Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.
Molecular Pharmacology Fast Forward. Published on October 27, 2004 as doi:10.1124/mol.104.006973

P-glycoprotein substrate binding domains are located at the transmembrane domain : transmembrane domain interfaces - A combined photoaffinity labeling - protein homology modeling approach

Karin Pleban, Stephan Kopp, Edina Csaszar, Michael Peer, Thomas Hrebicek, Andreas Rizzi, Gerhard F. Ecker and Peter Chiba

Institute of Medical Chemistry (K.P., S.K., M.P., P.C.), Medical University of Vienna, Department of Pharmaceutical Chemistry (K.P., G.F.E.), Institute of Analytical Chemistry (T.H., A.R.) and the Mass Spectrometry Facility, Max F. Perutz Laboratories, University Departments at the Vienna Biocenter (E.C.), University of Vienna, Austria

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

Running title: Substrate binding domain of P-glycoprotein

Address correspondence to:

Peter Chiba, Institute of Medical Chemistry, Medical University of Vienna,

Waehringerstrasse 10, A-1090 Vienna, Austria

Phone: +43 1 4277 60806, Fax: + 43 1 4277 60889

e-mail: peter.chiba@meduniwien.ac.at

Abbreviations

ABC-transporter, ATP-binding-casette transporter; TMD, transmembrane domain; NBD, nucleotide binding domain; TM, transmembrane segment; MALDI-TOF mass spectrometry, matrix assisted laser desorption/ionization-time-of-flight mass spectrometry

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

Abstract

P-glycoprotein (P-gp) is an energy dependent multidrug efflux pump conferring resistance to cancer-chemotherapy. Characterization of the mechanism of drug transport at a molecular level represents an important prerequisite for the design of pump inhibitors, which resensitize cancer cells to standard chemotherapy. In addition, P-glycoprotein plays an important role for early ADMET (**A**bsorption, **D**istribution, **M**etabolism, **E**xcretion and **T**oxicity) profiling in drug development. A set of propafenone-type substrate photoaffinity ligands has been used in this study in conjunction with MALDI-TOF mass-spectrometry to define the substrate-binding-domain(s) of P-gp in more detail. Highest labeling was observed in transmembrane segments 3, 5, 8 and 11. A homology model for P-gp was generated based on the dimeric crystal structure of *Vibrio cholerae* MsbA (Vc-MsbA), an essential lipid transporter. Subsequently, the labeling pattern was projected onto the 3D atomic detail model of P-gp to allow a visualization of the binding domain(s). Labeling is predicted by the model to occur at the two TMD:TMD interfaces formed between the amino- and carboxy terminal half of P-gp. These interfaces are formed by TM segments 3 and 11 on one hand and TM segments 5 and 8 on the other hand. Available data on LmrA and AcrB, two bacterial multidrug efflux pumps, suggest that binding at domain interfaces may be a general feature of polyspecific drug efflux pumps.

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

Introduction

Multidrug resistance (MDR) represents a serious obstacle to successful cancer chemotherapy. Though multi-factorial in etiology, one type of MDR is associated with the overexpression of energy dependent membrane-bound pumps, which intercept and efflux drugs before they reach their intracellular target structures. P-glycoprotein (ABCB1) represents a paradigm ATP-dependent efflux pump expressed in human cancer cells. In addition to its expression in cancer cells, P-gp is also physiologically expressed in a number of tissues such as intestinal epithelial cells, at the brush border of renal tubule epithelial cells, the canalicular side of hepatocytes and in capillary endothelial cells forming the blood brain barrier. It thus interferes with oral drug absorption, drug delivery to the brain and enhances renal and biliary excretion. P-gp has therefore attracted considerable attention as a non-target in the field of drug development, since for a large number of active compounds interaction with P-glycoprotein might compromise their future development into a drug. Considerable energy has therefore been devoted to the characterization of molecular features, which make compounds P-gp substrates and to the definition of the molecular mechanism of drug transport by P-gp. A number of studies have dealt with the kinetics and thermodynamics of the transport process (Senior & Gadsby 1997, Sauna et al., 2000, 2001, Al-Shawi et al., 2003). In addition, site directed mutagenesis has been used to replace individual amino acid residues and characterize mutant P-gp with respect to ATPase activity, substrate binding and transport characteristics (Loo and Clarke 2004a, b and references contained therein, Hafkemeyer et al., 1998). P-gp has been overexpressed, purified and reconstituted successfully in order to obtain biological material for structural resolution of the transporter. This endeavor has

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

resulted in low resolution structures obtained by electron-cryo microscopy (Rosenberg et al., 2001, 2003). Nevertheless an atomic detail structure of P-gp or of any other ATP-dependent multidrug transporter is not available to date. Photoaffinity labeling studies have identified regions, which are accessible for substrates (for a review see Peer et al., 2004). Atomic details of the substrate binding domain and of the dynamics of the transport process, however, still remain elusive.

In the present study a set of substrate photoaffinity ligands related to the lead compound propafenone were used to photolabel P-gp. Following gel electrophoretic separation and in gel proteolytic degradation of the P-gp band, ligand-labeled peptides were identified by MALDI-TOF-mass spectrometry. A homology model for P-gp was established using the essential lipid transporter MsbA from *Vibrio cholerae* as a template. The information obtained by photoaffinity labeling studies was projected onto this model indicating two pseudosymmetric binding domains formed by TM segments 3 and 11 (corresponding to TM segment 3 of the N-terminal and TM5 of the C-terminal TMD) on one hand and TMs 5 and 8 (corresponding to TM segment 5 of the N-terminal and TM2 of the C-terminal half of P-gp) on the other hand. Data indicate that the initial binding of substrates occurs at the TMD:TMD interfaces. A reconciliation of the model with cross linking data of double cysteine mutants has been possible. Data indicate that entry of substrates *via* the TMD:TMD interface might represent a common property of multidrug efflux pumps.

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

Materials and Methods

Design and synthesis of compounds - The design and synthesis of propafenone-type photo-affinity ligands has been described previously (Ecker et al., 2002). Structures were confirmed by IR, MS, NMR spectroscopy and elemental analysis and are shown in Table 1.

Cell lines - Insect Sf9 cells were obtained from the American Type Culture Collection (Rockville, MD, USA).

Preparation of plasma membrane vesicles - Plasma membrane preparations of Sf9 cells transfected with the baculovirus construct containing the his-tagged mdr1-gene (Germann, 1990) were prepared by nitrogen cavitation and subsequent discontinuous sucrose-gradient centrifugation as described (Schmid et al., 1999).

Photolabeling of P-gp and gel electrophoretic separation conditions - Plasma membrane-vesicles were taken up in 50 mM Tris-HCl, pH 7.4 and preincubated with ligand at a concentration of 10nmol/l. Samples were preincubated at room temperature for 15 min. Subsequently, samples were placed on ice and irradiated intermittently with a 500W Hg-lamp (Lot-Oriel, Darmstadt, Germany) for 6 times 30 sec in the presence of the photo-activatable propafenones. Control samples were irradiated in the absence of photoligands. A 1mm glass plate was placed in the light path to filter most of the UV light with wavelengths below 300nm. After irradiation, samples were centrifuged at 50000 x g for 30 min at 4°C. Protein pellets (30µg of total plasma membrane protein per lane) were taken up in 1 x SDS/sample buffer, loaded on a 10% polyacrylamide gel and run at 35mA for 60min using a Hoefer Mighty Small II SE250 unit (APBiotech, Vienna, Austria). Gels were fixed in methanol : glacial acetic acid : water (50:10:40) for 30 min

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

and washed overnight in double distilled water. Bands were visualized by silver staining as described (Ecker et al., 2002) and the P-gp band was excised and subjected to in gel-digestion as described below.

Chemicals for in-gel digestion and mass spectrometry - High quality water for the in-gel digestion and the mass spectrometric experiments was prepared using a Milli-Q water purification system (Millipore, Bedford, Massachusetts, USA). Ammonium hydrogen carbonate was obtained from FLUKA (Sigma-Aldrich, Vienna, Austria), dithiothreitol (DTT) for the reduction of proteins before in-gel digestion was purchased from Serva (Novex, San Diego, CA, USA), iodoacetamide was supplied by Sigma (Sigma-Aldrich, Vienna, Austria), chymotrypsin was obtained from Roche (sequencing grade, Roche Diagnostics GmbH, Mannheim, Germany). Acetone was from AppliChem (Darmstadt, Germany), acetonitrile, methanol and isopropanol were purchased from Sigma (Sigma-Aldrich, Vienna, Austria; HPLC grade), trifluoroacetic acid (TFA) was obtained from Pierce (Rockford, Illinois, USA) and formic acid was supplied by VWR (VWR Intl, Darmstadt, Germany). The α-cyano-4-hydroxycinnamic acid (CHCA) matrix for the MALDI-TOF measurements was purchased from Bruker (Bruker Daltonik GmbH, Bremen, Germany) and used without further purification. Nitrocellulose was purchased from Bio-Rad (Bio-Rad Laboratories, Hercules, CA, USA). Protein and peptide standards used to calibrate the MALDI mass spectrometer were obtained from Bruker (Part No.: 206195). In-gel digestions were performed in analogy to the procedure described in Ecker et al. (2002).

MALDI-TOF mass spectrometry - A Bruker REFLEX III (Bruker Daltonik GmbH, Bremen, Germany) MALDI-TOF instrument, equipped with a standard nitrogen

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

laser (337 nm) was used for mass-spectrometry. The spectra were recorded in reflectron mode, positive ionization and with an acceleration voltage of 25 kV. The laser power was varied on a relative scale of 0-100 and was kept at the threshold value to obtain appropriate signal intensity. The calibration of the instrument was done externally. Samples were prepared with a 75:25 (v/v) mixture of CHCA matrix (saturated solution in acetone) and nitrocellulose (10 mg/ml solution in acetone-isopropanol 50:50 (v/v)). A 1µl aliquot of the mixture was placed onto the sample slide and allowed to dry at room temperature. In-gel digested P-gp (0.5µl) was mixed with 0.5µl 0.1 % TFA on this thin layer of matrix crystals and vacuum dried. Samples were washed with ice-cold 0.1 % TFA. Hydrophobic peptides were purified and concentrated on Poros 20 R1 material (Applied Biosystems, Foster City, CA) loaded into GeLoader tips (Eppendorf-Netheler-Hinz-GmbH, Hamburg, Germany). The chromatography material was conditioned with 0.1 % TFA and the peptides were eluted with CHCA matrix (saturated solution in 0.1 % TFA-acetonitrile 50:50 (v/v)) directly onto the sample slide. Each spectrum was produced by accumulating data from 90-120 consecutive laser shots. Spectra were interpreted with the aid of the Mascot (Matrix Science Ltd, London, UK) or MS-Fit (Clauser et al., 1999) software using the NR database (NCBI Resources, NIH, Bethesda, MD, USA). Ligand modified masses were matched to peptide masses generated by *in silico* digests of the protein with the aid of a custom program developed in our laboratory.

Data analysis

Statistical analysis was performed using the unpaired Student’s t-test with a 95% confidence interval for the sample mean. The number of independent observations was 6.

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

Generation of a protein homology model of P-gp using Vc-MsbA as template

Primary sequences of the proteins are deposited in the Swissprot - TrEMBL database with the accession numbers P08183 (P-gp), Q03518 (TAP1) and Q9KQW9 (Vc-MsbA). Crystal structures were obtained from the PDB (Brookhaven Protein Databank). Sequence alignments were generated in InsightII, release 2000 (Accelrys, San Diego, CA) implemented on a Silicon Graphics Octane R12000 workstation. A geometric validation of the 3D model was carried out using the ProStat/Structure Check and/or Verify3D modules in InsightII. Minimization at splice points was performed using the CHARMm forcefield. Graphics were generated in MOE (Chemical Computing Group, Cologne, Germany) and VMD (Humphrey et al, 1996).

In order to identify putative transmembrane helices we initially submitted the N- and C-terminal halves of Pgp to the MPEx membrane protein explorer (Jaysinghe et al, 2000). Subsequently the N- and C-terminal halves of P-gp (1- 614; 697- 1280) were aligned with Vc-MsbA separately. Amino acids 1-33 and the 6 C-terminal amino acids were not considered in the model as they have no counterpart in MsbA. In addition the linker region extending from amino acid position 629 to 697 was deleted from the sequence, since evidence for a functional relevance of this protein sequence is lacking (Seigneuret et al., 2003). The alignment was performed using the Blosum62 matrix with a gap penalty of 11 and a gap extension penalty of 1, yielding an alignment identical to that previously suggested by Chang (Chang 2003). This alignment was carefully checked in order to avoid deletions or insertions in conserved regions and in transmembrane helices and, if necessary, corrected manually. The sequences corresponding to the ICL (intracellular loop) 2 and ICL4 (AA250-284 and AA893-927) were not considered in the

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

homology model as they are not resolved in the Vc-MsbA crystal structure indicating a highly flexible nature of this loop (Chang 2003). Finally the homology model was minimized at splice points with 500 cycles steepest descent algorithm followed by 550 cycles of conjugate gradient minimization. For final model assembly, the N- and C-terminal P-gp halves were superimposed onto Vc-MsbA and merged.

In addition to the three full length crystal structures of bacterial ABC-transporters also the X-ray structures of a number of nucleotide binding domains (NBDs) of other ABC-transporters are available. These structures suggest that the NBD conformation observed in Vc-MsbA is a result of crystallization conditions and likely does not represent a physiologically relevant conformation. Therefore, we decided to model the NBDs based on the crystal structure of TAP1, which in this domain shows highest sequence homology with P-gp. The NBD model for P-gp was obtained in analogy to the procedure described in detail for LmrA (Ecker et al., 2004). Sequences of the N-terminal and C-terminal NBD were aligned separately using the Blosum 62 matrix with a gap penalty of 11 and a gap extension penalty of 1. The alignment was manually refined to avoid deletions and/or insertions in conserved regions. The models were again minimized at splice points with a cycle of 500 steps Steepest Descent Algorithm minimization.

Generation of the Pgp model was therefore performed in three steps: in the first step, a model of the TMDs and the NBDs was generated using Vc-MsbA as template. In the second step an alternate NBD model was generated, which, in a third step, replaced the Vc-MsbA-type NBDs in analogy to a method previously published for the bacterial transporter LmrA (for details see Ecker et al., 2004). The quality of the models was

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

assessed by a ProStat structure check, indicating that 98 % of the residues were in allowed regions of the Ramachandran plot.

Cross linking data and low resolution electron microscopy studies suggest that the transmembrane domains of P-gp form a funnel shaped pore that is closer at its cytoplasmatic end than at the extracellular surface. As the model presented in Fig. 5 (Model A) shows an inverted funnel shaped pore, we created an alternate model by rotating the two TMD domains of P-gp about a point in the leaflet interface. Subsequent repositioning of the NBDs and recalculation of the TMD-NBD linker regions gave the final model (Model B) as shown in Figure 6. Both models conform to the majority, but not all distance constraints obtained by cysteine cross-linking data. Distances in our model appear to be similar to those obtained in a previously published P-gp model (Stenham et al., 2003) (Table 2).

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

Results

Propafenones are substrates of P-glycoprotein

Propafenone analogues have previously been shown to be accumulated in inside-out plasma membrane vesicles in an ATP-dependent and cyclosporine A inhibitable manner (Schmid et al. 1999). A net flux of compounds was only measurable, when the compounds carried a permanent positive charge at the nitrogen atom. In contrast, compounds with a tertiary nitrogen atom diffuse through membranes quickly and thus a net flux of these compounds is not detectable (Schmid et al., 1999). However, similar to quaternary compounds, propafenone analogues containing a tertiary nitrogen atom caused a shift in fluorescence when P-gp was reacted with the conformation sensitive antibody UIC2 (Suk Pak and Peter Chiba, unpublished). In addition, a number of studies from our group demonstrated that propafenones inhibit the transport of known P-gp substrates including rhodamine123, daunorubicin (Chiba et al., 1996, Ecker et al., 1999, Ecker et al., 2002), mitoxantrone and vinblastine (Stephan Kopp and Peter Chiba, unpublished). Photolabeling of plasma membrane vesicles with the radioactive analog [^{3}H]GPV51 was demonstrated to be highly specific for P-gp and other plasma membrane proteins were not labeled to any significant extent. Also non radioactive GPV51 competed with binding of [^{3}H]GPV51 in a concentration dependent manner (Ecker et al., 2002).

Identification of substrate-binding domains of P-gp

P-gp was overexpressed in Sf9 insect cells using a baculovirus expression system (Germann et al., 1990). Plasma membrane vesicles were prepared from these cells by nitrogen cavitation and labelled with propafenone-type photoligands. Structures of these

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

photoaffinity ligands are given in Table 1. As shown previously, these benzophenone type ligands exhibited an unchanged quantitative structure activity relationship when compared to parental phenylpropiophenones (Ecker et al., 2002). Upon irradiation, compounds are photoactivated in the carbonyl group, which represents a pharmacophoric substructure of these compounds (Ecker et al., 2002). After irradiation, membrane proteins were separated by SDS-PAGE and the bands were visualized by silver staining. The 140kD protein corresponding to core glycosylated P-gp was excised and proteolytically degraded in-gel by chymotrypsin digestion. The ensuing peptide fragments were eluted from the gel and identified by high resolution MALDI-TOF (matrix-assisted laser desorption/ionization time-of-flight) mass spectrometry. Since the efficiency of the photochemical reaction is low, the majority of peaks in the mass spectrum corresponded to unmodified peptide fragments and allowed to assess the coverage of the sequence with unmodified peptide fragments. The α-helical TM-segments of the TMDs, which are considered to be responsible for solute binding and translocation, were in their entirety represented by ligand-unmodified peptide fragments. Overall, the sequence coverage in the TMDs exceeded 95%, while that in the NBD was approximately 80%. Mass peaks were assigned to P-gp by comparison with a list of theoretical peptide masses obtained by *in silico* proteolytic degradation of the protein with chymotrypsin. Since chymotrypsin recognition sites are frequent in TM-segments, experiments were performed under incomplete digest conditions. This prevented loss of sequence coverage in protein regions with clustered protease recognition sites. Under these conditions up to seven missing cleavage sites were observed and thus *in silico* digest files were generated using this information.

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

Ligand-modified peptide fragments were assigned after subtracting the exact mass of the ligand. As evidence accumulated that more than one ligand can be adopted in binding pockets of multispecific proteins at the same time (see Neyfakh 2002 and Yu et al, 2002 for recent reviews), binding of more than one ligand to the same peptide fragment was also considered. Indeed a number of masses were candidate peaks for fragments modified by two ligands. These peptides were included in the frequency distribution analysis shown below.

Six ligands of different mass were used for the experiments, in order to improve the signal to noise ratio. The rational of these experiments was as follows: peptide fragments which are covalently modified by the photoligand, shift from their original position in the mass spectrum to an m/z, which is increased by the ligand mass. Since only a small fraction of the peptides is photochemically modified, two peaks are then detected, one at an m/z corresponding to the unmodified peptide and one with an m/z corresponding to the sum of the masses of the peptide fragment and the ligand. With a certain albeit low probability, this peak might be assigned incorrectly. Use of ligands with different mass and determination of consensus binding regions, however, minimizes any potential influence of incorrect peak assignments. We did not attempt to use other classes of photoactivatable compounds, since quantitative structure activity relationship studies for these ligands are not available.

The number of fragments containing a particular amino acid was counted and plotted as a function of this amino acid position as described in Ecker et al. (2004). The resulting frequency distribution analysis of photolabelling is shown in Fig. 1. Predicted TM segments are shown in magenta. The remainder of the sequence is shown in black.

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

Benzophenone type ligands have initially been reported to be non-selective (see Dorman and Prestwich, 2000). Recently, a preferred reaction with methionine residues in proteins has been reported (Rihakova et al. 2002, Wahlstrom et al. 2002). Methionine residues in Fig. 1 were therefore highlighted in red. Indeed, a number of methionine residues were identified to be located in peaks (M197, M440, M701 M791, M969 and M1027) or close to peaks of the trace (M68, M69, M105, M111, M796, M948 and M949). This indicates that the frequency distribution analysis of Fig. 1 correctly reflects the preference of the photochemical reaction. It is however important to note that not all peaks contain methionine residues, which is in agreement with the ability of benzophenones to react with aliphatic residues (Dorman & Prestwich, 2000). For example predicted TM segments 5 and 8 are strongly labelled, but do not contain methionine residues. Fig. 1 also shows that the presence of a methionine residue is not a sufficient precondition for reaction with the photoligands, since a number of methionine residues are not labelled (M89, M192, M450, M497, M878, M1010). These residues are thus inaccessible for the photoligands.

Inspection of Fig. 1 indicates four major sites of labelling within predicted TM segments. These are putative TMs 3 and 5 in the N-terminal half and TM segments 8 and 11 in the C-terminal half of P-gp. Strong labelling is also detected in the loop region connecting TM segments 9 and 10. Minor labelling is observed N- and C-terminal of predicted TM1, N-terminal of TM2, in putative TM6, at the N-terminus of putative TM7 and C-terminal of TM8. The methionine residue M969, which is located in the loop region connecting TM11 and 12 (ECL6) is located in the tip of a small riding peak C-terminal of the strongly labelled TM11. According to the P-gp models presented in detail

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

below (Figs. 5 and 6), these sites of minor labelling as well as the strongly labelled loop region between TM9 and 10, are located close to the predicted extra- and intracellular membrane-water interfaces. Since the affinity ligands are expected to electrostatically interact with phosphate head groups of the membrane phospholipids in protonated form, this labelling at the membrane surface did not come as a surprise. However these sites were scattered and only had in common their position at the respective membrane-water interfaces. Labelling at these positions, which lie outside the predicted transmembrane segments, is thus unlikely to be related to substrate transport.

Peak areas in Fig. 1 provide a measure of the density of labelled peptide fragments found in a particular protein region and were thus used to further discriminate between regions of high and low labelling. Peak areas were expressed as the sum of those amino acid residues in ligand modified fragments, which overlapped with the respective TM segment. Data are presented for each of the six photoligands individually to show the reproducibility of the results. In order to decrease the influence of surface labelled residues (by protrusion of labelled fragments from loops into TM-segments), the three most N-terminal and C-terminal amino acid residues of each predicted TM-helix were excluded from calculation. Figs. 2A-F represent data for each of the six photoaffinity ligands, respectively. For ligands GPV51, GPV317, GPV319, GPV442 and BP11 highest labeling was invariably observed for putative TM segments 3, 5, 8 and 11, indicating that these four TM segments are involved in the initial substrate binding event in nonenergized P-gp i.e. in the absence of ATP. For the quaternary compound GPV708, which carries a permanent positive charge at the nitrogen atom, labelling was only observed in putative TM3 and 11, while TM segments 5 and 8 were not labelled (Fig.

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

2F). This was an initial indication for a pair wise contribution of TMs3/11 and 5/8 to substrate binding.

Fig. 3 shows a composite representation of peak-area data for all six ligands. The average percentage of labelling in the respective TM segment was calculated as a percentage of labelling in all TM segments. Clearly, TM segments 3, 5, 8 and 11 were responsible for close to 70 percent of total labelling.

The probability of detection of labeled peptide fragments within a particular protein region might depend on the recovery of peptide fragments from the same protein region. In order to evaluate a potential influence of this variable on the photolabeling results, the expected retrieval of peptide fragments from individual TM segments was determined, based on *in silico* digests. Results are shown in Fig. 4. The experimentally determined average photolabeling (as shown in Fig. 3) was plotted as a function of the expected coverage of individual TM-segments. Strongly labeled TM-segments (closed symbols) showed broad overlap with poorly labeled TMs with respect to expected fragment density. Results were therefore not biased for the probability of retrieval of peptide fragments from different TM-segments.

Generation of a three dimensional atomic detail model of P-gp

A protein homology model of P-gp was generated to allow projection of labelling data on a 3D model of the protein as described in detail in the methods section. The crystal structure of dimeric Vc-MsbA became available recently (Chang, 2003). This structure shows higher resolution than the structure of *E. coli* MsbA, which has been used as a template for a previous modeling study (Stenham et al. 2003). Vc-MsbA is an essential lipid transporter, which shares structural and functional homology with P-gp as well as LmrA, an ATP driven multidrug efflux pump from *L. lactis* (Reuter et al., 2003). The initial P-gp model (model A) is shown as a stereo image in Fig. 5. The model predicts that the TM segments form a helical bundle, which lines a central water-filled chamber with access to the extracellular space. TMs 3, 5, and 6 as well as 9, 11 and 12 are predicted to have broad access to the central cavity, while the access of TM segments 2 and 8 seems restricted. The TMD:TMD interfaces involve contributions of TM segments 2, 3 and 11 on one side and 5, 8 and 9 on the other side of the aqueous pore.

In contrast to the template, which lines a central cavity that is narrower at the extracytoplasmic face, P-gp has been shown to be funnel shaped (Loo & Clarke, 2001, Rosenberg et al., 2001, 2003) with the TM segments being closer at their cytoplasmic ends than at the extracellular surface of the membrane. Therefore, the TMDs were repositioned manually by slightly moving the TMDs apart at the extracellular surface and narrowing them at the cytoplasmic face. This and subsequent recalculation of the TMD:NBD linker led to the model depicted in side and top view in Fig. 6. This model (model B) satisfied the majority of the distance constraints established by cysteine cross linking studies better than model A. Table 2 gives a synopsis of distance constraints obtained by site directed mutagenesis and subsequent cysteine cross linking experiments. The table lists the positions of mutated residues and distances in models A and B. These distances are similar to those reported for a P-gp model generated with a different template structure (*E. coli* MsbA) (Stenham et al., 2003).

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

Relating the substrate photoaffinity labeling information to the 3D atomic detail model of P-gp

As indicated above, putative TM segments 3, 5, 8 and 11 showed highest accessibility for substrate-type photoaffinity ligands with the peak scoring positions being M197 in TM3, A311 in TM5, T769 in TM8 and F951 in TM11. These amino acid residues were displayed in the model shown in Fig. 6 with their Connolly surfaces colored in yellow and red, respectively. This revealed a number of remarkable facts. Firstly, when looking at the stereo image of the side view of the model, one is able to appreciate that these residues are located at approximately the same height and predicted to be buried deep within the membrane (membrane boundaries indicated by horizontal dotted lines) approximately halfway between the inner and the outer membrane surface. Secondly, M197 in TM3 and F951 in TM11) are in close vicinity of each other, but located on different halves of P-gp (the latter are depicted in dark blue (N-terminal half) and cyan (C-terminal half), respectively). Thus helices 3 and 11 contribute to one of the two TMD:TMD interfaces. The peak scoring residues are shown in yellow in the stereo image. Both the side and the top view of the P-gp model demonstrate the close spatial vicinity of these residues, which show a C_α-carbon distance of less than 9Å. Thirdly, A311 in TM 5 and T769 in TM8 are located at the other TMD:TMD interface and have a C_α-carbon distance of less than 14Å. Thus labeling data suggest the presence of two substrate binding domains at two pseudo symmetrical TMD:TMD interfaces formed by TM segments 3 of the N-terminal TMD and TM11 of the C-terminal TMD and by TM5 of the N-terminal and TM8 of the C-terminal NBD. Interestingly, only one of these interfaces contains methionine residues (TM3 and 11), while the other (TM5 and TM8) is

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

methionine less. This indicates that labeling at the domain interfaces does not depend on the presence of methionine residues. In addition, the labeling is pseudo symmetrical, since substrate binding at the TMD:TMD interfaces occurs to TM3 of the N-terminal and TM5 of the C-terminal half (corresponding to TM11) but TM2 of the C-terminal TMD (corresponding to TM8) and TM5 of the N-terminal TMD. This rotational pseudo-symmetry is in agreement with the low resolution structure of P-gp obtained by electron cryo microscopy (Rosenberg et al., 2001, 2003) and with recent data by Loo, Bartlett and Clarke, showing that cross links between equivalent AA-positions in TMs2/11 and 5/8 are influenced differently by the same drug substrates (Loo et al., 2004b). In addition, ligand GPV708 only labels the interface formed by TMs3 and 11, while the other interface domain is not labelled.

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973

Discussion

The P-gp protein homology model generated in this study (Fig. 6) conforms to the majority of distance constraints obtained by previous systematic cross linking studies of double cysteine mutants (Table 2). In particular, the spatial proximity of TM helices 5/8 and 2/11 has been shown by two different groups (Stenham et al. 2003, Loo et al., 2004a,b). Stenham et al. replaced three consecutive amino acid residues in the extracellular halves of each, TM segments 2/11 and 5/8, to allow cross links independent of the relative rotational positioning of these helices towards each other and showed that cross links were indeed detected.

Loo, Bartlett and Clarke used double cysteine mutants in TM5/8 and 2/11 respectively (Loo et al., 2004a, b). In particular, cross linking between N296 and G774 and I299 and F770 as well as between I299 and G774 and G300 and F300 were observed when these residues were replaced by cysteines and subsequently cross linked with the zero-length cross linker copper phenanthroline. With respect to putative TMs 2 and 11 cross links were found for the double cysteine mutants V133/G939, C137C/A935C and L138C/A935C. As in other P-gp models (Seigneuret and Garnier-Suillerot, 2003, Stenham et al., 2003) the distance between helix 2 and 11 exceeds 7Å, the latter being the maximal distance for cross linking of residues with the zero-length cross linker copper phenanthroline. Despite this fact, the model depicted in Fig. 6 accurately predicts the spatial proximity of the helix pairs 5/8 and 2/11.

Loo, Bartlett and Clarke also generated one hundred TM3-TM11 double cysteine mutants in the inner half of these helices, but no cross links were detected with copper-phenanthroline. In consistence with these results the model in Fig. 6 predicts helices 3

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

and 11 to be closer at their extracytoplasmic half, while they are more distant at the cytoplasmic half.

A protein homology model must be viewed as an approximation of the structure of a protein rather than an accurate 3D structure and the exact positioning of TM segments 2/3/11 and 5/8/9 at the TMD:TMD interfaces might be somewhat different from that shown in our model. The symmetrical Vc-MsbA template necessarily yields a symmetrical model, though for P-gp an asymmetry has been documented experimentally (Rosenberg et al. 2001, Loo et al, 2004b). Nevertheless, the scarcity of available membrane protein structures (as of August 2004, 82 unique membrane proteins are structurally resolved at atomic level, none of the resolved structures however being an ATP-dependent multidrug transporter) makes these computational models particularly valuable as a starting point for the integration of photoaffinity labeling data and 3D structure. The models are also capable of guiding the design of mutant P-gp, to validate data obtained by cross linking studies and last not least to use these cross links as constraints in molecular dynamics simulations.

In the present study, photoaffinity labeling allowed the identification of an involvement of four TM helices, TMs 3, 5, 8 and 11 in the primary substrate binding event. These helices are located at the TMD:TMD interfaces and conformational changes at these interfaces might provide an explanation why such a large number of compounds can be adopted and transported by P-gp. A schematic top-view of P-gp is shown in Fig. 7. The amino and carboxy terminal halves of the transporter are symbolized by blue and cyan arches. Individual helices 1-6 and 7-12 are represented in the same color, but with darker shading. Relative positions and colors are identical to those in Fig. 6. The central

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

pore is depicted in black. The pairs of helices involved in substrate binding also refer to the model in Fig. 6 and are indicated by yellow and red numbering. Arrows indicate the putative path by which substrates access their binding sites.

Binding of substrates at interfaces has previously been shown for AcrB, a proton motive force driven multidrug efflux pump from *E. coli*, which has been co-crystallized in the presence of four substrates - ethidium, ciprofloxacine, dequalinium and rhodamine 6G (Yu et al. 2003). We recently showed, that the ATP dependent multidrug transporter LmrA from *L. lactis* also binds substrates at the TMD:TMD interfaces (Ecker et al. 2004). Given the fact that at least three different multidrug transporters bind substrates at domain interfaces suggests that this might be a common property of polyspecific drug efflux pumps. The fact that a number of peptide fragments were identified to contain more than one covalently bound photoaffinity ligand suggests, that similar to AcrB and other polyspecific proteins, more than one substrate molecule can be adopted in the binding domains at a time (see Neyfakh 2002 and Schumacher and Brennan 2003 for reviews). Thus different amino acid residues in the binding sites are likely to be able to substitute for each other to accommodate structurally diverse substrate molecules. Evidence for the simultaneous binding of two different drugs in the binding pocket of P-gp has previously been presented by Loo et al. (2003). Acceptance or exclusion of substrates on basis of defined physicochemical parameters such as π-π, hydrophobic and dipolar interactions (Chiba et al, 1996, Ecker et al., 1999, Seelig & Landwojtowicz 2000) together with the plasticity of the TMD:TMD interfaces might be the underlying principle of multispecificity. For the regulatory proteins BmrR of *Bacillus subtilis* and QacR a repressor from *Staphylococcus aureus*, x-ray studies in complex with ligands have

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

demonstrated binding in rather large binding pockets (Schumacher and Brennan, 2003). The binding was mediated mostly by hydrophobic interactions with key electrostatic contributions from glutamic acid residues. The latter can be ruled out for substrate P-gp interaction, since TM segments do not contain charged amino acid residues. However, dipolar interactions might substitute for electrostatic interactions. In the apolar environment of the membrane stable hydrogen bonds can be formed and ligand based design has shown that hydrogen bonding properties contribute to the binding of propafenone analogs (Chiba et al., 1996) and other substance groups (Seelig & Landwojtowicz 2000). Interestingly, the protonable nitrogen atom in propafenones contributes to biological activity by its hydrogen bonding and not by its electrostatic properties (Ecker et al, 1999).

For multispecific proteins loose binding pockets are able to accommodate more than one ligand at the same time, whereby different orientations for one and the same ligand are possible. Our results indicate that this also applies for P-gp. A number of peptide fragments contain two covalently bound ligands, suggesting the simultaneous presence of more than one substrate molecule in the binding pocket. This necessitates the notion of different but coexisting substrate orientations.

The random nature of ATP-hydrolysis in one of two nucleotide binding sites has been shown (Ambudkar et a., 2003, Al-Shawi 2003). If substrate binding also follows random characteristics is unclear. Our findings are compatible with both, binding of substrates to one P-gp molecule at two sites in random manner, or simultaneous binding at both binding domains. Interestingly, ligand GPV708 only labels the TM3/11interface

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

and thus clearly binds in nonrandom manner providing a proof of principle for a pair wise contribution of TM-segments 3/11 and 5/8 to two binding domains.

In conclusion, the present study provides evidence that substrates access the central cavity of P-gp *via* the TMD:TMD interfaces and that these interfaces are asymmetric with respect to the contribution of TM-helices in the two halves of P-gp. Photoligands used in this study combine the advantage of well defined quantitative structure activity relationships and an unmodified core structure as compared to propafenone. This motivated sole use of these compounds in the present study. Nevertheless, heterologous displacement studies with known P-gp substrates indicate a general nature of our results extending to other substance classes. Future studies will be directed at identifying possible changes in the labeling pattern at different steps of the catalytic cycle of P-gp. These studies might allow the evolution of a concept which links P-gp structure to the dynamics of the transport process.

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973

Acknowledgements

The baculovirus expression vector pVL941-MDR1/A was generously provided by Dr. M. M. Gottesman (National Cancer Institute, Bethesda, MD). The technical assistance of F.C. and N.C. during July 2004 is highly appreciated.

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

References

Al-Shawi MK, Polar MK, Omote H, Figler RA (2003) Transition state analysis of the coupling of drug transport to ATP hydrolysis by P-glycoprotein. *J Biol Chem* **278**:52629-40.

Ambudkar SV, Kimchi-Sarfaty C, Sauna ZE and Gottesman MM (2003) P-glycoprotein: from genomics to mechanism. *Oncogene* **22**:7468-7485.

Campbell JD, Biggin PC, Baaden M and Sansom MS (2003) Extending the structure of an ABC transporter to atomic resolution: modeling and simulation studies of MsbA. *Biochemistry* **42**:3666-73.

Chang G (2003) Structure of MsbA from Vibrio cholera: a multidrug resistance ABC transporter homolog in a closed conformation. *J Mol Biol* **330**:419-30.

Chiba P, Burghofer S, Richter E, Tell B, Moser A and Ecker G (1995) Synthesis, pharmacologic activity, and structure-activity relationships of a series of propafenone-related modulators of multidrug resistance. *J Med Chem* **38**:2789-93.

Chiba P, Ecker G, Schmid D, Drach J, Tell B, Goldenberg S and Gekeler V (1996) Structural requirements for activity of propafenone-type modulators in P-glycoprotein-mediated multidrug resistance. *Mol Pharmacol* **96**: 1122-30.

Clauser KR, Baker PR, and Burlingame AL (1999) Role of accurate mass measurements (+/-10ppm) in protein identification strategies emplying MS or MS/MS and database searching. *Anal Chem* **71**:2871-2882.

Dorman G, and Prestwich GD (2000) Using photolabile ligands in drug discovery and development. *Trends Biotechnol* **18**:64-77.

Ecker G, Chiba P, Hitzler M, Schmid D, Visser K, Cordes HP, Csollei J, Seydel JK, and Schaper KJ (1996) Structure-activity relationship studies on benzofurane analogs of propafenone-type modulators of tumor cell multidrug resistance. *J Med Chem* **39**:4767-4774.

Ecker G, Huber M, Schmid D, and Chiba P (1999) The importance of a nitrogen atom in modulators of multidrug resistance. *Mol Pharmacol* **56**:791-796.

Ecker G, Csaszar E, Kopp S, Plagens B, Holzer W, Ernst W and Chiba P (2002) Identification of ligand-binding regions of P-glycoprotein by activated-pharmacophore photoaffinity laebling and matrix -assisted laser desorption/ionization-time of flight mass spectrometry. *Mol Pharmacol* **61**:637-48.

Ecker GF, Pleban K, Kopp S, Csaszar E, Poelarends GJ, Putman M, Kaiser D, Konings WN and Chiba P (2004) A three dimensional model for the substrate binding domain of the mutlidrug ABC transporter LmrA. *Mol Pharmacol* Aug 10 (epub)

Germann UA, Willingham MC, Pastan I, and Gottesman MM (1990) Expression of the human multidrug transporter in insect cells by a recombinant baculovirus. *Biochemistry* **29**:2295-2303.

Hafkemeyer P, Dey S, Ambudkar SV, Hrycyna CA, Pastan I, Gottesman MM. (1998) Contribution to substrate specificity and transport of nonconserved residues in transmembrane domain 12 of human P-glycoprotein. *Biochemistry* **37**:16400-9.

Humphrey W, Dalke A, Schulten K. (1996) VMD: visual molecular dynamics. *J Mol Graph* **14**:33-38, 27-28.

Jaysinghe S, Hristova K, and White SH (2000) http://blanco.biomol.uci.edu/mpex.

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

Loo TW and Clarke DM (1997) Drug-stimulated ATPase activity of human P-glycoprotein requires movement between transmembrane segments 6 and 12. *J Biol Chem* **272**:20986-89.

Loo TW and Clarke DM (2000) The packing of the transmembrane segments of human multidrug resistance P-glycoprotein is revealed by disulfide cross-linking analysis. *J Biol Chem* **275**:5253-56.

Loo TW and Clarke DM (2001) Determining the dimensions of the drug-binding domain of human P-glycoprotein using thiol cross-linking compounds as molecular rulers. *J Biol Chem* **276**:36877-80.

Loo TW and Clarke DM (2002) Vanadate trapping of nucleotide at the ATP-binding sites of human multidrug resistance P-glycoprotein exposes different residues to the drug-binding site. *Proc Natl Acad Sci* **99**:3511-16.

Loo TW, Bartlett MC and Clarke DM (2003) Simultaneous binding of two different drugs in the binding pocket of the human multidrug resistance P-glycoprotein. J Biol Chem **278**:39706-10.

Loo TW, Bartlett MC and Clarke DM (2004a) Disulfide cross-linking analysis shows that transmembrane segments 5 and 8 of human P-glycoprotein are close together on the cytoplasmic side of the membrane. *J Biol Chem* **279**:7692-97.

Loo TW, Barlett MC and Clarke DM (2004b) Val133 and Cys137 in transmembrane segment 2 are close to Arg935 and Gly939 in transmembrane segment 11 of human P-glycoprotein. *J Biol Chem* **279**:18232-38.

Neyfakh AA (2002) Mystery of multidrug transporters: the answer can be simple. *Mol Microbiol* **44**:1123-30.

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

Peer M, Csaszar E, Vorlaufer E, Kopp S and Chiba P (2004) Photoaffinity labeling of P-glycoprotein. *Minireviews in Medicinal Chemistry*, in press

Reuter G, Janvilisri T, Venter H, Shahi S, Balakrishnan L, van Veen HW (2003) The ATP binding cassette multidrug transporter LmrA and lipid transporter MsbA have overlapping substrate specificities. *J Biol Chem* **278**:35193-8.

Rihakova L, Deraet M, Auger-Messier M, Perodin J, Boucard AA, Guilemette G., Leduc R, Lavigne P, and Escher E (2002) Methionine proximity assay, a novel method for exploring peptide ligand-receptor interaction. *J Recept Signal Transduct Res* **22**:297-313.

Rosenberg MF, Velarde G, Ford RC, Martin C, Berridge G, Kerr ID, Callaghan R, Schmidlin A, Wooding C, Linton KJ and Higgins CF (2001) Repacking of the transmembrane domains of P-glycoprotein during the transport ATPase cycle. *EMBO J* **20**:5615-25.

Rosenberg MF, Kamis AB, Callaghan R, Higgins CF and Ford RC (2003) Three-dimensional structures of the mammalian multidrug resistance P-glycoprotein demonstrate major conformational changes in the transmembrane domains upon nucleotide binding. *J Biol Chem* **278**:8294-9.

Sauna ZE and Ambudkar SV (2000) Evidence for a requirement for ATP hydrolysis at two distinct steps during a single turnover of the catalytic cycle of human P-glycoprotein. *Proc Natl Acad Sci* **97**:2515-20.

Sauna ZE and Ambudkar SV (2001) Characterization of the catalytic cycle of ATP hydrolysis by human P-glycoprotein. *J Biol Chem* **276**:11653-61.

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

Schmid D, Ecker G, Kopp S, Hitzler M, and Chiba P (1999) Structure-activity relationship studies of propafenone analogs based on P-glycoprotein ATPase activity measurements. *Biochem Pharmacol* **58**:1447-56.

Schumacher MA, Brennan RG (2003) Deciphering the molecular basis of multidrug recognition: crystal structures of the Staphylococcus aureus multidrug binding transcription regulator QacR. *Res Microbiol* **154**:69-77.

Seelig A, Landwojtowicz E (2000) Structure-activity relationship of P-glycoprotein substrates and modifiers. *Eur J Pharm Sci* **12**(1):31-40.

Seigneuret M, Garnier-Suillerot A (2003) A structural model for the open conformation of the mdr1 P-glycoprotein based on the MsbA crystal structure. *J Biol Chem* **278**:30115-24.

Senior AE, Gadsby DC (1997) ATP hydrolysis cycles and mechanism in P-glycoprotein and CFTR. *Semin Cancer Biol.* **8**:143-50.

Stenham DR, Campbell JD, Sansom MS, Higgins C F, Kerr ID, Linton, KJ (2003) An atomic detail model for the human ATP binding cassette transporter P-glycoprotein derived from disulfide cross-linking and homology modeling. *FASEB J* **17**: 2287-2289.

Wahlstrom JL, Randall MA, Lawson JD, Lyons DE, Siems WF, Crouch GJ, Barr R, Facemyer KC and Cremo CR (2002) Structural model of the regulatory domain of smooth muscle heavy meromyosin. *J Biol Chem* **278**:5123-31.

Yu EW, McDermott G, Zgurskaya HI, Nikaido H, Koshland DE Jr. (2003) Structural basis of multiple drug-binding capacity of the AcrB multidrug efflux pump. *Science* **300**:976-80.

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

Yu EW, Aires JR, and Nikaido H (2002) AcrB multidrug efflux pump of Escherichia coli: composite substrate-binding cavity of exceptional flexibility generates its extremely wide substrate specificity. *J Bacteriology* **185**:5657-5664.

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

Footnotes

This work was supported by grants from the Austrian Science Fund (grant 17014-B11) and the Austrian National Bank (grant 10654). The baculovirus expression vector pVL941-MDR1/A was generously provided by Dr. M. M. Gottesman (National Cancer Institute, Bethesda, MD). The technical assistance of F.C. and N.C. during July 2004 is highly appreciated.

Author to whom proofs should be sent:

Peter Chiba, Institute of Medical Chemistry, Medical University of Vienna, Waehringerstrasse 10, A-1090 Vienna, Austria

Phone: +43 1 4277 60806, Fax: +43 1 4277 60889

E-mail: peter.chiba@univie.ac.at

The coordinates of the P-gp model are available from the authors upon request.

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

Legends for Figures

Figure 1

Frequency distribution analysis of labeled peptide fragments. The number of labeled fragments in which each particular amino acid residue in the primary sequence of P-gp is found, is plotted as a function of this amino acid position. The position of putative TM segments are shown in magenta, the remainder of the sequence is in black. A schematic representation below the graph allows easier orientation. TMD1 and 2 represent the N-terminal and C-terminal transmembrane domains, the motor domains NBD1 and 2 are shown in blue. The linker region is indicated in grey and the C-terminal hexa-his-tag is depicted in black. The positions of methionine residues are highlighted by red symbols (x). Highest labeling is observed for putative TMs 3, 5, 8 and 11. In addition the loop region connecting TM9 and 10 (extracellular loop (ECL) 5) has a high labeling score.

Figure 2

Cumulative labeling frequency expressed as the area under curve for each individual ligand and TM-segment. In order to exclude accessory surface labeling, the first three extra- and intracytoplasmic amino acid residues of each TM segments were excluded from calculations. Ordinate numbers represent the cumulative number of amino acid residues found in fragments localizing to the respective TM-segment. (A) Ligand GPV51, (B) ligand GPV317, (C) ligand GPV319, (D) ligand GPV442, (E) ligand BP11, (F) ligand GPV708. For structures of the ligands see Table 1. Invariably, highest scores

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

were obtained for TMs 3, 5, 8 and 11 with ligands GPV51, 317, 319, 442 and BP11. GPV708 (Fig. 5F) only labeled TM segments 3 and 11.

Figure 3

Average relative labeling frequencies of the TM-segments obtained with six photoligands. Mean and standard deviations are given. Since ligand GPV708 did not label TMs 5 and 8, it was excluded from calculations for these two TM helices. As for the individual ligands depicted in Fig. 2, highest scores were obtained for TMs 3, 5, 8 and 11 and differences in labeling between each of these four TM-segments and all other TM-segments were statistically significant with $p < 0.001$.

Figure 4

Theoretical peptide fragment density for the individual TM segments of P-gp. The prediction is based on an incomplete *in silico* digest of P-gp with chymotrypsin. Values were obtained by counting the number of amino acid residues in fragments overlapping with each individual TM segment. Expected fragment densities showed broad overlap for TM segments with high labeling (TMs 3, 5, 8, and 11, shown as closed symbols) and those with poor labeling (helices 1, 2, 4, 6, 7, 9, 10 and 12, shown as open symbols). Data indicate that photolabeling results are not biased for differences in recovery of peptide fragments from different TM segments.

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

Figure 5

Stereo image of a P-gp homology model (model A) generated on basis of the x-ray structure of Vc-MsbA. The nucleotide binding domain was modeled based on the structure of TAP1 as described in the Methods section. TM-segments form a helical bundle which lines a central pore that connects to the extracellular environment. TMs 1 and 7 can easily be identified by a short α helix, which lies just N-terminal and in a 90° angle relative to these TM-segments. The short helix is almost parallel to the membrane and defines the approximate location of the inner membrane-water interface. The two dotted horizontal lines give the approximate position of the membrane boundaries.

Figure 6

Stereo image of a side and top view of a refined P-gp model (model B). This model was generated to account for experimental evidence that the central pore is wider at the extracytoplasmic face of the membrane, than at the cytoplasmic face. The N-terminal half of P-gp is depicted in dark blue, the C-terminal half is in cyan. Residues with highest labeling scores are depicted in yellow (residues M197 in TM3 and F951 in TM11) and red (A311 in TM5 and T769 in TM8), respectively. These residues are located at the two TMD:TMD interfaces. The two residues depicted in yellow M197, F951) have C_{α}-carbon distances of less than 9Å, while the residues depicted in red (A311 and T769) are less than 14Å apart. The side view indicates that these four residues are located at similar depth within the membrane.

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

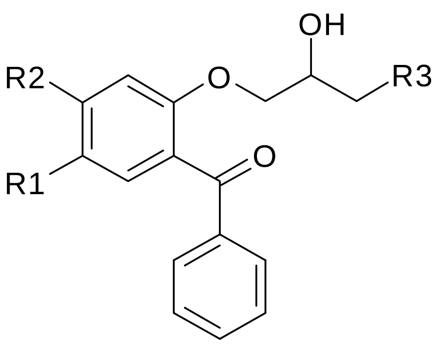
Figure 7

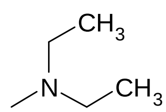
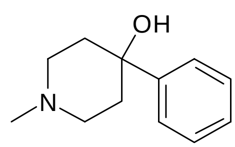
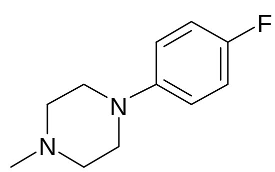
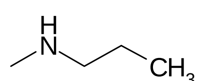
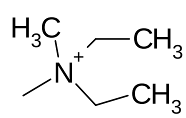
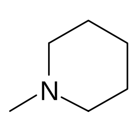
Schematic representation of a top view of P-gp. The N- and C-terminal halves are symbolized by dark blue and cyan arches, respectively. Circles and numbers indicate the TM-segments. Labeled TMs are numbered in yellow and red. These helices are located at the two TMD:TMD interfaces. The red and yellow arrows symbolize the path along which substrates are considered to access their binding sites. See the Discussion section for details.

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

Table 1

Chemical structure and exact mass of benzophenone-type photoligands



Compound	R1	R2	R3	exact mass
GPV 51	H	H		327.183
GPV 317	H	H		431.210
GPV 319	H	H		434.200
GPV 442	$-CH_3$	H		327.183
GPV 708	H	H		342.210
BP 11	H	$-OCH_3$		369.194

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

Table 2

Distances of residues in the models that have previously been shown to be cross linked in double cysteine mutants

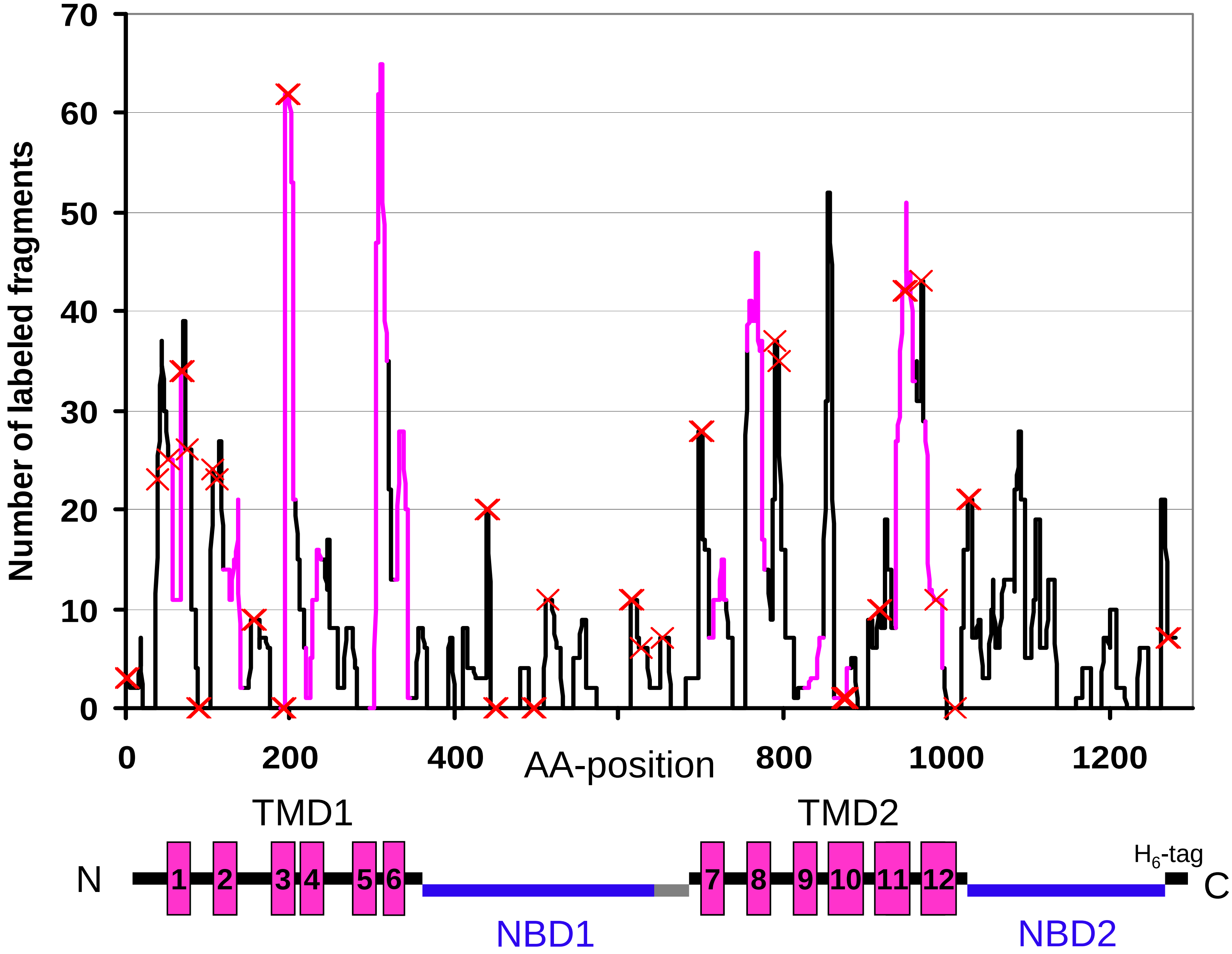
TM-helix a	Residue a	TM-helix b	Residue b	X-link distance (Å)	distance model A (Å)	distance model B (Å)	reference
2	Y117-S119	**11**	G955-F957	12	12	17	Stenham et al. 2003
2	V133[1], C137[2]	**11**	A935[1], G939[2]	7	18-26	15-25	Loo et al. 2004b
4	S222	**10**	I868, G872	9-25	40-41	40-41	Loo & Clarke 2001
4	L227, V231-A233, I235, L236	**11**	S993	7	28-31	22-25	Loo & Clarke 2000
5	G317-T319	**8**	N753-F755	11	12	16	Stenham et al. 2003
5	N296[1], I299[2,3], G300[4]	**8**	G774[1,3], F770[2,4]	7	15-19	15-21	Loo et al. 2004a
5	I306	**10**	I868, G872	13-25	24-25	20-21	Loo & Clarke 2001
5	I306	**11**	T945	13-25	28	22	Loo & Clarke 2001
5	I306[1], A295[2], I299[2]	**12**	V982[1], G984[1], S993[2]	7-25	20-29	10-22	Loo & Clarke 2000, 2001
6	L339 P350	**10**	I868, G872 V874-M876	7-25	28-33	23-31	Loo & Clarke 2000, 2001
6	L339 P350	**11**	F942, T945 G939	7-25	20-29	9-22	Loo & Clarke 2000, 2001
6	L339[1], F343[2], G346[3], P350[4]	**12**	V982[1], A985[1], M986[2], G989[3], S993[4]	7-25	18-32	10-27	Loo & Clarke 2000, 2001

TM-helix a and TM-helix b denote the pairs of putative transmembrane α-helices in which cross-links between two cysteines introduced at residue a and residue b could be obtained. The source reference is given in the rightmost column. In order to allow a more convenient identification of the mutated residues, they are denoted with their original (non-mutated) one-letter amino acid code followed by the position in the primary

Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973
This article has not been copyedited and formatted. The final version may differ from this version.

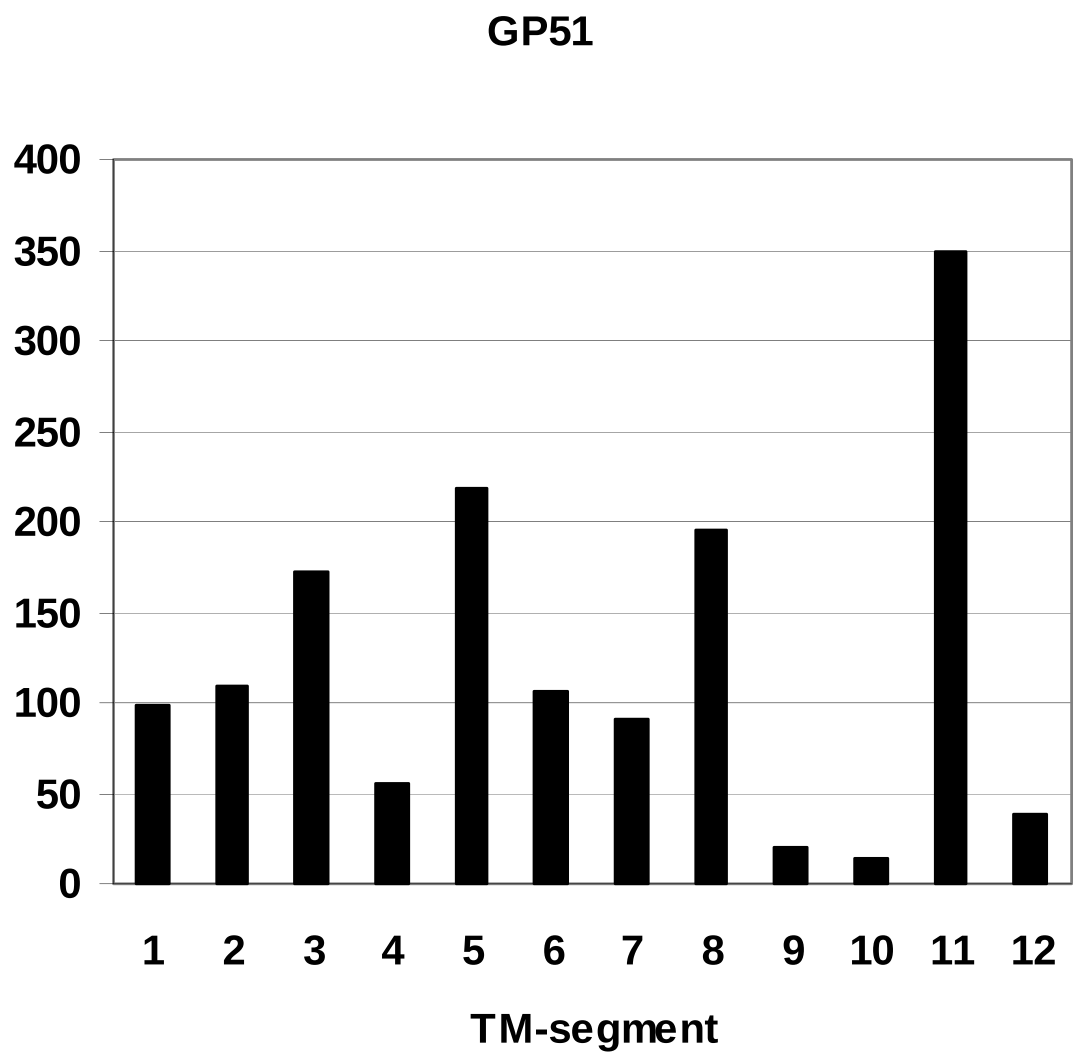
sequence. All distances are minimum distances given in Å and are measured between the C_α atoms of the specified residues. In instances where more than one pair of residues has been cross-linked, the superscripts 1-4 are used to identify these pairs and the minimum and maximum distances are given. In case of the triple cysteine mutants (Stenham et al., 2003) the minimum distance is given.

Figure 1



Number of labeled fragments
70
60
50
40
30
20
10
0
0
200
400
AA-position
800
1000
1200
TMD1
TMD2
N
1
2
3
4
5
6
7
8
9
10
11
12
H_6-tag
C
NBD1
NBD2

Figure 2A



Molecular Pharmacology Fast Forward. Published on October 27, 2004 as DOI: 10.1124/mol.104.006973

Figure 2B

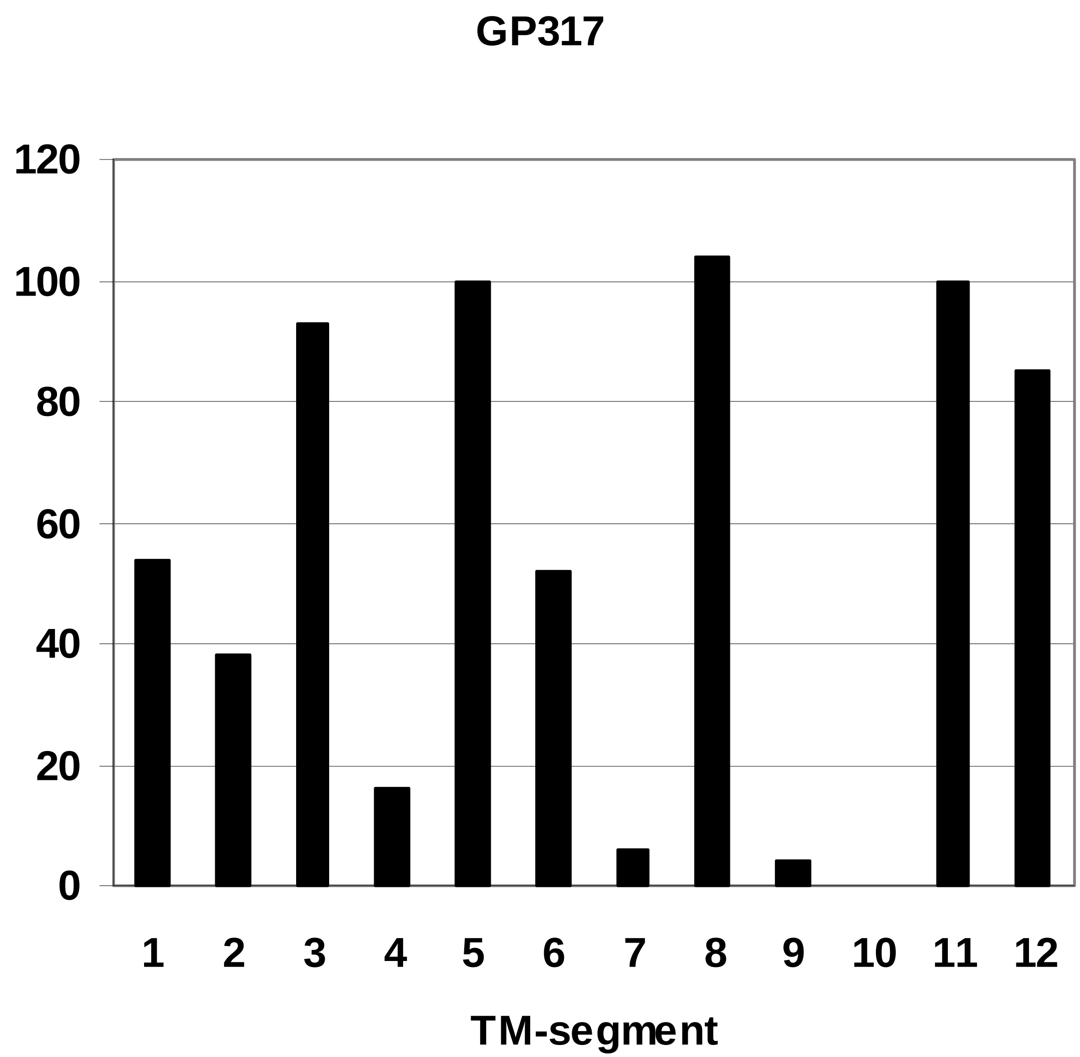


Figure 2C

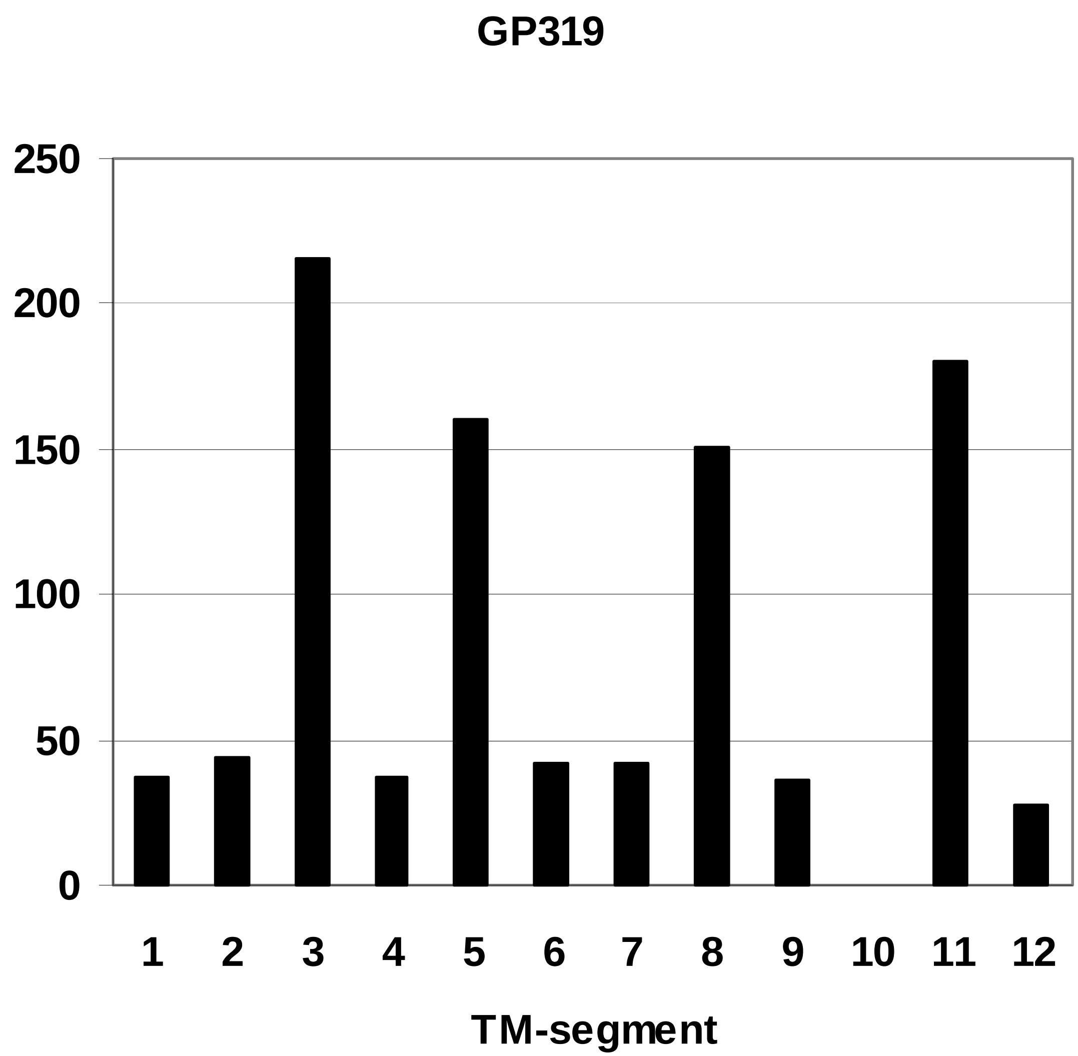


Figure 2D

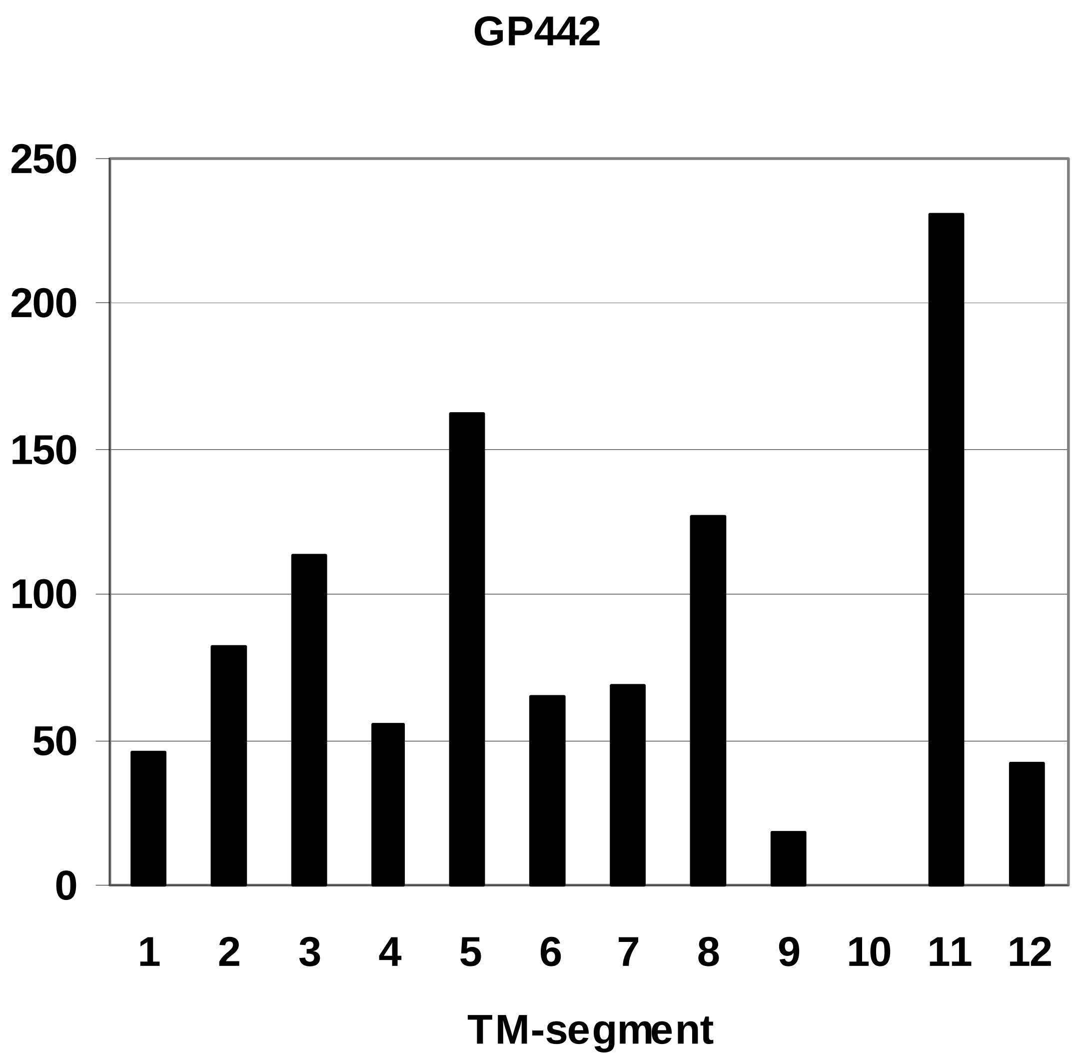


Figure 2E

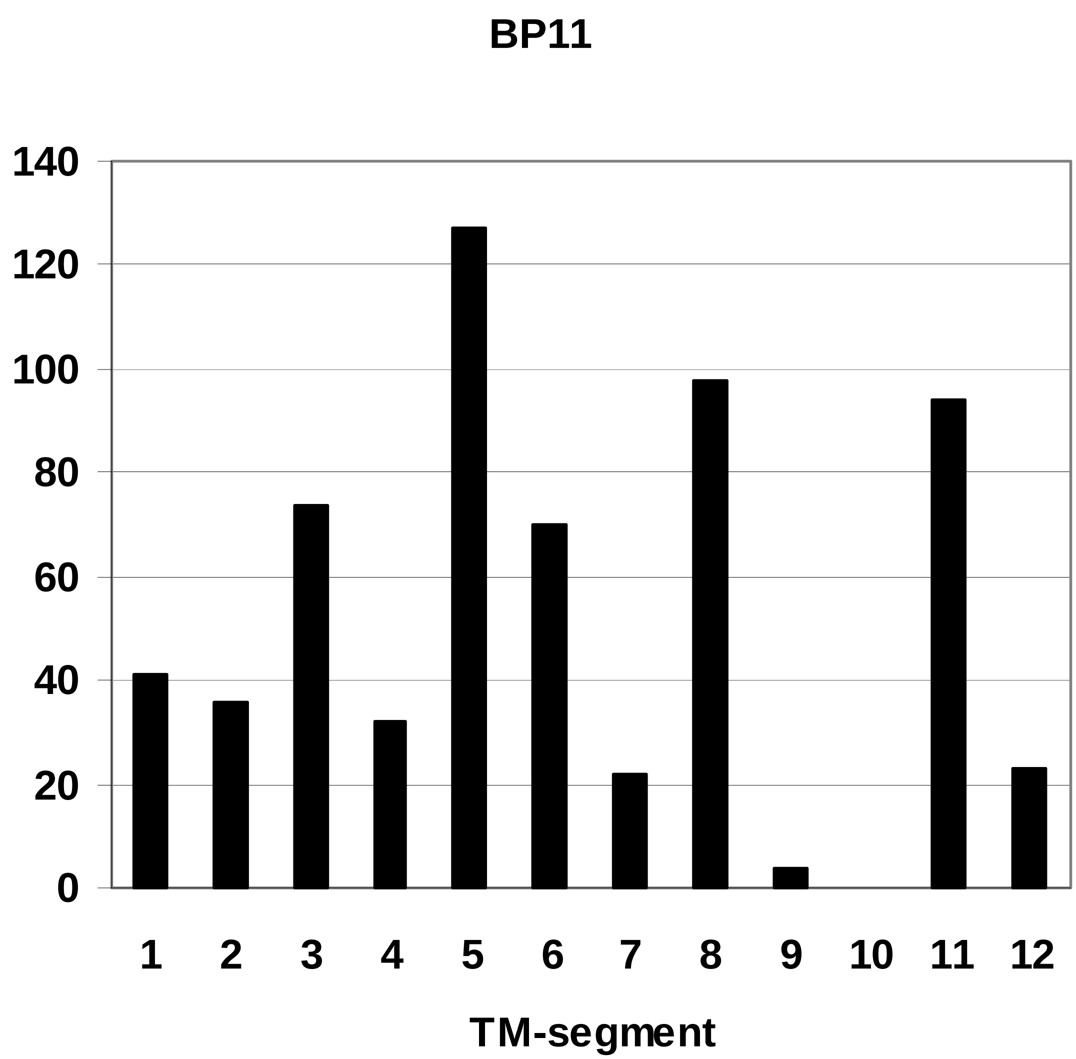


Figure 2F

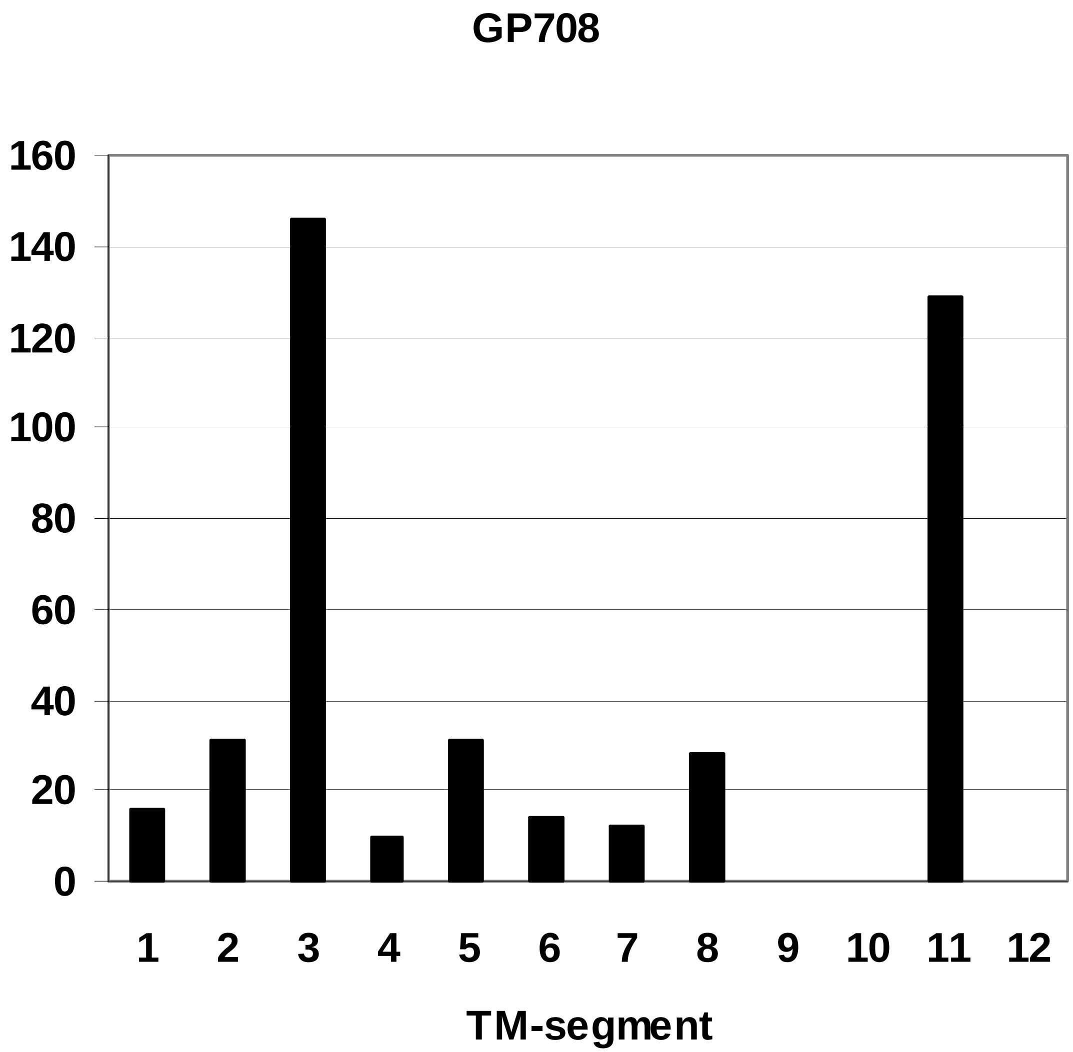
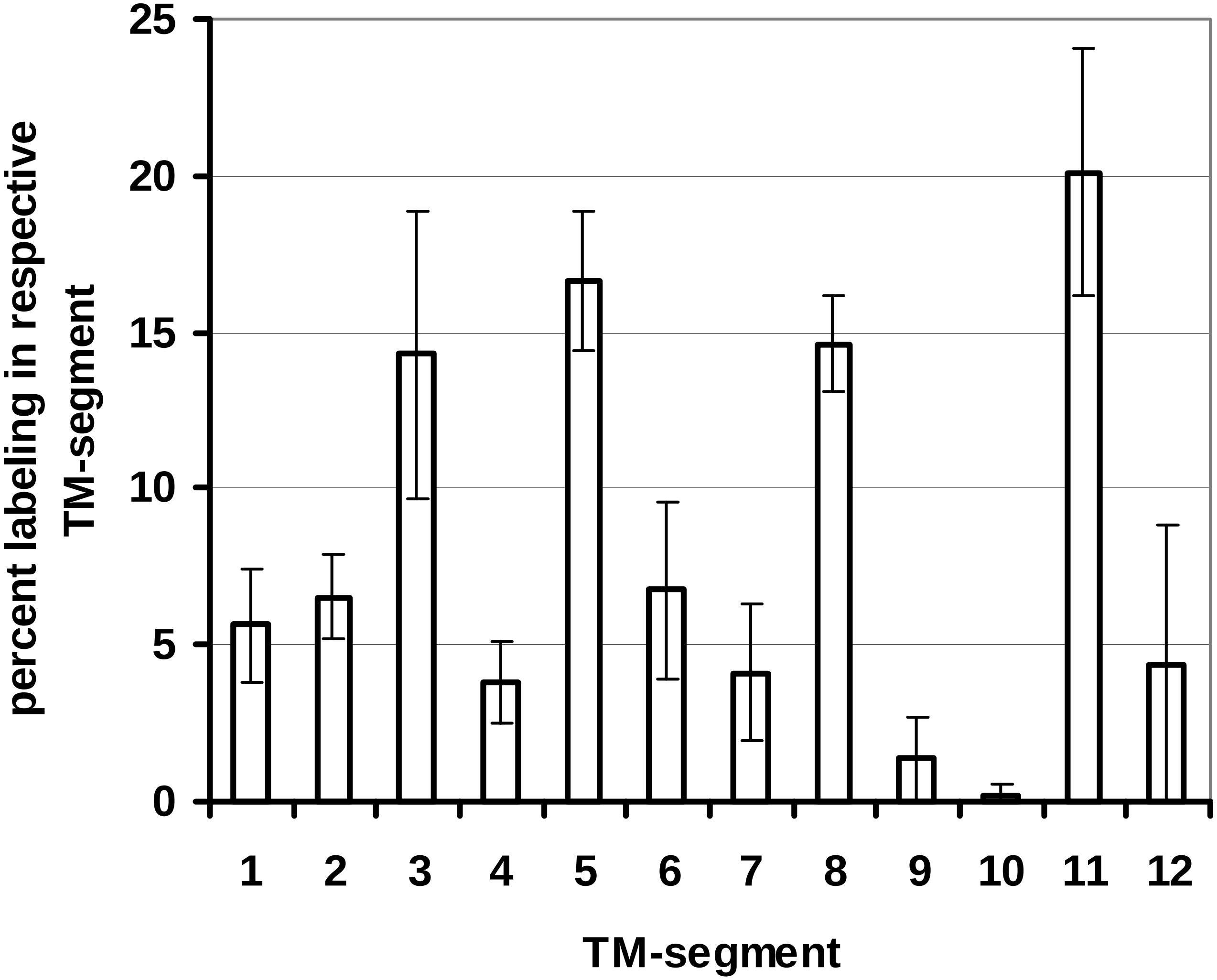
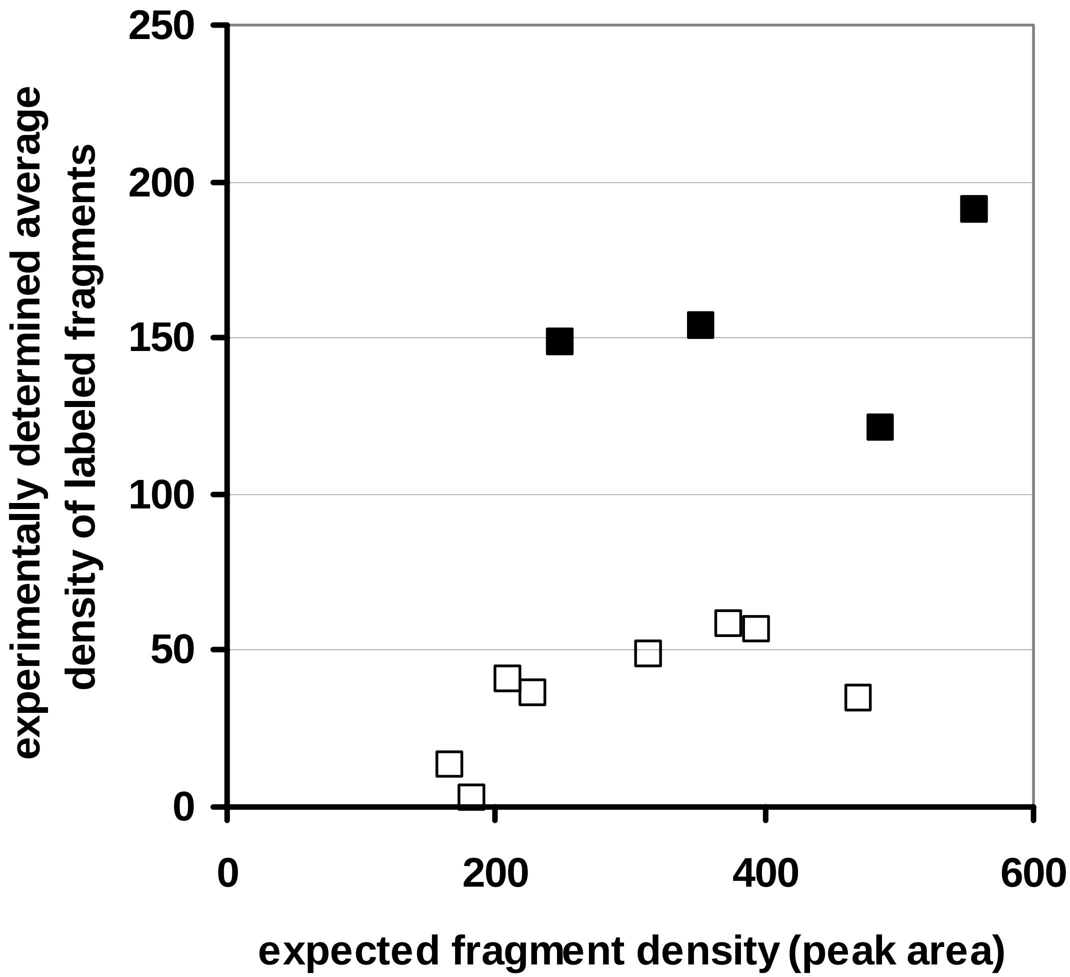

GP708
160
140
120
100
80
60
40
20
0
1
2
3
4
5
6
7
8
9
10
11
12
TM-segment

Figure 3



25
20
15
10
5
0
percent labeling in respective TM-segment
1
2
3
4
5
6
7
8
9
10
11
12
TM-segment

Figure 4



250
200
150
100
50
0
0
200
400
600
experimentally determined average density of labeled fragments
expected fragment density (peak area)

Figure 5

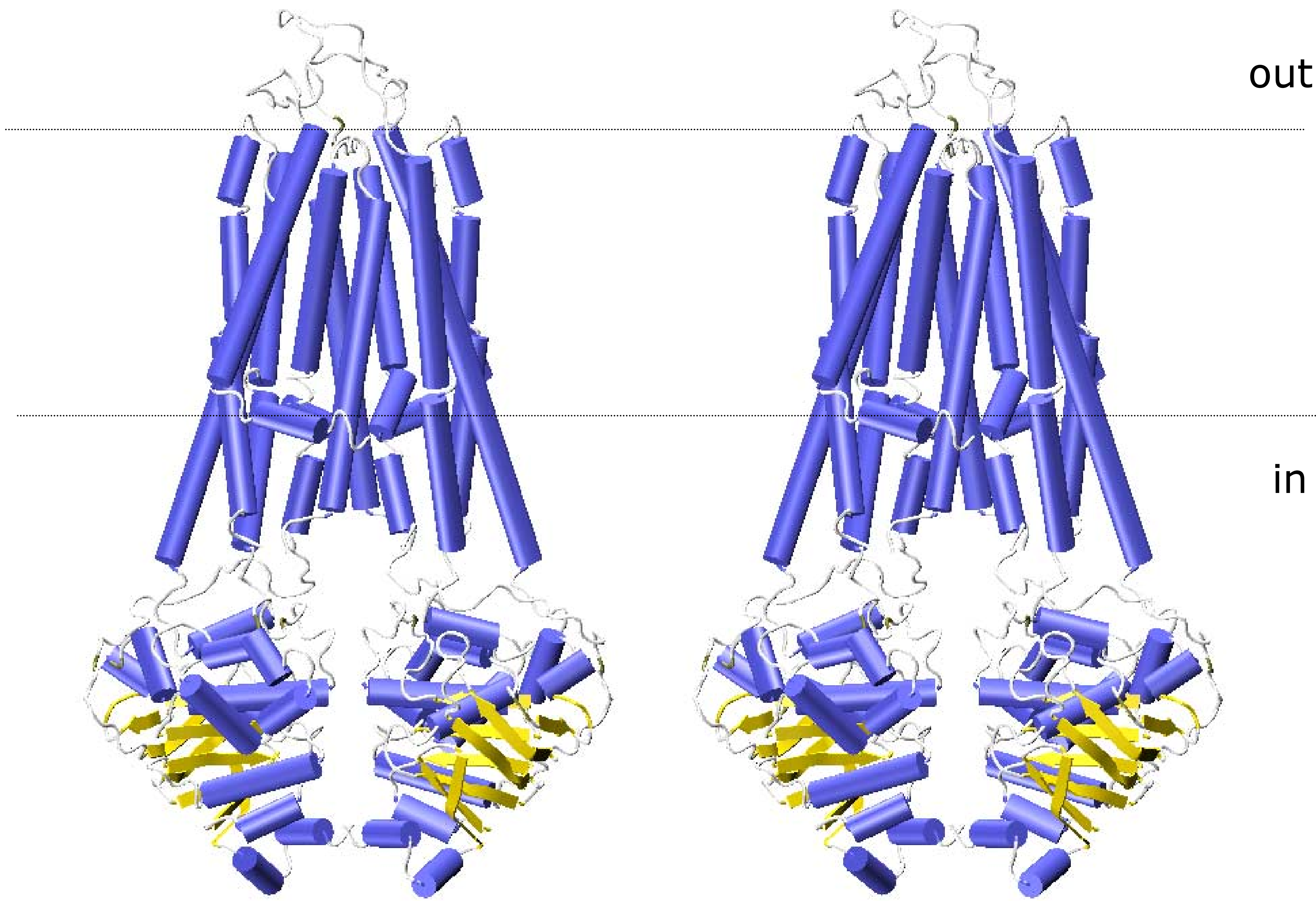


Figure 6

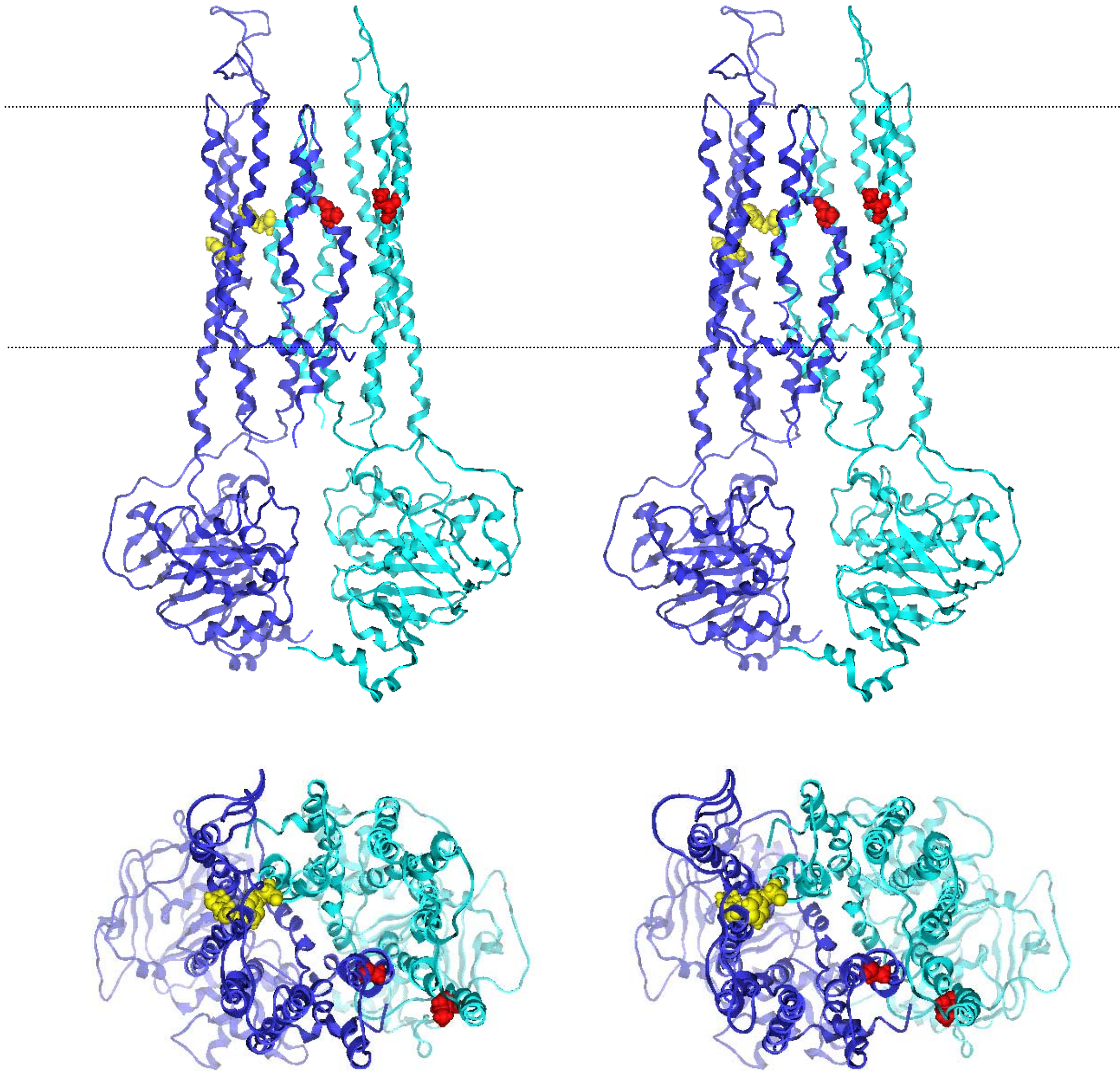


Figure 7