer, since the amplification of abcg2 gene alone is sufficient to circumvent drug toxicity in drug-selected cell lines (Miyake et al., 1999; Knutsen et al., 2000). It was suggested that ABCG2 forms a homodimer bridge by disulfide bonds (Kage et al., 2002).

* Corresponding author.
E-mail address: Zsolt.Bikadi@virtuadrug.com (Z. Bikádi).

[1] *Abbreviations used:* MDR, multidrug resistance; BCRP, ABCG2, breast cancer resistance protein; TM, transmembrane; NB, nucleotide binding; DAS, Dense Alignment Surface.

1047-8477/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2007.12.001

The active dimer is spontaneously formed in insect cells overexpressing the transporter (Ozvegy et al., 2002).

Although general mechanism of ABC transporters may exist, there are also differences in ligand–protein interaction and in translocation mechanism as well within the three major human ABC transporters (Clark et al., 2006). Therefore, detailed characterization of a given ABC transporter structure is needed for exploring the exact mechanism of function. Thus, efforts have been made in order to explore the three-dimensional structure of ABC proteins.

There are crystal structures of isolated NB domains available. The first published NB structure was the ATP-binding subunit of the histidine permease (Hung et al., 1998). Since then, besides a number of structures of bacterial ABC transporter NBs, human ABC transporter NBs have been solved as well. The structure of TAP1–NB (ABCB2), was determined in complex with Mg^{2+}/ADP (Gaudet and Wiley, 2001) and NB1 of human CFTR (ABCC7) with Mg^{2+}/ATP (Lewis et al., 2005). In addition, the structure of human wild-type MRP1–NB1 (Ramaen et al., 2006) was published recently. As can be expected from the relatively highly conserved amino acid sequence of NBs throughout the ABC transporter family, the three-dimensional structures of NBs were found to be highly conservative. The close interaction of the two NBs in ABC transporters is likely to result in the formation of a fully competent catalytic site. It has been found that the interaction of the two NB units is an essential requirement for the catalytic reaction. Several lines of evidence indicate that both NBs can bind ATP, and both catalytic sites are active (Moody et al., 2002).

To date, no complete high-resolution structure of eukaryotic ABC transporters is available as eukaryotic transporters are more refractory to structural analysis than bacterial transporters. A low resolution (25 Å) structure of P-gp of was reported in 1997 determined by electron diffraction (Higgins et al., 1997). This structure depicts that P-gp has a large central chamber within the membrane. It is closed at the cytoplasmic side of the membrane and it opens to the lipid phase. More recently, crystal structure of P-gp from Chinese hamster was reported at 8 Å resolution obtained by cryo-electron crystallography (Rosenberg et al., 2005). Unfortunately, this low resolution does not provide sufficient information on the exact structure and mechanism of function.

Full-length crystal structures containing NB and TM domains of bacterial ABC-transporters (*Escherichia coli* vitamin B12 importer BtuCD (Locher et al., 2002), molybdate importer from *Archaeoglobus fulgidus* (Hollenstein et al., 2007), a metal chelate importer from *Haemophilus influenzae* (Pinkett et al., 2007) and the multidrug exporter Sav1866 from *Staphylococcus aureus* (Dawson and Locher, 2006)) have been determined. (It should be noted that the crystal structure of bacterial MsbA lipid flippases has recently been distracted.)

In our study, a BCRP model was built based on the crystal structure of Sav1866 as template of the transmembrane region and as a template for generating the final homodimer model.

Our new model combines current knowledge on BCRP structure obtained from transmembrane prediction, secondary structure prediction and mutagenesis data. Furthermore, known ligands of BCRP were docked to our homology model using blind docking method. The results were shown to conform experimental data on BCRP structure, therefore might contribute to understanding the mechanism of BCRP action.

2. Materials and methods

The whole procedure carried out in order to construct a homology model is described in Fig. 1.

2.1. Generation of homology model of BCRP

Primary sequences of the proteins were taken from the Swiss-Prot TrEMBL database (http://au.expasy.org/sprot/). Crystal structures were obtained from the Brookhaven Protein DataBank (http://www.rcsb.org/pdb/). The initial sequence alignment of NB and TM regions of BCRP with the available NB (PDB codes: 1Q12, 2CBZ, 1JJ7, 1L2T) and TM (PDB code: 2HYD) structural templates were carried out using the ClustalW program (Thompson et al., 1994). Four regions in BCRP have been identified, namely N-terminal hydrophilic segments (amino acids 1–40); NB segment (41–299); linker region (299–363); TM region (364–655). Amino acids 1–58 and the linker region between NB and TM domains were not considered in the final model because no appropriate templates have been found.

2.2. Modeling the NB domain of BCRP

Available structural templates for NB domain were identified by Blast Search (http://www.ncbi.nlm.nih.gov/BLAST) and using an iterative hidden Markov model-based method (SAM-T04) (Karplus et al., 2005). Best templates are listed in Table 1. Multiple sequence alignment was carried out using SAM-T04 method (Karplus et al., 2005). The NB domain of BCRP has been modeled on the basis of the crystal structure of the ATP-binding protein of the *E. coli* maltose/maltodextrin transporter (PDB code: 1Q12) (Chen et al., 2003a) possessing the highest sequence identity with the NB domain of BCRP. The homology modeling was based on the SAM-T04 multiple alignment. Full-length transporters were omitted from NB modeling, because the NB conformation observed might be the result of crystallization conditions and probably does not represent a physiologically relevant conformation.

Three-dimensional (3D) models were built by the MODELLER6 package (Sali and Blundell, 1993). A bundle of 10 models from random generation of the starting structure was calculated. Subsequently, the best model (with the